SECOND EDITION Changing of the second edition

Robin Smith





Chemical Process Design and Integration

Chemical Process Design and Integration

Second Edition

Robin Smith

School of Chemical Engineering and Analytical Science, The University of Manchester, UK

WILEY

This edition first published 2016 © 2016 John Wiley & Sons, Ltd

Registered office

John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, United Kingdom

For details of our global editorial offices, for customer services and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com.

The right of the author to be identified as the author of this work has been asserted in accordance with the Copyright, Designs and Patents Act 1988.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, except as permitted by the UK Copyright, Designs and Patents Act 1988, without the prior permission of the publisher.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Designations used by companies to distinguish their products are often claimed as trademarks. All brand names and product names used in this book are trade names, service marks, trademarks or registered trademarks of their respective owners. The publisher is not associated with any product or vendor mentioned in this book.

Limit of Liability/Disclaimer of Warranty: While the publisher and author have used their best efforts in preparing this book, they make no representations or warranties with respect to the accuracy or completeness of the contents of this book and specifically disclaim any implied warranties of merchantability or fitness for a particular purpose. It is sold on the understanding that the publisher is not engaged in rendering professional services and neither the publisher nor the author shall be liable for damages arising herefrom. If professional advice or other expert assistance is required, the services of a competent professional should be sought. The advice and strategies contained herein may not be suitable for every situation. In view of ongoing research, equipment modifications, changes in governmental regulations, and the constant flow of information relating to the use of experimental reagents, equipment, and devices, the reader is urged to review and evaluate the information provided in the package insert or instructions for each chemical, piece of equipment, reagent, or device for, among other things, any changes in the instructions or indication of usage and for added warnings and precautions. The fact that an organization or Website is referred to in this work as a citation and/or a potential source of further information does not mean that the author or the publisher endorses the information the organization or Website may provide or recommendations it may make. Further, readers should be aware that Internet Websites listed in this work may have changed or disappeared between when this work was written and when it is read. No warranty may be created or extended by any promotional statements for this work. Neither the publisher nor the author shall be liable for any damages arising herefrom.

Library of Congress Cataloging-in-Publication Data

Names: Smith, Robin (Chemical engineer)

Title: Chemical process design and integration / Robin Smith.

Description: Second edition. | Chichester, West Sussex, United Kingdom : John Wiley & Sons, Inc., 2016. | Includes index. | Description based on print version record and CIP data provided by publisher; resource not viewed.

Identifiers: LCCN 2015034820 (print) | LCCN 2015032671 (ebook) | ISBN 9781118699089 (ePub) | ISBN 9781118699096 (Adobe PDF) | ISBN 9781119990147 (hardback) |

ISBN 9781119990130 (paper)

Subjects: LCSH: Chemical processes. | BISAC: TECHNOLOGY & ENGINEERING / Chemical & Biochemical.

Classification: LCC TP155.7 (print) | LCC TP155.7 .S573 2016 (ebook) | DDC 660/.28–dc23 LC record available at http://lccn.loc.gov/2015034820

A catalogue record for this book is available from the British Library.

ISBN: 9781119990147

Set in 9.5/12 pt TimesLTStd-Roman by Thomson Digital, Noida, India

1 2016

To the next generation George, Oliver, Ava and Freya

Contents CONTENTS

Preface xiii Acknowledgements xv Nomenclature xvii

1 The Nature of Chemical Process Design and Integration 1

- 1.1 Chemical Products 1
- 1.2 Formulation of Design Problems 3
- **1.3** Synthesis and Simulation 4
- 1.4 The Hierarchy of Chemical Process Design and Integration 6
- 1.5 Continuous and Batch Processes 8
- 1.6 New Design and Retrofit 11
- 1.7 Reliability, Availability and Maintainability 11
- **1.8** Process Control 12
- 1.9 Approaches to Chemical Process Design and Integration 13
- 1.10 The Nature of Chemical Process Design and Integration – Summary 16 References 17

2 Process Economics 19

- 2.1 The Role of Process Economics 19
- 2.2 Capital Cost for New Design 19
- 2.3 Capital Cost for Retrofit 25
- 2.4 Annualized Capital Cost 26
- 2.5 Operating Cost 27
- 2.6 Simple Economic Criteria 30
- 2.7 Project Cash Flow and Economic Evaluation 31
- 2.8 Investment Criteria 33
- 2.9 Process Economics–Summary 34
- 2.10 Exercises 34 References 36

3 Optimization 37

- **3.1** Objective Functions 37
- **3.2** Single-Variable Optimization 40
- **3.3** Multivariable Optimization 42
- 3.4 Constrained Optimization 45
- 3.5 Linear Programming 47
- **3.6** Nonlinear Programming 49
- **3.7** Structural Optimization 50
- **3.8** Solution of Equations Using Optimization 54
- **3.9** The Search for Global Optimality 55

- **3.10** Optimization Summary 56
- 3.11 Exercises 56 References 58
- 4 Chemical Reactors I Reactor Performance 59
- 4.1 Reaction Path 59
- 4.2 Types of Reaction Systems 61
- 4.3 Measures of Reactor Performance 63
- 4.4 Rate of Reaction 64
- 4.5 Idealized Reactor Models 65
- 4.6 Choice of Idealized Reactor Model 73
- 4.7 Choice of Reactor Performance 76
- **4.8** Reactor Performance Summary 77
- 4.9 Exercises 78 References 79

5 Chemical Reactors II – Reactor Conditions 81

- 5.1 Reaction Equilibrium 81
- 5.2 Reactor Temperature 85
- **5.3** Reactor Pressure 92
- 5.4 Reactor Phase 93
- 5.5 Reactor Concentration 94
- 5.6 Biochemical Reactions 99
- 5.7 Catalysts 99
- 5.8 Reactor Conditions Summary 102
- 5.9 Exercises 103 References 105

6 Chemical Reactors III – Reactor Configuration 107

- 6.1 Temperature Control 107
- 6.2 Catalyst Degradation 111
- 6.3 Gas-Liquid and Liquid-Liquid Reactors 112
- 6.4 Reactor Configuration 116
- 6.5 Reactor Configuration For Heterogeneous Solid-Catalyzed Reactions 121
- 6.6 Reactor Configuration Summary 122
- 6.7 Exercises 122 References 123

7 Separation of Heterogeneous Mixtures 125

7.1 Homogeneous and Heterogeneous Separation 125

- 7.2 Settling and Sedimentation 126
- 7.3 Inertial and Centrifugal Separation 130
- 7.4 Electrostatic Precipitation 131
- 7.5 Filtration 133
- 7.6 Scrubbing 134
- 7.7 Flotation 135
- 7.8 Drying 136
- 7.9 Separation of Heterogeneous
- Mixtures Summary 137 7.10 Exercises 137
- References 138

8 Separation of Homogeneous Fluid Mixtures I – Distillation 139

- 8.1 Vapor–Liquid Equilibrium 139
- 8.2 Calculation of Vapor-Liquid Equilibrium 141
- 8.3 Single-Stage Separation 146
- 8.4 Distillation 146
- 8.5 Binary Distillation 150
- 8.6 Total and Minimum Reflux Conditions for Multicomponent Mixtures 155
- 8.7 Finite Reflux Conditions for Multicomponent Mixtures 162
- 8.8 Column Dimensions 164
- 8.9 Conceptual Design of Distillation 174
- 8.10 Detailed Design of Distillation 176
- 8.11 Limitations of Distillation 179
- 8.12 Separation of Homogeneous Fluid Mixtures by Distillation – Summary 180
- 8.13 Exercises 180 References 183

9 Separation of Homogeneous Fluid Mixtures II – Other Methods 185

- 9.1 Absorption and Stripping 185
- 9.2 Liquid–Liquid Extraction 189
- 9.3 Adsorption 196
- 9.4 Membranes 199
- 9.5 Crystallization 211
- 9.6 Evaporation 215
- 9.7 Separation of Homogeneous Fluid Mixtures by Other Methods – Summary 217
- 9.8 Exercises 217 References 219

10 Distillation Sequencing 221

- 10.1 Distillation Sequencing using Simple Columns 221
- 10.2 Practical Constraints Restricting Options 221
- 10.3 Choice of Sequence for Simple Nonintegrated Distillation Columns 222
- 10.4 Distillation Sequencing using Columns With More Than Two Products 229
- 10.5 Distillation Sequencing using Thermal Coupling 231

- 10.6 Retrofit of Distillation Sequences 236
- 10.7 Crude Oil Distillation 237
- 10.8 Structural Optimization of Distillation Sequences 239
- 10.9 Distillation Sequencing Summary 242
- 10.10 Exercises 242 References 245

11 Distillation Sequencing for Azeotropic Distillation 247

- 11.1 Azeotropic Systems 247
- **11.2** Change in Pressure 247
- 11.3 Representation of Azeotropic Distillation 248
- **11.4** Distillation at Total Reflux Conditions 250
- 11.5 Distillation at Minimum Reflux Conditions 255
- **11.6** Distillation at Finite Reflux Conditions 256
- 11.7 Distillation Sequencing Using an Entrainer 259
- 11.8 Heterogeneous Azeotropic Distillation 264
- 11.9 Entrainer Selection 267
- 11.10 Multicomponent Systems 270
- 11.11 Trade-Offs in Azeotropic Distillation 270
- 11.12 Membrane Separation 270
- 11.13 Distillation Sequencing for Azeotropic Distillation – Summary 271
- 11.14Exercises272References273

12 Heat Exchange 275

- 12.1 Overall Heat Transfer Coefficients 275
- 12.2 Heat Exchanger Fouling 279
- 12.3 Temperature Differences in Shell-and-Tube Heat Exchangers 281
- 12.4 Heat Exchanger Geometry 288
- 12.5 Allocation of Fluids in Shell-and-Tube Heat Exchangers 294
- 12.6 Heat Transfer Coefficients and Pressure Drops in Shell-and-Tube Heat Exchangers 294
- 12.7 Rating and Simulation of Heat Exchangers 301
- 12.8 Heat Transfer Enhancement 307
- 12.9 Retrofit of Heat Exchangers 313
- 12.10 Condensers 316
- 12.11 Reboilers and Vaporizers 321
- 12.12 Other Types of Heat Exchangers 326
- 12.13 Fired Heaters 328
- 12.14 Heat Exchange Summary 345
- 12.15 Exercises 346 References 348

13 Pumping and Compression 349

- 13.1 Pressure Drops in Process Operations 349
- 13.2 Pressure Drops in Piping Systems 349
- **13.3** Pump Types 355
- 13.4 Centrifugal Pump Performance 356
- 13.5 Compressor Types 363
- 13.6 Reciprocating Compressors 366

- 13.7 Dynamic Compressors 367
- **13.8** Staged Compression 369
- **13.9** Compressor Performance 370
- **13.10** Process Expanders 372
- 13.11 Pumping and Compression Summary 374
- **13.12** Exercises 374
- References 375

14 Continuous Process Recycle Structure 377

- 14.1 The Function of Process Recycles 377
- 14.2 Recycles with Purges 382
- 14.3 Hybrid Reaction and Separation 385
- 14.4 The Process Yield 386
- 14.5 Feed, Product and Intermediate Storage 38814.6 Continuous Process Recycle
 - Structure Summary 389
- 14.7 Exercises 389 References 391

15 Continuous Process Simulation and Optimization 393

- 15.1 Physical Property Models for Process Simulation 393
- 15.2 Unit Models for Process Simulation 394
- 15.3 Flowsheet Models 400
- 15.4 Simulation of Recycles 400
- 15.5 Convergence of Recycles 402
- 15.6 Design Specifications 408
- 15.7 Flowsheet Sequencing 408
- 15.8 Model Validation 408
- **15.9** Process Optimization 408
- 15.10 Continuous Process Simulation and Optimization – Summary 413
- 15.11 Exercises 413 References 416

16 Batch Processes 417

- 16.1 Characteristics of Batch Processes 417
- 16.2 Batch Reactors 417
- **16.3** Batch Distillation 420
- **16.4** Batch Crystallization 431
- 16.5 Batch Filtration 432
- 16.6 Batch Heating and Cooling 433
- 16.7 Optimization of Batch Operations 436
- 16.8 Gantt Charts 442
- 16.9 Production Schedules for Single Products 442
- 16.10 Production Schedules for Multiple Products 444
- 16.11 Equipment Cleaning and Material Transfer 445
- 16.12 Synthesis of Reaction and Separation Systems for Batch Processes 446
- 16.13 Storage in Batch Processes 452
- 16.14 Batch Processes Summary 452
- 16.15 Exercises 452 References 455

17 Heat Exchanger Networks I – Network Targets 457

- 17.1 Composite Curves 457
- 17.2 The Heat Recovery Pinch 461
- 17.3 Threshold Problems 464
- **17.4** The Problem Table Algorithm 466
- 17.5 Non-global Minimum Temperature Differences 472
- 17.6 Process Constraints 473
- **17.7** Utility Selection 475
- **17.8** Furnaces 477
- 17.9 Cogeneration (Combined Heat and Power Generation) 480
- 17.10 Integration of Heat Pumps 485
- 17.11 Number of Heat Exchange Units 486
- 17.12 Heat Exchange Area Targets 489
- 17.13 Sensitivity of Targets 493
- 17.14 Capital and Total Cost Targets 493
- 17.15 Heat Exchanger Network Targets Summary 496
- 17.16 Exercises 496 References 499

18 Heat Exchanger Networks II – Network Design 501

- **18.1** The Pinch Design Method 501
- 18.2 Design for Threshold Problems 507
- 18.3 Stream Splitting 507
- 18.4 Design for Multiple Pinches 511
- 18.5 Remaining Problem Analysis 516
- 18.6 Simulation of Heat Exchanger Networks 518
- 18.7 Optimization of a Fixed Network Structure 520
- 18.8 Automated Methods of Heat Exchanger Network Design 523
- 18.9 Heat Exchanger Network Retrofit with a Fixed Network Structure 525
- 18.10 Heat Exchanger Network Retrofit through Structural Changes 530
- 18.11 Automated Methods of Heat Exchanger Network Retrofit 536
- 18.12 Heat Exchanger Network Design Summary 538
- 18.13 Exercises 539 References 542

19 Heat Exchanger Networks III – Stream Data 543

- **19.1** Process Changes for Heat Integration 543
- 19.2The Trade-Offs Between Process Changes, Utility
Selection, Energy Cost and Capital Cost 543
- 19.3 Data Extraction 544
- 19.4Heat Exchanger Network StreamData Summary551
- 19.5 Exercises 551 References 553

20 Heat Integration of Reactors 555

- 20.1 The Heat Integration Characteristics of Reactors 555
- 20.2 Appropriate Placement of Reactors 557
- 20.3 Use of the Grand Composite Curve for Heat Integration of Reactors 558
- 20.4 Evolving Reactor Design to Improve Heat Integration 560
- 20.5 Heat Integration of Reactors Summary 561
- 20.6 Exercises 561 Reference 561

21 Heat Integration of Distillation 563

- 21.1 The Heat Integration Characteristics of Distillation 563
- 21.2 The Appropriate Placement of Distillation 563
- 21.3 Use of the Grand Composite Curve for Heat Integration of Distillation 564
- 21.4 Evolving the Design of Simple Distillation Columns to Improve Heat Integration 564
- 21.5 Heat Pumping in Distillation 567
- 21.6 Capital Cost Considerations for the Integration of Distillation 567
- 21.7 Heat Integration Characteristics of Distillation Sequences 568
- 21.8 Design of Heat Integrated Distillation Sequences 571
- **21.9** Heat Integration of Distillation Summary 572
- 21.10 Exercises 572 References 575

22 Heat Integration of Evaporators and Dryers 577

- 22.1 The Heat Integration Characteristics of Evaporators 577
- 22.2 Appropriate Placement of Evaporators 577
- 22.3 Evolving Evaporator Design to Improve Heat Integration 577
- 22.4 The Heat Integration Characteristics of Dryers 579
- 22.5 Evolving Dryer Design to Improve Heat Integration 579
- 22.6 A Case Study 581
- 22.7 Heat Integration of Evaporators and Dryers – Summary 581
- 22.8 Exercises 582 References 582

23 Steam Systems and Cogeneration 583

- **23.1** Boiler Feedwater Treatment 585
- 23.2 Steam Boilers 589
- 23.3 Gas Turbines 595
- **23.4** Steam Turbines 602
- 23.5 Steam Distrubution 609
- **23.6** Site Composite Curves 612

- 23.7 Cogeneration Targets 623
- 23.8 Power Generation and Machine Drives 627
- 23.9 Utility Simulation 631
- 23.10 Optimizing Steam Systems 633
- 23.11 Steam Costs 638
- 23.12 Steam Systems and Cogeneration Summary 641
- 23.13 Exercises 642 References 645

24 Cooling and Refrigeration Systems 647

- 24.1 Cooling Systems 647
- 24.2 Once-Through Water Cooling 647
- 24.3 Recirculating Cooling Water Systems 647
- 24.4 Air Coolers 650
- 24.5 Refrigeration 656
- 24.6 Choice of a Single-Component Refrigerant for Compression Refrigeration 662
- 24.7 Targeting Refrigeration Power for Pure Component Compression Refrigeration 665
- 24.8 Heat Integration of Pure Component Compression Refrigeration Processes 669
- 24.9 Mixed Refrigerants for Compression Refrigeration 673
- 24.10 Expanders 677
- 24.11 Absorption Refrigeration 681
- 24.12 Indirect Refrigeration 682
- 24.13 Cooling Water and Refrigeration Systems – Summary 682
- 24.14 Exercises 683 References 685

25 Environmental Design for Atmospheric Emissions 687

- 25.1 Atmospheric Pollution 687
- 25.2 Sources of Atmospheric Pollution 688
- 25.3 Control of Solid Particulate Emissions to Atmosphere 690
- 25.4 Control of VOC Emissions 690
- 25.5 Control of Sulfur Emissions 703
- 25.6 Control of Oxides of Nitrogen Emissions 708
- 25.7 Control of Combustion Emissions 711
- 25.8 Atmospheric Dispersion 714
- 25.9 Environmental Design for Atmospheric Emissions – Summary 716
- 25.10 Exercises 717 References 720

26 Water System Design 721

- 26.1 Aqueous Contamination 724
- 26.2 Primary Treatment Processes 725
- 26.3 Biological Treatment Processes 729
- 26.4 Tertiary Treatment Processes 732
- 26.5 Water Use 733
- 26.6 Targeting for Maximum Water Reuse for Single Contaminants for Operations with Fixed Mass Loads 735

- 26.7 Design for Maximum Water Reuse for Single Contaminants for Operations with Fixed Mass Loads 737
- 26.8 Targeting for Maximum Water Reuse for Single Contaminants for Operations with Fixed Flowrates 747
- 26.9 Design for Maximum Water Reuse for Single Contaminants for Operations with Fixed Flowrates 751
- 26.10 Targeting and Design for Maximum Water Reuse Based on Optimization of a Superstructure 758
- 26.11 Process Changes for Reduced Water Consumption 760
- 26.12 Targeting for Minimum Wastewater Treatment Flowrate for Single Contaminants 761
- 26.13 Design for Minimum Wastewater Treatment Flowrate for Single Contaminants 765
- 26.14 Regeneration of Wastewater 767
- 26.15 Targeting and Design for Effluent Treatment and Regeneration Based on Optimization of a Superstructure 772
- 26.16 Data Extraction 773
- 26.17 Water System Design Summary 775
- 26.18 Exercises 776 References 779

27 Environmental Sustainability in Chemical Production 781

- 27.1 Life Cycle Assessment 781
- 27.2 Efficient Use of Raw Materials Within Processes 786
- 27.3 Efficient Use of Raw Materials Between Processes 792
- 27.4 Exploitation of Renewable Raw Materials 794
- 27.5 Efficient Use of Energy 795
- 27.6 Integration of Waste Treament and Energy Sytems 805
- 27.7 Renewable Energy 806
- 27.8 Efficient Use of Water 807
- 27.9 Sustainability in Chemical Production – Summary 807
- 27.10 Exercises 808 References 809

28 Process Safety 811

- 28.1 Fire 811
- 28.2 Explosion 812
- **28.3** Toxic Release 813
- 28.4 Hazard Identification 813
- 28.5 The Hierarchy of Safety Management 815
- 28.6 Inherently Safer Design 815
- 28.7 Layers of Protection 819
- 28.8 Hazard and Operability Studies 822
- 28.9 Layer of Protection Analysis 823
- 28.10 Process Safety Summary 823

28.11 Exercises 824 References 825

Appendix A Physical Properties in Process Design 827

- A.1 Equations of State 827
- A.2 Phase Equilibrium for Single Components 831
- A.3 Fugacity and Phase Equilibrium 831
- A.4 Vapor–Liquid Equilibrium 831
- A.5 Vapor–Liquid Equilibrium Based on Activity Coefficient Models 833
- A.6 Group Contribution Methods for Vapor–Liquid Equilibrium 835
- A.7 Vapor–Liquid Equilibrium Based on Equations of State 837
- A.8 Calculation of Vapor–Liquid Equilibrium 838
- A.9 Liquid–Liquid Equilibrium 841
- A.10 Liquid–Liquid Equilibrium Activity Coefficient Models 842
- A.11 Calculation of Liquid–Liquid Equilibrium 842
- A.12 Choice of Method for Equilibrium Calculations 844
- A.13 Calculation of Enthalpy 846
- A.14 Calculation of Entropy 847
- A.15 Other Physical Properties 848
- A.16 Physical Properties in Process Design – Summary 850
- A.17 Exercises 851 References 852

Appendix B Materials of Construction 853

- B.1 Mechanical Properties 853
- B.2 Corrosion 854
- B.3 Corrosion Allowance 855
- B.4 Commonly Used Materials of Construction 855
- B.5 Criteria for Selection 859
- B.6 Materials of Construction Summary 860 References 860

Appendix C Annualization of Capital Cost 861

Reference 861

Appendix DThe Maximum Thermal Effectiveness for 1–2Shell-and-Tube Heat Exchangers863

References 863

Appendix E Expression for the Minimum Number of 1–2 Shell-and-Tube Heat Exchangers for a Given Unit 865

References 866

Appendix F Heat Transfer Coefficient and Pressure Drop in Shell-and-Tube Heat Exchangers 867

F.1 Heat Transfer and Pressure Drop Correlations for the Tube Side 867 F.2 Heat Transfer and Pressure Drop Correlations for the Shell Side 869 References 873

Appendix G Gas Compression Theory 875

- G.1 Modeling Reciprocating Compressors 875
- G.2 Modeling Dynamic Compressors 877

G.3 Staged Compression 877 References 879

Appendix H Algorithm for the Heat Exchanger Network Area Target 881

Index 883

Preface to the Second Edition PREFACE TO THE SECOND EDITION

his book deals with the design and integration of chemical processes. The Second Edition has been rewritten, restructured and updated throughout from the First Edition. At the heart of the book are the conceptual issues that are fundamental to the creation of chemical processes and their integration to form complete manufacturing systems. Compared with the First Edition, this edition includes much greater consideration of equipment and equipment design, including materials of construction, whilst not sacrificing understanding of the overall conceptual design. Greater emphasis has also been placed on physical properties, process simulation and batch processing. Increasing environmental awareness has dictated the necessity of a greater emphasis on environmental sustainability throughout. The main implication of this for process design is greater efficiency in the use of raw materials, energy and water and a greater emphasis on process safety. Consideration of integration has not been restricted to individual processes, but integration across processes has also been emphasized to create environmentally sustainable integrated manufacturing systems. Thus, the text integrates equipment, process and manufacturing system design. This edition has been rewritten to make it more accessible to undergraduate students of chemical engineering than the First Edition, as well as maintaining its usefulness to

postgraduate students of chemical engineering and to practicing chemical engineers.

As with the first edition, this edition as much as possible emphasizes understanding of process design methods, as well as their application. Where practical, the derivation of design equations has been included, as this is the best way to understand the limitations of those equations and to ensure their wise application.

The book is intended to provide a practical guide to chemical process design and integration for students of chemical engineering at all levels, practicing process designers and chemical engineers and applied chemists working in process development. For undergraduate studies, the text assumes basic knowledge of material and energy balances and thermodynamics, together with basic spread-sheeting skills. Worked examples have been included throughout the text. Most of these examples do not require specialist software and can be solved either by hand or using spreadsheet software. A suite of Excel spreadsheets has also been made available to allow some of the more complex example calculations to be performed more conveniently. Finally, a number of exercises has been added at the end of each chapter to allow the reader to practice the calculation procedures. A solutions manual is available.

Robin Smith

Acknowledgements ACKNOWLEDGEMENTS

The author would like to express gratitude to a number of people who have helped in the preparation of the Second Edition.

From the University of Manchester: Mary Akpomiemie, Adisa Azapagic, Stephen Doyle, Victor Manuel Enriquez Gutierrez, Oluwagbemisola Oluleye, Kok Siew Ng, Li Sun and Colin Webb

Interns at the University of Manchester: Rabia Amaaouch, Béatrice Bouchon, Aymeric Cambrillat, Leo Gandrille, Kathrin Holzwarth, Guillemette Nicolas and Matthias Schmid

From National Technical University of Athens: Antonis Kokossis

From the University of Huddersfield: Grant Campbell From Nova Process Ltd: Stephen Hall

From Norwegian University of Science and Technology: Truls Gundersen

From BP: Paul Oram

Gratitude is also expressed to Lucy Adams, Ellen Gleeson and Tim Mummery for help in the preparation of the figures and the text.

Finally, gratitude is expressed to all of the member companies of the Process Integration Research Consortium, both past and present. Their support has made a considerable contribution to research in the area, and hence to this text.

Nomenclature NOMENCLATURE

а	Activity (–), or constant in cubic equation of state	B_C	Baffle cut for shell-and-tube heat exchangers
	$(N \cdot m^4 \cdot kmol^{-2})$ or	ROD	Biological oxygen demand $(kg \cdot m^{-3} mg \cdot l^{-1})$
	correlating coefficient (units depend on	C	Capital cost law coefficient (–) or
	application) or	U	correlating coefficient (units depend on
	cost law coefficient (\$) or		application) or
	order of reaction $(-)$		order of reaction (-)
a	Group interaction parameter in the UNIEAC	(p	Drag coefficient (-)
u _{mn}	Model (K)	e _D	Eanning friction factor (_)
a. a.	Profile control parameters in optimization (_)	c_f	Smooth tube Eanning friction factor $(-)$
a_1, a_2	Absorption factor in absorption $(-)$ or	c_{fS}	Loss coefficient for nine or nine fitting (-)
Л	annual cash flow (\$) or	C_L	Concentration $(kg, m^{-3}, kmol, m^{-3}, nnm)$ or
	constant in vanor pressure correlation	C	constant in vapor pressure correlation (\mathbf{K}) or
	$(N \cdot m^{-2} \text{ bar})$ or		number of components (separate systems) in
	(\mathbf{N} , \mathbf{m} , \mathbf{D} , \mathbf{D}), \mathbf{D}		number of components (separate systems) in network design (_)
A	Cross sectional area of column (m^2)	C	Base capital cost of equipment (\$)
A_C	Appual cash flow $(\$, v^{-1})$	C_B	Environmental discharge concentration (nnm)
A _{CF}	Annual cash now $(\mathfrak{z} \cdot \mathfrak{y}^{-})$	Ce C	Environmental discharge concentration (ppin)
A_D	Area occupied by distination downcomer (iii) Appual discounted cash flow (f_{1}, v_{1}^{-1})	C_E	Equipment capital cost (\$), of unit cost of energy ($\$$ kW^{-1} $\$$ MW^{-1})
A _{DCF}	Amount discounted cash now $(\mathbf{x} \cdot \mathbf{y})$	C	$ \begin{array}{c} \text{unit cost of energy } (5 \cdot \mathbf{K} \mathbf{W} , 5 \cdot \mathbf{M} \mathbf{W}) \\ \text{Eived conital cost of complete installation } (5) \\ \end{array} $
AFIN	Area of fins (m)	C_F	Fixed capital cost of complete instantation (5)
A_I	Heat transfer area on the inside of tubes (m^2) or	C_P	Specific near capacity at constant pressure $(k_{\rm L}, k_{\rm S}^{-1}, W^{-1}, k_{\rm L}, k_{\rm S}^{-1}, W^{-1})$
	(III), of $(112)^{-3}$		$(KJ \cdot Kg \cdot K , KJ \cdot KIIIOI \cdot K)$
4	M_{embrane} and (m, m, m^2)	C_P	Mean neat capacity at constant pressure $(\mathbf{k} \mathbf{k} = 1, \mathbf{k}^{-1} \mathbf{k} \mathbf{k} = 1, \mathbf{k}^{-1})$
A_M	Memorane area (m) (m^2)	C	$(KJ \cdot Kg \cdot K , KJ \cdot KIIIOI \cdot K)$
ANETWORK	Heat exchanger network area (m) Heat transfer area on the outside of tubes (m^2)	\mathcal{C}_S	Corrected superficial velocity in distillation $(m - e^{-1})$
A _O	Expand outside root area of a fund tube (iii)	C	(III · S) Specific heat conscitu at constant volume
A _{ROOT}	Exposed outside foot area of a finited tube (iii) Heat exchanger area for an individual shall (m^2)	C_V	Specific field capacity at constant volume $(k_{\rm L} k_{\rm g}^{-1} K^{-1} k_{\rm L} k_{\rm mol}^{-1} K^{-1})$
ASHELL	Appuelization factor for capital cost ()	C^*	$(\mathbf{K} \cdot \mathbf{K} \mathbf{g} \cdot \mathbf{K}, \mathbf{K} \cdot \mathbf{K})$ Solubility of solute in solvent $(\mathbf{k} \mathbf{g}, \mathbf{k} \mathbf{g}, \mathbf{s} \mathbf{g})$
AI	capital cost law coefficient (units depend on cost		Cycles of concentration for a cooling tower (-)
	law) or	CC	Cumulative cost $(\$, t^{-1})$
	constant in cubic equation of state	COD	Chemical oxygen demand $(kg \cdot m^{-3} - mg \cdot l^{-1})$
	$(m^3, kmol^{-1})$ or	COP	Coefficient of performance $(-)$
	correlating coefficient (units depend on	COP	Coefficient of performance of an absorption heat
	application) or		nump (–)
	order of reaction (-)	COP	Coefficient of performance of an absorption heat
h.	Bottoms flowrate of Component <i>i</i> (kmol \cdot s ⁻¹	COLAHI	transformer (–)
\mathcal{O}_{i}	$k \text{mol} \cdot h^{-1}$	$COP_{4,p}$	Coefficient of performance of absorption
R	Baffle spacing in shell-and-tube heat exchangers		refrigeration (-)
Б	(m) or	COP	Coefficient of performance of a compression
	Bottoms flowrate in distillation $(ka_{s})^{-1}$	COLCHP	heat nump (_)
	$kg_{\rm s}h^{-1}$ kmol s ⁻¹ kmol h ⁻¹) or	COP	Coefficient of performance of a heat nump (-)
	breadth of device (m) or	COP	Coefficient of performance of a refrigeration
	constant in vanor pressure correlation	COI REF	system (_)
	$(N, K, m^{-2} \text{ bar}, K)$ or	СР	Canacity parameter in distillation (m, s^{-1}) or
	moles remaining in batch distillation (kmcl)	CI	best capacity flowrote (LW, K^{-1}, MW, K^{-1})
	mores remaining in batch distination (kinol)		near capacity nowrate (KW·K, WW·K)

CP_{EX}	Heat capacity flowrate of heat engine exhaust $(kW \cdot K^{-1}, MW \cdot K^{-1})$	F_{TC}	Correction factor for tube count in shell-and- tube heat exchangers (–)
CW	Cooling water	F_{Tmin}	Minimum acceptable F_T for noncountercurrent
d	Diameter (µm, m), or		heat exchangers (–)
	correlating coefficient (units depend on	F_{XY}	Factor to allow for inclination in structured
	application)		packing (–)
d_C	Column inside diameter (m)	F_{σ}	Factor to allow for inadequate wetting of
d_i	Distillate flowrate of Component <i>i</i> (kmol \cdot s ⁻¹ ,		packing (–)
	$kmol \cdot h^{-1}$)	g	Acceleration due to gravity (9.81 m \cdot s ⁻²)
d_I	Inside diameter of pipe or tube (m)	g_{ij}	Energy of interaction between Molecules i
d_P	Distillation and absorption packing size (m)		and j in the NRTL equation $(kJ \cdot kmol^{-1})$
d_R	Outside tube diameter for a finned tube at the	G	Free energy (kJ), or
	root of fins (m)		gas flowrate (kg \cdot s ⁻¹ , kmol \cdot s ⁻¹)
D	Distillate flowrate $(kg \cdot s^{-1}, kg \cdot h^{-1}, kmol \cdot s^{-1},$	\overline{G}_i	Partial molar free energy of Component i
	$\operatorname{kmol} \cdot \operatorname{h}^{-1}$)		$(kJ \cdot kmol^{-1})$
D_B	Tube bundle diameter for shell-and-tube heat	\overline{G}_i^O	Standard partial molar free energy of
	exchangers (m)	r.	Component i (kJ · kmol ⁻¹)
D_S	Inside shell diameter for shell-and-tube heat	GCV	Gross calorific value of fuel $(J \cdot m^{-3}, kJ \cdot m^{-3}, $
	exchangers (m)		$\mathbf{J} \cdot \mathbf{kg}^{-1}, \mathbf{kJ} \cdot \mathbf{kg}^{-1})$
DCFRR	Discounted cash flowrate of return (%)	h	Settling distance of particles (m)
e	Wire diameter (m)	h_B	Boiling heat transfer coefficient for the tube
E	Activation energy of reaction $(kJ \cdot kmol^{-1})$, or		bundle $(\mathbf{W} \cdot \mathbf{m}^{-2} \cdot \mathbf{K}^{-1}, \mathbf{kW} \cdot \mathbf{m}^{-2} \cdot \mathbf{K}^{-1})$
	entrainer flowrate in azeotropic and extractive	h_C	Condensing film heat transfer coefficient
	distillation (kg \cdot s ⁻¹ , kmol \cdot s ⁻¹), or		$(\mathbf{W} \cdot \mathbf{m}^{-2} \cdot \mathbf{K}^{-1}, \mathbf{kW} \cdot \mathbf{m}^{-2} \cdot \mathbf{K}^{-1})$
	exchange factor in radiant heat transfer (-), or	h_I	Film heat transfer coefficient for the inside
	extract flowrate in liquid-liquid extraction		$(\mathbf{W} \cdot \mathbf{m}^{-2} \cdot \mathbf{K}^{-1}, \mathbf{kW} \cdot \mathbf{m}^{-2} \cdot \mathbf{K}^{-1})$
	$(\text{kg} \cdot \text{s}^{-1}, \text{kmol} \cdot \text{s}^{-1}), \text{ or }$	h_{IF}	Fouling heat transfer coefficient for the inside
	stage efficiency in separation (–)		$(\mathbf{W} \cdot \mathbf{m}^{-2} \cdot \mathbf{K}^{-1}, \mathbf{kW} \cdot \mathbf{m}^{-2} \cdot \mathbf{K}^{-1})$
E_O	Overall stage efficiency in distillation and	h_L	Head loss in a pipe or pipe fitting (m)
	absorption (–)	h_{NB}	Nucleate boiling heat transfer coefficient
EP	Economic potential $(\$ \cdot y^{-1})$		$(\mathbf{W} \cdot \mathbf{m}^{-2} \cdot \mathbf{K}^{-1}, \mathbf{kW} \cdot \mathbf{m}^{-2} \cdot \mathbf{K}^{-1})$
f	Fuel-to-air ratio for gas turbine (–)	h_O	Film heat transfer coefficient for the outside
f_i	Capital cost installation factor for Equipment		$(\mathbf{W} \cdot \mathbf{m}^{-2} \cdot \mathbf{K}^{-1}, \mathbf{k}\mathbf{W} \cdot \mathbf{m}^{-2} \cdot \mathbf{K}^{-1})$
	<i>i</i> (–), or	h_{OF}	Fouling heat transfer coefficient for the outside
	feed flowrate of Component <i>i</i> (kmol \cdot s ⁻¹ ,		$(\mathbf{W} \cdot \mathbf{m}^{-2} \cdot \mathbf{K}^{-1}, \mathbf{kW} \cdot \mathbf{m}^{-2} \cdot \mathbf{K}^{-1})$
	kmol \cdot h ⁻¹), or	h_{RAD}	Radiant heat transfer coefficient ($W \cdot m^{-2} \cdot K^{-1}$,
	fugacity of Component i (N \cdot m ⁻² , bar)		$kW \cdot m^{-2} \cdot K^{-1}$
f_P	Capital cost factor to allow for design	h_W	Heat transfer coefficient for the tube wall
	pressure (–)		$(\mathbf{W} \cdot \mathbf{m}^{-2} \cdot \mathbf{K}^{-1}, \mathbf{kW} \cdot \mathbf{m}^{-2} \cdot \mathbf{K}^{-1})$
f_T	Capital cost factor to allow for design	H	Enthalpy $(kJ, kJ \cdot kg^{-1}, kJ \cdot kmol^{-1})$, or
	temperature (–)		height (m), or
F	Feed flowrate $(kg \cdot s^{-1}, kg \cdot h^{-1}, kmol \cdot s^{-1},$		Henry's Law Constant (N \cdot m ⁻² , bar, atm), or
	kmol \cdot h ⁻¹), or		stream enthalpy $(kJ \cdot s^{-1}, MJ \cdot s^{-1})$
	future worth a sum of money allowing for	H_F	Height of fin (m)
	interest rates (\$), or	H_T	Tray spacing (m)
	number of degrees of freedom $(-)$, or	\overline{H}_{i}^{O}	Standard heat of formation of Component <i>i</i>
F	volumetric flowrate $(m^2 \cdot s^{-1}, m^2 \cdot h^{-1})$	0	$(kJ \cdot kmol^{-1})$
F _{FOAM}	Foaming factor in distillation $(-)$	ΔH^O	Standard heat of reaction (J, kJ)
F_{LV}	Liquid–vapor flow parameter in distillation (–)	ΔH_{COMB}	Heat of combustion $(J \cdot kmol^{-1}, kJ \cdot kmol^{-1})$
r' _{RAD}	Fraction of near absorbed in fired heater radiant	ΔH_{COMB}^0	Standard heat of combustion at 298 K
F	Security for the large for the		$(J \cdot kmol^{-1}, kJ \cdot kmol^{-1})$
F _{SC}	correction factor for shell construction in shell-	ΔH_{FUEL}	Heat to bring fuel to standard temperature
F	and-tube neat exchangers (-)		$(J \cdot kmol^{-1}, kJ \cdot kg^{-1})$
Γ _T	contection factor for noncountercurrent flow in	ΔH_{IS}	Isentropic enthalpy change of an expansion
	shen-and-tube neat exchangers (-)		$(\mathbf{J} \cdot \mathbf{kmol}^{-1}, \mathbf{kJ} \cdot \mathbf{kg}^{-1})$

ΔH_P	Heat to bring products from standard	т
	temperature to the final temperature	
	$(\mathbf{J} \cdot \mathbf{kmol}^{-1}, \mathbf{kJ} \cdot \mathbf{kg}^{-1})$	
ΔH_R	Heat to bring reactants from their initial	m_C
	temperature to standard temperature	
	$(\mathbf{J} \cdot \mathbf{kmol}^{-1}, \mathbf{kJ} \cdot \mathbf{kmol}^{-1})$	m_{CO}
ΔH_{STEAM}	Enthalpy difference between generated steam	m_{EX}
	and boiler feedwater (kW, MW)	m_{FU}
ΔH_{VAP}	Latent heat of vaporization $(kJ \cdot kg^{-1}, kJ \cdot kmol^{-1})$	m_{ma}
HETP	Height equivalent of a theoretical plate (m)	т
HP	High pressure	mw
HR	Heat rate for gas turbine $(kI \cdot kWh^{-1})$	mw
i	Fractional rate of interest on money $(-)$, or	
	number of ions (–)	mu
I	Total number of hot streams $(-)$	
I	Total number of cold streams (-)	mur
s k	Reaction rate constant (units depend on order of	mwr
ĸ	reaction) or	mwi
	step number in a numerical calculation $(-)$ or	М
	thermal conductivity (W, m^{-1} , K^{-1}	171
	$\frac{1}{2} \sum_{k=1}^{N} \frac{1}{k} \sum_{k=1}^{N-1} $	
l-	$\mathbf{K} \mathbf{W} \cdot \mathbf{H} \mathbf{I} \cdot \mathbf{K}$) Fin thermal conductivity ($\mathbf{W} = \mathbf{m}^{-1} \mathbf{K}^{-1}$	MD
κ_F	Fin thermal conductivity ($\mathbf{w} \cdot \mathbf{m} \cdot \mathbf{K}$,	MP
1	KW·M·K)	MC
$K_{G,i}$	Mass transfer coefficient in the gas phase $(1 - 1)^{-2}$ $\mathbf{p} = 1 = 1$	n
,	$(\text{kmol} \cdot \text{m}^2 \cdot \text{Pa}^2 \cdot \text{s}^2)$	
k _{ij}	Interaction parameter between Components <i>i</i>	
	and j in an equation of state (-)	
$k_{L,i}$	Mass transfer coefficient of Component <i>i</i> in the	Ν
	liquid phase $(m \cdot s^{-1})$	
k_0	Frequency factor for heat of reaction (units	
	depend on order of reaction)	
k_W	Wall thermal conductivity $(W \cdot m^{-1} \cdot K^{-1})$,	
	$kW \cdot m^{-1} \cdot K^{-1}$	
Κ	Overall mass transfer coefficient	
	$(\operatorname{kmol} \cdot \operatorname{Pa}^{-1} \cdot \operatorname{m}^{-2} \cdot \operatorname{s}^{-1}), \text{ or }$	N_F
	rate constant for fouling $(m^2 \cdot K \cdot W^{-1} \cdot day^{-1})$, or	N_i
	total number of enthalpy intervals in heat	N_{i0}
	exchanger networks (-)	N_{min}
K_a	Equilibrium constant of reaction based on	N_P
	activity (–)	N_R
K_i	Ratio of vapor-to-liquid composition at	N_{SH}
	equilibrium for Component i (–)	
$K_{M,i}$	Equilibrium partition coefficient of membrane	N_R
	for Component $i(-)$	N_S
K_p	Equilibrium constant of reaction based on	N_T
Ĩ	partial pressure in the vapor phase (–)	N_{TR}
K_T	Parameter for terminal settling velocity $(m \cdot s^{-1})$	N_{UN}
K _x	Equilibrium constant of reaction based on mole	
	fraction in the liquid phase (–)	NC
$K_{\rm v}$	Equilibrium constant of reaction based on mole	
,	fraction in vapor phase (–)	NCV
L	length (m), or	
	liquid flowrate (kg \cdot s ⁻¹ , kmol \cdot s ⁻¹), or	NPS
	number of independent loops in a	NPL
	network (–)	Nu
Lw	Distillation tray weir length (m)	n
-w LP	Low pressure	ľ
	P	

	Mass flowrate $(kg \cdot s^{-1})$, or
	molar flowrate (kmol \cdot s ⁻¹), or
	number of items (-)
2	Mass flowrate of water contaminant
	$(\mathbf{g}\cdot\mathbf{h}^{-1},\mathbf{g}\cdot\mathbf{d}^{-1})$
COND	Mass of condensate (kg)
X	Mass flowrate of exhaust $(kg \cdot s^{-1})$
UEL	Mass of fuel (kg)
ıax	Maximum mass flowrate $(kg \cdot s^{-1})$
TEAM	Mass flowrate of steam $(\text{kg} \cdot \text{s}^{-1})$
V	Mass flowrate of pure water $(t \cdot h^{-1}, t \cdot d^{-1})$
VL	Limiting mass flowrate of pure water
	$(\mathbf{t} \cdot \mathbf{h}^{-1}, \mathbf{t} \cdot \mathbf{d}^{-1})$
Vmin	Minimum mass flowrate of fresh water
	$(\mathbf{t} \cdot \mathbf{h}^{-1}, \mathbf{t} \cdot \mathbf{d}^{-1})$
VT	Target mass flowrate of fresh water $(t \cdot h^{-1}, t \cdot d^{-1})$
VTLOSS	Target mass flowrate of fresh water involving a
	water loss $(t \cdot h^{-1}, t \cdot d^{-1})$
	Constant in capital cost correlations (-), or
	molar mass $(kg \cdot kmol^{-1})$, or
	number of variables (-)
D	Medium pressure
STEAM	Marginal cost of steam $(\$ \cdot t^{-1})$
	Number of items (–), or
	number of years (-), or
	polytropic coefficient (-), or
	slope of Willans Line $(kJ \cdot kg^{-1}, MJ \cdot kg^{-1})$
	Number of compression stages (-), or
	number of independent equations (-), or
	number of moles (kmol), or
	number of theoretical stages (-), or
	rate of transfer of a component
	$(\text{kmol} \cdot \text{s}^{-1} \cdot \text{m}^{-3})$, or
	rotational speed (s^{-1}, min^{-1})
,	Number of fins per unit length (m^{-1})
	Number of moles of Component i (kmol)
)	Initial number of moles of Component <i>i</i> (kmol)
iin	Minimum number of theoretical stages (-)
,	Number of tube passes (–)
	Number of tube rows (–)
HELLS	Number of number of 1–2 shells in shell-and-
	tube heat exchangers (-)
	Number of tube rows (–)
	Specific speed of centrifugal pump (-)
	Number of tubes (–)
R	Number of tubes per row (–)
NITS	Number of units in a heat exchanger
	network (–)
2	Number of components in a multicomponent
	mixture (–)
CV	Net calorific value of fuel $(J \cdot m^{-3}, kJ \cdot m^{-3}, m^{-3})$
	$J \cdot kg^{-1}, kJ \cdot kg^{-1}$
PSH	Net positive suction head (m)
V	Net present value (\$)
!	Nusselt number $(-)$
	Partial pressure (N \cdot m ⁻² , bar), or
	helical pitch (m)

p_C	Pitch configuration factor for tube layout (–)	Q_{INP}
p_T	Tube pitch (m)	Q_{HP}
Р	Present worth of a future sum of money (\$), or	
	pressure (N \cdot m ⁻² , bar), or	Q_{LP}
	probability (–), or	Q_{LOS}
	thermal effectiveness of 1-2 shell-and-tube heat	
	exchanger (–)	Q_{OU}
P_C	Critical pressure (N \cdot m ⁻² , bar)	Q_{RAL}
P_{max}	Maximum thermal effectiveness of 1-2 shell-	Q_{REA}
	and-tube heat exchangers (–)	Q_{REE}
$P_{M,i}$	Permeability of Component <i>i</i> for a membrane	Q_{REC}
	$(\text{kmol} \cdot \text{m} \cdot \text{s}^{-1} \cdot \text{m}^{-2} \cdot \text{bar}^{-1},$	Q_{SITI}
	kg solvent $\cdot m^{-1} \cdot s^{-1} \cdot bar^{-1}$)	Q_{STE}
$\overline{P}_{M,i}$	Permeance of Component <i>i</i> for a membrane	r
	$(\mathbf{m}^3 \cdot \mathbf{m}^{-2} \cdot \mathbf{s}^{-1} \cdot \mathbf{bar}^{-1})$	
P_{N-2N}	Thermal effectiveness over N_{SHELLS} number of	
	1–2 shell-and-tube heat exchangers in series (–)	r_i
P_{1-2}	Thermal effectiveness over each 1–2 shell-and-	
6 4 T	tube heat exchanger in series (-)	
P^{SAT}	Saturated liquid–vapor pressure ($N \cdot m^{-2}$, bar)	
Pr	Prandtl number (–)	
ΔP	Pressure drop (N \cdot m ⁻² , bar)	
ΔP_{FLOOD}	Pressure drop under flooding conditions	R
	$(N \cdot m^{-2}, bar)$	
q	Heat flux $(W \cdot m^{-2}, kW \cdot m^{-2})$, or	
	thermal condition of the feed in distillation (-), or	
	Wegstein acceleration parameter for the	
	convergence of recycle calculations (–)	
q_C	Critical heat flux $(W \cdot m^{-2}, kW \cdot m^{-2})$	
q_{C1}	Critical heat flux for a single tube $(W \cdot m^{-2})$,	
	$kW \cdot m^{-2}$)	
q_i	Individual stream heat duty for Stream <i>i</i>	
	$(kJ \cdot s^{-1})$, or	
	pure component property measuring the	
	molecular van der Waals surface area for	
	Molecule <i>i</i> in the UNIQUAC Equation of	R_{AF}
	UNIFAC Model (-)	R_{min}
q_{RAD}	Radiant heat flux $(W \cdot m^{-2}, kW \cdot m^{-2})$	R_F
Q	Heat duty (kW, MW)	
Q_{ABS}	Absorber heat duty (kW, MW)	
Qc	Cooling duty (kW, MW)	R _{SITE}
Qc_{min}	Target for cold utility (kW, MW)	ROI
Q_{COND}	Condenser heat duty (kW, MW)	Re
Q_{CONV}	Convective heat duty (kW, MW)	S
Q_{EVAP}	Evaporator heat duty (kW, MW)	
Q_{EX}	Heat duty for heat engine exhaust	S
	(kW, MW)	
Q_{FEED}	Heat duty to the feed (kW, MW)	
Q_{FUEL}	Heat from fuel in a furnace, boiler, or gas	
	turbine (kW, MW)	
Q_{GEN}	Heat pump generator heat duty (kW, MW)	
Q_H	Heating duty (kW, MW)	
Q_{Hmin}	Target for hot utility (kW, MW)	
Q_{HE}	Heat engine heat duty (kW, MW)	
Q_{HEN}	Heat exchanger network heat duty	
	(kW, MW)	

	Heat input from fuel (kW_MW)
NPUI	Heat duty on high processing stoom (kW/ MW) or
4P	Theat duty off high-pressure steam (Kw, Ww), of
	heat pump heat duty (kw, Mw)
LP	Heat duty on low-pressure steam (kW, MW)
LOSS	Stack loss from furnace, boiler, or gas turbine
	(kW, MW)
OUTPUT	Heat output to steam generation (kW, MW)
RAD	Radiant heat duty (kW, MW)
REACT	Reactor heating or cooling duty (kW, MW)
DED	Reboiler heat duty (kW, MW)
	Heat recovery (kW_MW)
KEC	Site heating demand (kW_MW)
SITE	Hast input for steem generation (kW, MW)
STEAM	Molor ratio () or
	pressure ratio (–), or
	radius (m)
	Pure component property measuring the
	molecular van der Waals volume for Molecule i
	in the UNIQUAC Equation and UINFAC
	Model (–), or
	rate of reaction of Component <i>i</i> (kmol ^{-1} · s ^{-1}), or
	recovery of Component <i>i</i> in separation $(-)$
	Fractional recovery of a component in
	separation (-) or
	heat capacity ratio of 1, 2 shall and tube heat
	avalance () or
	exchanger (-), or
	ratinate nowrate in liquid–liquid extraction
	$(kg \cdot s^{-1}, kmol \cdot s^{-1}), or$
	ratio of heat capacity flowrates (–), or
	reflux ratio in distillation (–), or
	removal ratio in water treatment (-), or
	residual error (units depend on application), or
	universal gas constant (8314.5
	$\mathbf{N} \cdot \mathbf{m} \cdot \mathbf{kmol}^{-1} \cdot \mathbf{K}^{-1} = \mathbf{J} \cdot \mathbf{kmol}^{-1} \cdot \mathbf{K}^{-1}, 8.3145$
	$kJ \cdot kmol^{-1} \cdot K^{-1}$
E	Mass ratio of air to fuel $(-)$
	Minimum reflux ratio (-)
nin	Fouling resistance in heat transfer
e'	(m^{-2}, K, W^{-1}) or
	$(\Pi \cdot \mathbf{K} \cdot \mathbf{W}), \Theta$
	Site a server to heat ratio ()
SITE	Site power-to-neat ratio (–)
Л	Return on investment (%)
	Reynolds number (–)
	Reactor space velocity $(s^{-1}, \min^{-1}, h^{-1})$, or
	steam-to-air ratio for gas turbine (–)
	Entropy $(kJ \cdot K^{-1}, kJ \cdot kg^{-1} \cdot K^{-1})$,
	$kJ \cdot kmol^{-1} \cdot K^{-1}$), or
	number of streams in a heat exchanger
	network (–), or
	reactor selectivity (–), or
	reboil ratio for distillation $(-)$, or
	selectivity of a reaction (-), or
	slack variable in optimization (units depend on
	application) or
	application), or $a_{1}^{-1} = a_{1}^{-1} = a_{1}^{-1}$
	Solvent nowrate (kg \cdot s , kmol \cdot s), or
	surpping factor in absorption (–)

~	
S_C	Number of cold streams (–)
S_H	Number of hot streams (–)
S_W	Dimensionless swirl parameter (-)
t	Batch time (s, h), or
	time (s, h)
Т	Temperature (°C, K)
Ting	Absorber temperature ($^{\circ}C$ K)
	Normal bailing point $(^{\circ}C, K)$
I BPT	Critical terra entern (K) an
I_C	Critical temperature (K), or
_	temperature of heat sink (\mathcal{C} , K)
T_{COND}	Condenser temperature (°C, K)
T_E	Equilibrium temperature (°C, K)
T_{EVAP}	Evaporation temperature (°C, K)
T_{FEED}	Feed temperature (°C, K)
T_{GEN}	Heat pump generator temperature (°C, K)
T_H	Temperature of heat source (°C, K)
T_R	Reduced temperature $T/T_{C}(-)$
	Repoiler temperature ($^{\circ}C$ K)
	Stream supply temperature $(^{\circ}C)$
15 T	Subtraction temperature of boiling liquid ($^{\circ}C$, K)
I SAT	Stream target temperature (°C)
1 _T	Stream target temperature (C)
I _{TFT}	Theoretical fiame temperature (°C, K)
T_W	Wall temperature (°C, K)
T _{WBT}	Wet bulb temperature (°C)
T^*	Interval temperature (°C)
ΔT_{LM}	Logarithmic mean temperature difference
	(°C, K)
ΔT_{min}	Minimum temperature difference (°C, K)
$\Delta T_{THRESHOLD}$	Threshold temperature difference (°C, K)
TAC	Total annual cost $(\$ \cdot y^{-1})$
TOD	Total oxygen demand (kg \cdot m ⁻³ , mg \cdot l ⁻¹)
И.:	Interaction parameter between Molecule <i>i</i> and
ij	Molecule <i>i</i> in the UNIOUAC Equation
	$(kL, kmol^{-1})$
I I	(V) which f
U	$kW m^{-2} V^{-1}$
	$\mathbf{K}\mathbf{W} \cdot \mathbf{\Pi} \cdot \mathbf{K}$)
v	Velocity $(\mathbf{m} \cdot \mathbf{s})$
v_D	Downcomer liquid velocity $(m \cdot h^{-1})$
v_S	Shell-side fluid velocity $(\mathbf{m} \cdot \mathbf{s}^{-1})$, or
	superficial vapor velocity $(m \cdot s^{-1})$
v_T	Terminal settling velocity $(m \cdot s^{-1})$, or
	tube-side tube velocity $(m \cdot s^{-1})$
v_V	Superficial vapor velocity in empty column
	$(\mathbf{m} \cdot \mathbf{s}^{-1})$
V	Molar volume $(m^3 \cdot kmol^{-1})$, or
	vapor flowrate $(kg \cdot s^{-1}, kmol \cdot s^{-1})$, or
	volume (m^3) or
	volume of gas or vapor adsorbed (m^3, kg^{-1})
V.	Minimum vapor flow (k_{0}, e^{-1}, k_{0})
' min VE	Vanor fraction ()
vГ	vapor fraction $(-)$
W	Mass of adsorbate per mass of adsorbent (-)
W	Shatt power (kW, MW), or
	shaft work (kJ, MJ)
W_{GEN}	Power generated (kW, MW)
W_{GT}	Power generated by gas turbine (kW, MW)

W_{LOSS}	Power loss in gas or steam turbines (kW, MW)
W_{SITE}	Site power demand (kW, MW)
x	Control variable in optimization problem (units
	depend on application), or
	liquid-phase mole fraction (-)
X_F	Final value of control variable in optimization
	problem (units depend on application), or
	mole fraction in the feed (–)
X_B	Mole fraction in the distillation bottoms (-)
x_D	Mole fraction in the distillate (–)
x_0	Initial value of control variable in optimization
	problem (units depend on application)
Χ	Reactor conversion (-), or
	wetness fraction of steam (-)
X_E	Equilibrium reactor conversion (-)
X_{OPT}	Optimal reactor conversion (–)
X_P	Fraction of maximum thermal effectiveness
	P_{max} allowed in a 1–2 shell-and-tube heat
	exchanger (–)
XP	Cross-pinch heat transfer in heat exchanger
	network (kW, MW)
у	Integer variable in optimization (-), or
	twist ratio for twisted tape (-), or
	vapor-phase mole fraction (–)
y_F	Distance between fins (m)
z	Elevation (m), or
	feed mole fraction (–)
Ζ	Compressibility of a fluid (-)

Greek Letters

α	Constant in cubic equation of state (-), or
	constants in vapor pressure correlation (units depend
	on which constant), or
	fraction open of a valve (-), or
	helix angle of wire to the tube axis (degrees), or
	relative volatility between a binary pair (-)
α_{ij}	Ideal separation factor or selectivity of membrane
	between Components i and j (–), or
	parameter characterizing the tendency of Molecule i
	and Molecule <i>j</i> to be distributed in a random fashion
	in the NRTL equation (-), or
	relative volatility between Components i and j (–)
α_{LH}	Relative volatility between light and heavy key
	components (-)
α_P	Packing surface area $(m^2 \cdot m^3)$
β_{ij}	Separation factor between Components i and j (–)
γ	Logic variable in optimization (–), or
	ratio of heat capacities for gases and vapors (-)
γi	Activity coefficient for Component i (–)
δ	Thickness (m)
δ_F	Fin thickness (m)
δ_M	Membrane thickness (m)
ε	Emissivity (–), or
	extraction factor in liquid-liquid extraction (–), or

	pipe roughness (mm)	е
η	Carnot factor (–), or	
	efficiency (–)	
η_{AHP}	Absorber heat pump efficiency (–)	
η_{AHT}	Absorber heat transformer efficiency (-)	E
η_{BOILER}	Boiler efficiency (–)	
η_C	Carnot efficiency (–)	EVAP
η_{CHP}	Compression heat pump efficiency (-)	EX
η_{COGEN}	Cogeneration efficiency (-)	final
η_F	Fin efficiency (–)	F
η_{GT}	Efficiency of gas turbine (-)	
η_{IS}	Isentropic efficiency of compression or	
	expansion (-)	FIN
η_{MECH}	Mechanical efficiency of steam turbine (-)	FG
η_P	Polytropic efficiency of compression or expansion (-)	G
η_{POWER}	Power generation efficiency (–)	H
η_{ST}	Efficiency of steam turbine (-)	HP
η_W	Weighted fin efficiency (-)	
θ	Angle (degrees), or	i
	fraction of feed permeated through membrane (-), or	
	root of the Underwood Equation (-)	in
θ	Logic variable in optimization (–)	Ι
λ	Ratio of latent heats of vaporization (-)	IF
λ_{ij}	Energy parameter characterizing the interaction of	IMP
	Molecule <i>i</i> with Molecule j (kJ · kmol ⁻¹)	IS
μ	Fluid viscosity (kg \cdot m ⁻¹ \cdot s ⁻¹ , mN \cdot s \cdot m ⁻² = cP)	J
π	Osmotic pressure $(N \cdot m^{-2}, bar)$	
ρ	Density $(\text{kg} \cdot \text{m}^{-3}, \text{kmol} \cdot \text{m}^{-3})$	Κ
σ	Stephan–Boltzmann constant ($W \cdot m^{-2} \cdot K^{-4}$), or	
	surface tension $(mN \cdot m^{-1} = mJ \cdot m^{-2} = dyne \cdot cm^{-1})$	L
τ	Reactor space time (s, min, h), or	LP
	residence time (s, min, h)	M
$ au_W$	Wall shear stress $(N \cdot m^{-2})$	max
v_k^i	Number of interaction Groups k in Molecule i (–)	min
Φ	Fugacity coefficient (-), or	М
	logic variable in optimization (-)	MIX
ω	Acentric factor (-)	Ν

Subscripts

axial	Axial direction
В	Blowdown, or
	bottoms in distillation
BFW	Boiler feedwater
BW	Bridgewall
cont	Contribution
С	Cold stream, or
	contaminant
CN	Condensing
COND	Condensing conditions
СР	Cold plane, or
	continuous phase
CW	Cooling water
D	Distillate in distillation
DS	De-superheating

е	Enhanced, or
	end zone on the shell side of a heat exchanger, or
	environment, or
	equivalent
Ε	Evaporation, or
	extract in liquid-liquid extraction
EVAP	Evaporator conditions
EX	Exhaust
final	Final conditions in a batch
\overline{F}	Feed. or
	final, or
	fluid
FIN	Fin on a finned tube
FG	Flue gas
G	Gas phase
H	Hot stream
HP	Heat pump, or
	high pressure
i	Component number, or
	stream number
in	Inlet
I	Inside
IF	Inside fouling
IMP	Impeller
IS	Isentropic
J	Component number, or
	stream number
Κ	Enthalpy interval number in heat exchanger
	networks
L	Liquid phase
LP	Low pressure
M	Stage number in distillation and absorption
max	Maximum
min	Minimum
M	Makeup
MIX	Mixture
N	Stage number in distillation and absorption
out	Outlet
0	Outside, or
	standard conditions
OF	Outside fouling
р	Stage number in distillation and absorption
prod	Products of reaction
Р	Particle, or
	permeate
PINCH	Pinch conditions
react	Reactants
R	Raffinate in liquid-liquid extraction
REACT	Reaction
ROOT	Root of a finned tube
S	Solvent in liquid-liquid extraction
SAT	Saturated conditions
SF	Supplementary firing
SUP	Superheated conditions
SW	Swirl direction
Т	Treatment

Te	Tube side enhanced
TW	Treated water
V	Vapor phase
w	Window section on the shell side of a heat
	exchanger
W	Conditions at the tube wall, or
	water
WBT	Wet bulb conditions
WW	Waste water
∞	Conditions at distillate pinch point

Superscripts

Ι	Phase I
II	Phase II
III	Phase III
IDEAL	Ideal behavior
L	Liquid phase
0	Standard conditions
V	Vapor phase
*	Adjusted parameter

Chapter 1

The Nature of Chemical Process Design and Integration

1.1 Chemical Products

Chemical products are essential to modern living standards. Almost all aspects of everyday life are supported by chemical products in one way or another. However, society tends to take these products for granted, even though a high quality of life fundamentally depends on them.

When considering the design of processes for the manufacture of chemical products, the market into which they are being sold fundamentally influences the objectives and priorities in the design. Chemical products can be divided into three broad classes:

- 1) *Commodity or bulk chemicals.* These are produced in large volumes and purchased on the basis of chemical composition, purity and price. Examples are sulfuric acid, nitrogen, oxygen, ethylene and chlorine.
- 2) Fine chemicals. These are produced in small volumes and purchased on the basis of chemical composition, purity and price. Examples are chloropropylene oxide (used for the manufacture of epoxy resins, ion-exchange resins and other products), dimethyl formamide (used, for example, as a solvent, reaction medium and intermediate in the manufacture of pharmaceuticals), *n*-butyric acid (used in beverages, flavorings, fragrances and other products) and barium titanate powder (used for the manufacture of electronic capacitors).
- **3)** Specialty or effect or functional chemicals. These are purchased because of their effect (or function), rather than their chemical composition. Examples are pharmaceuticals, pesticides, dyestuffs, perfumes and flavorings.

Because commodity and fine chemicals tend to be purchased on the basis of their chemical composition alone, they can be considered to be *undifferentiated*. For example, there is nothing to choose between 99.9% benzene made by one manufacturer and that made by another manufacturer, other than price and delivery issues. On the other hand, specialty chemicals tend to be purchased on the basis of their effect or function and therefore can be considered to be *differentiated*. For example, competitive pharmaceutical products are differentiated according to the efficacy of the product, rather than chemical composition. An adhesive is purchased on the basis of its ability to stick things together, rather than its chemical composition, and so on.

However, in practice few products are completely undifferentiated and few completely differentiated. Commodity and fine chemical products might have impurity specifications as well as purity specifications. Traces of impurities can, in some cases, give some differentiation between different manufacturers of commodity and fine chemicals. For example, 99.9% acrylic acid might be considered to be an undifferentiated product. However, traces of impurities, at concentrations of a few parts per million, can interfere with some of the reactions in which it is used and can have important implications for some of its uses. Such impurities might differ between different manufacturing processes. Not all specialty products are differentiated. For example, pharmaceutical products like aspirin (acetylsalicylic acid) are undifferentiated. Different manufacturers can produce aspirin, and there is nothing to choose between these products, other than the price and differentiation created through marketing of the product. Thus, the terms undifferentiated and differentiated are more relative than absolute terms.

The scale of production also differs between the three classes of chemical products. Fine and specialty chemicals tend to be produced in volumes less than $1000 \text{ t} \cdot \text{y}^{-1}$. By contrast, commodity chemicals tend to be produced in much larger volumes than this. However, the distinction is again not so clear. Polymers are differentiated products because they are purchased on the basis of their mechanical properties, but can be produced in quantities significantly higher than $1000 \text{ t} \cdot \text{y}^{-1}$.

Chemical Process Design and Integration, Second Edition. Robin Smith. © 2016 John Wiley & Sons, Ltd. Published 2016 by John Wiley & Sons, Ltd. Companion Website: www.wiley.com/go/smith/chemical2e

When a new chemical product is first developed, it can often be protected by a patent in the early years of its commercial exploitation. For a product to be eligible to be patented, it must be novel, useful and unobvious. If patent protection can be obtained, this effectively gives the producer a monopoly for commercial exploitation of the product until the patent expires. Patent protection lasts for 20 years from the filing date of the patent. Once the patent expires, competitors can join in and manufacture the product. If competitors cannot wait until the patent expires, then alternative competing products must be developed.

Another way to protect a competitive edge for a new product is to protect it by secrecy. The formula for Coca-Cola has been kept a secret for over 100 years. Potentially, there is no time limit on such protection. However, for the protection through secrecy to be viable, competitors must not be able to reproduce the product from chemical analysis. This is likely to be the case only for certain classes of specialty chemicals and food products for which the properties of the product depend on both the chemical composition and the method of manufacture.

Figure 1.1 illustrates different product *life cycles* (Sharratt, 1997; Brennan, 1998). The general trend is that when a new product is introduced into the market, the sales grow slowly until the market is established and then more rapidly once the market is established. If there is patent protection, then competitors will not be able to exploit the same product commercially until the patent expires, when competitors can produce the same product and take market share. It is expected that competitive products will cause sales to diminish later in the product life cycle until sales become so low that a company would be expected to withdraw from the market. In Figure 1.1, Product *A* appears to be a poor



Figure 1.1

Product life cycles. (Adapted from Sharratt PN, 1997, Handbook of Batch Process Design, Chapman & Hall, reproduced by permission.) product that has a short life with low sales volume. It might be that it cannot compete well with other competitive products and alternative products quickly force the company out of that business. However, a low sales volume is not the main criterion to withdraw a product from the market. It might be that a product with low volume finds a market niche and can be sold for a high value. On the other hand, if it were competing with other products with similar functions in the same market sector, which keeps both the sale price and volume low, then it would seem wise to withdraw from the market. Product B in Figure 1.1 appears to be a better product, showing a longer life cycle and higher sales volume. This has patent protection but sales decrease rapidly after patent protection is lost, leading to loss of market through competition. Product C in Figure 1.1 is an even better product. This shows high sales volume with the life of the product extended through reformulation of the product (Sharratt, 1997). Finally, Product D in Figure 1.1 shows a product life cycle that is typical of commodity chemicals. Commodity chemicals tend not to exhibit the same kind of life cycles as fine and specialty chemicals. In the early years of the commercial exploitation, the sales volume grows rapidly to a high volume, but then volume does not decline and enters a mature period of slow growth, or, in some exceptional cases, slow decline. This is because commodity chemicals tend to have a diverse range of uses. Even though competition might take away some end uses, new end uses are introduced, leading to an extended life cycle.

The different classes of chemical products will have very different *added value* (the difference between the selling price of the product and the purchase cost of raw materials). Commodity chemicals tend to have low added value, whereas fine and specialty chemicals tend to have high added value. Commodity chemicals tend to be produced in large volumes with low added value, while fine and specialty chemicals tend to be produced in small volumes with high added value.

Because of this, when designing a process for a commodity chemical, it is usually important to keep operating costs as low as possible. The capital cost of the process will tend to be high relative to a process for fine or specialty chemicals because of the scale of production.

When designing a process for specialty chemicals, priority tends to be given to the product, rather than to the process. This is because the unique function of the product must be protected. The process is likely to be small scale and operating costs tend to be less important than with commodity chemical processes. The capital cost of the process will be low relative to commodity chemical processes because of the scale. The time to market for the product is also likely to be important with specialty chemicals, especially if there is patent protection. If this is the case, then anything that shortens the time from basic research, through product testing, pilot plant studies, process design, construction of the plant to product manufacture will have an important influence on the overall project profitability.

All this means that the priorities in process design are likely to differ significantly, depending on whether a process is being designed for the manufacture of a commodity, fine or specialty chemical. In commodity chemicals, there is likely to be relatively little product innovation, but intensive process innovation. Also, equipment will be designed for a specific process step. On the other hand, the manufacture of fine and specialty chemicals might involve:

- selling into a market with low volume;
- a short product life cycle;
- a demand for a short time to market, and therefore less time is available for process development, with product and process development proceeding simultaneously.

As a result, the manufacture of fine and specialty chemicals is often carried out in multipurpose equipment, perhaps with different chemicals being manufactured in the same equipment at different times during the year. The life of the equipment might greatly exceed the life of the product.

The development of pharmaceutical products demands that high-quality products must be manufactured during the development of the process to allow safety and clinical studies to be carried out before full-scale production. Pharmaceutical production represents an extreme case of process design in which the regulatory framework controlling production makes it difficult to make process changes, even during the development stage. Even if significant improvements to processes for pharmaceuticals can be suggested, it might not be feasible to implement them, as such changes might prevent or delay the process from being licensed for production.

1.2 Formulation of Design Problems

Before a process design can be started, the design problem must be formulated. Formulation of the design problem requires a product specification. If a well-defined chemical product is to be manufactured, then the specification of the product might appear straightforward (e.g. a purity specification). However, if a specialty product is to be manufactured, it is the functional properties that are important, rather than the chemical properties, and this might require a *product design* stage in order to specify the product (Seider *et al.*, 2010; Cussler and Moggridge, 2011).

The initial statement of the design problem is often ill defined. For example, the design team could be asked to expand the production capacity of an existing plant that produces a chemical that is a precursor to a polymer product, which is also produced by the company. This results from an increase in the demand for the polymer product and the plant producing the precursor currently being operated at its maximum capacity. The design team might well be given a specification for the expansion. For example, the marketing department might assess that the market could be expanded by 30% over a two-year period, which would justify a 30% expansion in the process for the precursor. However, the 30% projection can easily be wrong. The economic environment can change, leading to the projected increase being either too large or too small. It might also be possible to sell the polymer precursor in the market to other manufacturers of the polymer and justify an expansion even larger than 30%. If the polymer precursor can be sold in the marketplace, is the current purity specification of the company suitable for the marketplace? Perhaps the marketplace demands a higher purity than the current company specification. Perhaps the current specification is acceptable, but if the specification could be improved, the product could be sold for a higher value and/or at a greater volume. An option might be to not expand the production of the polymer precursor to 30%, but instead to purchase it from the market. If it is purchased from the market, is it likely to be up to the company specifications or will it need some purification before it is suitable for the company's polymer process? How reliable will the market source be? All these uncertainties are related more to market supply and demand issues than to specific process design issues.

Closer examination of the current process design might lead to the conclusion that the capacity can be expanded by 10% with a very modest capital investment. A further increase to 20% would require a significant capital investment, but an expansion to 30% would require an extremely large capital investment. This opens up further options. Should the plant be expanded by 10% and a market source identified for the balance? Should the plant be expanded to 20% similarly? If a real expansion in the marketplace is anticipated and expansion to 30% would be very expensive, why not be more aggressive and, instead of expanding the existing process, build an entirely new process? If a new process is to be built, then what should be the process technology? New process technology might have been developed since the original plant was built that enables the same product to be manufactured at a much lower cost. If a new process is to be built, where should it be built? It might make more sense to build it in another country that would allow lower operating costs, and the product could be shipped back to be fed to the existing polymer process. At the same time, this might stimulate the development of new markets in other countries, in which case, what should be the capacity of the new plant?

Thus, from the initial ill-defined problem, the design team must create a series of very specific options and these should then be compared on the basis of a common set of assumptions regarding, for example, raw materials and product prices. Having specified an option, this gives the design team a well-defined problem to which the methods of engineering and economic analysis can be applied.

In examining a design option, the design team should start out by examining the problem at the highest level, in terms of its feasibility with the minimum of detail to ensure the design option is worth progressing (Douglas, 1985). Is there a large difference between the value of the product and the cost of the raw materials? If the overall feasibility looks attractive, then more detail can be added, the option re-evaluated, further detail added, and so on. Byproducts might play a particularly important role in the economics. It might be that the current process produces some byproducts that can be sold in small quantities to the market. However, as the process is expanded, there might be market constraints for the new scale of production. If the byproducts cannot be sold, how does this affect the economics?

In summary, the original problem posed to process design teams is often ill defined, even though it might appear to be well defined in the original design specification. The design team must then formulate a series of plausible design options to be screened by the methods of engineering and economic analysis. These design options are formulated into very specific design problems. In this way, the design team turns the ill-defined problem into a series of well-defined design options for analysis.

1.3 Synthesis and Simulation

In a chemical process, the transformation of raw materials into desired chemical products usually cannot be achieved in a single step. Instead, the overall transformation is broken down into a number of steps that provide intermediate transformations. These are carried out through reaction, separation, mixing, heating, cooling, pressure change, particle size reduction or enlargement for solids. Once individual steps have been selected, they must be interconnected to carry out the overall transformation (Figure 1.2a). Thus, the *synthesis* of a chemical process involves two broad activities. First, individual transformation steps are selected. Second, these individual transformations are interconnected to form a complete process that achieves the required overall transformation. A *flowsheet* or *process flow diagram* (PFD) is a diagrammatic representation of the process steps with their interconnections.

Once the flowsheet structure has been defined, a *simulation* of the process can be carried out. A simulation is a mathematical model of the process that attempts to predict how the process would behave if it was constructed (Figure 1.2b). Material and energy balances can be formulated to give better definition to the inner workings of the process and a more detailed process design can be developed. Having created a model of the process, the flowrates, compositions, temperatures and pressures of the feeds can be



(a) Process design starts with the synthesis of a process to convert raw materials into desired products.



(b) Simulation predicts how a process would behave if it was constructed.

Figure 1.2

Synthesis is the creation of a process to transform feed streams into product streams. Simulation predicts how it would behave if it was constructed. assumed. The simulation model then predicts the flowrates, compositions, temperatures, pressures and properties of the products. It also allows the individual items of equipment in the process to be sized and predicts, for example, how much raw material is being used or how much energy is being consumed. The performance of the design can then be evaluated.

1) Accuracy of design calculations. A simulation adds more detail once a design has been synthesized. The design calculations for this will most often be carried out in a general purpose simulation software package and solved to a high level of precision. However, a high level of precision cannot usually be justified in terms of the operation of the plant after it has been built. The plant will almost never work precisely at its original design flowrates, temperatures, pressures and compositions. This might be because the raw materials are slightly different from what is assumed in the design. The physical properties assumed in the calculations might have been erroneous in some way, or operation at the original design conditions might create corrosion or fouling problems, or perhaps the plant cannot be controlled adequately at the original conditions, and so on, for a multitude of other possible reasons. The instrumentation on the plant will not be able to measure the flowrates, temperatures, pressures and compositions as accurately as the calculations performed. High precision might be required in the calculations for certain specific parts of the design. For example, a polymer precursor might need certain impurities to be very tightly controlled, perhaps down to the level of parts per million, or it might be that some contaminant in a waste stream might be exceptionally environmentally harmful and must be extremely well defined in the design calculations.

Even though a high level of precision cannot be justified in many cases in terms of the plant operation, the design calculations will normally be carried out to a reasonably high level of precision. The value of precision in design calculations is that the consistency of the calculations can be checked to allow errors or poor assumptions to be identified. It also allows the design options to be compared on a valid like-for-like basis.

Because of all the uncertainties in carrying out a design, the specifications are often increased beyond those indicated by the design calculations and the plant is *overdesigned*, or contingency is added, through the application of safety factors to the design. For example, the designer might calculate the number of distillation plates required for a distillation separation using elaborate calculations to a high degree of precision, only to add an arbitrary extra 10% to the number of plates for contingency. This allows for the feed to the unit not being exactly as specified, errors in the physical properties, upset conditions in the plant, control requirements, and so on. If too little contingency is added, the plant might not work. If too much contingency is added, the plant will not only be unnecessarily expensive but too much overdesign might make the plant difficult to operate and might lead to a less efficient plant. For example, the designer might calculate the size of a heat exchanger and then add in a large contingency and significantly oversize the heat exchanger. The lower fluid velocities encountered by the oversized heat exchanger can

cause it to have a poorer performance and to foul-up more readily than a smaller heat exchanger.

Too little overdesign might lead to the plant not working. Too much overdesign will lead to the plant becoming unnecessarily expensive, and perhaps difficult to operate and less efficient. A balance must be made between different risks.

- 2) Physical properties in process design. Almost all design calculations require physical properties of the solids, liquids and gases being fed, processed and produced. Physical properties can be critical to obtaining meaningful, economic and safe designs. When carrying out calculations in computer software packages there is most often a choice to be made for the physical property correlations and data. However, if poor decisions are made by the designer regarding physical properties, the design calculations can be meaningless or even dangerous, even though the calculations have been performed to a high level of precision. Using physical property correlations outside the ranges of conditions for which they were intended can be an equally serious problem. Appendix A discusses physical properties in process design in more detail.
- 3) Evaluation of performance. There are many facets to the evaluation of performance. Good economic performance is an obvious first criterion, but it is certainly not the only one. Chemical processes should be designed to maximize the sustainability of industrial activity. Maximizing sustainability requires that industrial systems should strive to satisfy human needs in an economically viable, environmentally benign and socially beneficial way (Azapagic, 2014). For chemical process design, this means that processes should make use of materials of construction that deplete the resource as little as practicable. Process raw materials should be used as efficiently as is economic and practicable, both to prevent the production of waste that can be environmentally harmful and to preserve the reserves of manufacturing raw materials as much as possible. Processes should use as little energy as is economic and practicable, both to prevent the build-up of carbon dioxide in the atmosphere from burning fossil fuels and to preserve the reserves of fossil fuels. Water must also be consumed in sustainable quantities that do not cause deterioration in the quality of the water source and the long-term quantity of the reserves. Aqueous and atmospheric emissions must not be environmentally harmful and solid waste to landfill must be avoided. The boundary of consideration should go beyond the immediate boundary of the manufacturing facility to maximize the benefit to society to avoid adverse health effects, unnecessarily high burdens on transportation, odour, noise nuisances, and so on.

The process must also meet required health and safety criteria. Start-up, emergency shutdown and ease of control are other important factors. Flexibility, that is, the ability to operate under different conditions, such as differences in feedstock and product specification, may be important. Availability, that is, the portion of the total time that the process meets its production requirements, might also be critically important. Uncertainty in the design, for example, resulting from poor design data, or uncertainty in the economic data might guide the design away from certain options. Some of these factors, such as economic performance, can be readily quantified; others, such as safety, often cannot. Evaluation of the factors that are not readily quantifiable, the intangibles, requires the judgment of the design team.

- **4)** *Materials of construction.* Choice of materials of construction affects both the mechanical design and the capital cost of equipment. Many factors enter into the choice of the materials of construction. Among the most important are (see Appendix B):
 - mechanical properties (particularly yield and tensile strength, compressive strength, ductility, toughness, hardness, fatigue limit and creep resistance);
 - effect of temperature on mechanical properties (both low and high temperatures),
 - ease of fabrication (machining, welding, and so on);
 - corrosion resistance;
 - availability of standard equipment in the material;
 - cost (e.g. if materials of construction are particularly expensive, it might be desirable to use a cheaper material together with a lining on the process side to reduce the cost).

Estimation of the capital cost and preliminary specification of equipment for the evaluation of performance requires decisions to be made regarding the materials of construction. The discussion of the more commonly used materials of construction is given in Appendix B.

- 5) Process safety. When evaluating a process design, process safety should be the prime consideration. Safety considerations must not be left until the design has been completed. Safety systems need to be added to the design later for the relief of overpressure, to trip the process under dangerous conditions, etc. However, by far the largest impact on process safety can be made early in the design through measures to make the design inherently safer. This will be discussed in detail in Chapter 28. Inherently safer design means avoiding the need for hazardous materials if possible, or using less of them, or using them at lower temperatures and pressures or diluting them with inert materials. One of the principal approaches to making a process inherently safer is to limit the inventory of hazardous material. The inventories to be avoided most of all are flashing flammable or toxic liquids, that is, liquids under pressure above their atmospheric boiling points (see Chapter 28).
- 6) *Optimization*. Once the basic performance of the design has been evaluated, changes can be made to improve the performance; the process is *optimized*. These changes might involve the synthesis of alternative structures, that is, *structural optimization*. Thus, the process is simulated and evaluated again, and so on, optimizing the structure. Each structure can be subjected to *parameter optimization* by changing operating conditions within that structure. This is illustrated in Figure 1.3.



Optimization can be carried out as structural or parameter optimization to improve the evaluation of the design.

From the project definition an initial design is synthesized. This can then be simulated and evaluated. Once evaluated, the design can be optimized in a parameter optimization through changing the continuous parameters of flowrate, composition, temperature and pressure to improve the evaluation. However, this parameter optimization only optimizes the initial design configuration, which might not be an optimal configuration. So the design team might return to the synthesis stage to explore other configurations in a structural optimization. Also, if the parameter optimization adjusts the settings of the conditions to be significantly different from the original assumptions, then the design team might return to the synthesis stage to consider other configurations in the structural optimization. The different ways this design process can be followed will be considered later in this chapter.

7) Keeping design options open. To develop a design concept requires design options to be first generated and then evaluated. There is a temptation to carry out preliminary evaluation early in the development of a design and eliminate options early that initially appear to be unattractive. However, this temptation must be avoided. In the early stages of a design the uncertainties in the evaluation are often too serious for early elimination of options, unless it is absolutely clear that a design option is not viable. Initial cost estimates can be very misleading and the full safety and environmental implications of early decisions are only clear once detail has been added. If it was possible to foresee everything that lay ahead, decisions made early might well be different. There is a danger in focusing on one option without rechecking the assumptions later for validity when more information is available. The design team must not be boxed in early by preconceived ideas. This means that design options should be left open as long as practicable until it is clear that options can be closed down. All options should be considered, even if they appear unappealing at first.

1.4 The Hierarchy of Chemical Process Design and Integration

Consider the process illustrated in Figure 1.4 (Smith and Linnhoff, 1988). The process requires a reactor to transform the FEED into PRODUCT (Figure 1.4a). Unfortunately, not all the FEED reacts. Also, part of the FEED reacts to form BYPRODUCT instead of the desired PRODUCT. A separation system is needed to isolate the *PRODUCT* at the required purity. Figure 1.4b shows one possible separation system consisting of two distillation columns. The unreacted FEED in Figure 1.4b is recycled and the PRODUCT and BYPRODUCT are removed from the process. Figure 1.4b shows a flowsheet where all heating and cooling is provided by external utilities (steam and cooling water in this case). This flowsheet is probably too inefficient in its use of energy and heat should be recovered. Thus, heat integration is carried out to exchange heat between those streams that need to be cooled and those that need to be heated. Figure 1.5 (Smith and Linnhoff, 1988) shows two possible designs for the *heat exchanger network*, but many other heat integration arrangements are possible.

The flowsheets shown in Figure 1.5 feature the same reactor design. It could be useful to explore the changes in reactor design. For example, the size of the reactor could be increased to increase the amount of FEED that reacts (Smith and Linnhoff, 1988). Now there is not only much less FEED in the reactor effluent but also more PRODUCT and BYPRODUCT. However, the increase in BYPRODUCT is larger than the increase in PRODUCT. Thus, although the reactor has the same three components in its effluent as the reactor in Figure 1.4a, there is less FEED, more PRODUCT and significantly more BYPRODUCT. This change in reactor design generates a different task for the separation system and it is possible that a separation system different from that shown in Figures 1.4 and 1.5 is now appropriate. Figure 1.6 shows a possible alternative. This also uses two distillation columns, but the separations are carried out in a different order.

Figure 1.6 shows a flowsheet without any heat integration for the different reactor and separation system. As before, this is probably too inefficient in the use of energy, and heat integration schemes can be explored. Figure 1.7 (Smith and Linnhoff, 1988) shows two of the many possible flowsheets involving heat recovery.

Different complete flowsheets can be evaluated by simulation and costing. On this basis, the flowsheet in Figure 1.5b might be more promising than the flowsheets in Figures 1.5a and 1.7a and b. However, the best flowsheet cannot be identified without first optimizing the operating conditions for each. The flowsheet in Figure 1.7b might have greater scope for improvement than that in Figure 1.5b, and so on.

Thus, the complexity of chemical process synthesis is twofold. First, can all possible structures be identified? It might be considered that all the structural options can be found by inspection, at least all of the significant ones. The fact that



(a) A reactor transforms FEED into PRODUCT and BYPRODUCT.



(b) To isolate the *PRODUCT* and recycle unreacted *FEED* we need a separation system.

Figure 1.4

Process design starts with the reactor. The reactor design dictates the separation and recycle problem. (Reproduced from Smith R and Linnhoff B, 1998, *Trans IChemE ChERD*, **66**: 195 by permission of the Institution of Chemical Engineers.)

even long-established processes are still being improved bears evidence to just how difficult this is. Second, can each structure be optimized for a valid comparison? When optimizing the structure, there may be many ways in which each individual task can be performed and many ways in which the individual tasks can be interconnected. This means that the operating conditions for a multitude of structural options must be simulated and optimized. At first sight, this appears to be an overwhelmingly complex problem.

It is helpful when developing a methodology if there is a clear picture of the nature of the problem. If the process requires a reactor, this is where the design starts. This is likely to be the only place in the process where raw material components are converted into components for the products. The chosen reactor design produces a mixture of unreacted feed materials, products and byproducts that need separating. Unreacted feed material is recycled. The reactor design dictates the separation and recycle problem. Thus, design of the separation and recycle system follows the reactor design. The reactor and separation and recycle system designs together define the process for heating and cooling duties. Thus, the heat exchanger network design comes next. Those heating and cooling duties that cannot be satisfied by heat recovery dictate the need for external heating and cooling *utilities* (furnace heating, use of steam, steam generation, cooling water, air cooling or refrigeration). Thus, utility selection and design follows the design of the heat recovery system. The selection and design of the utilities is made more complex by the fact that the process will most likely operate within the context of a site comprising a number of different processes that are all connected to a common utility system. The process and the utility system will both need water, for example, for steam generation, and will also produce aqueous effluents that will have to be brought to a suitable quality for discharge. Thus, the design of the water and aqueous effluent treatment system comes last. Again, the water and effluent treatment system must be considered at the site level as well as the process level.

This hierarchy can be represented symbolically by the layers of the "onion diagram" shown in Figure 1.8 (Linnhoff *et al.*, 1982). The diagram emphasizes the sequential, or hierarchical, nature of process design. Other ways to represent the hierarchy have also been suggested (Douglas, 1985).

Some processes do not require a reactor, for example, some processes just involve separation. Here, the design starts with the separation system and moves outward to the heat exchanger network, utilities, and so on. However, the same basic hierarchy prevails.



For a given reactor and separator design there are different possibilities for heat integration. (Reproduced from Smith R and Linnhoff B, 1998, *Trans IChemE ChERD*, **66**: 195 by permission of the Institution of Chemical Engineers.)

The synthesis of the correct structure and the optimization of parameters in the design of the reaction and separation systems are often the most important tasks of process design. Usually there are many options, and it is impossible to fully evaluate them unless a complete design is furnished for the "outer layers" of the onion model. For example, it is not possible to assess which is better, the basic scheme from Figure 1.4b or that from Figure 1.6, without fully evaluating all possible designs, such as those shown in Figures 1.5a and b and 1.7a and b, all completed, including utilities. Such a complete search is normally too timeconsuming to be practical.

Later, in Chapter 17, an approach will be presented in which some early decisions (i.e. decisions regarding reactor and separator options) can be evaluated without a complete design for the "outer layers".

1.5 Continuous and Batch Processes

When considering the processes in Figures 1.4 to 1.6, an implicit assumption was made that the processes operated continuously. However, not all processes operate continuously. In a *batch* process, the main steps operate discontinuously. In contrast with a continuous process, a batch process does not deliver its product continuously but in discrete amounts. This means that heat, mass, temperature, concentration and other properties vary with time. In practice, most batch processes are made up of a series of batch and *semi-continuous* steps. A semicontinuous step runs continuously with periodic start-ups and shutdowns.



Changing the reactor dictates a different separation and recycle problem. (Reproduced from Smith R and Linnhoff B, 1998, *Trans IChemE ChERD*, **66**: 195 by permission of the Institution of Chemical Engineers.)



Figure 1.7

A different reactor design not only leads to a different separation system but additional possibilities for heat integration. (Reproduced from Smith R and Linnhoff B, 1998, *Trans IChemE ChERD*, **66**: 195 by permission of the Institution of Chemical Engineers.)



The onion model of process design. A reactor is needed before the separation and recycle system can be designed, and so on.

Consider the simple process shown in Figure 1.9. Feed material is withdrawn from storage using a pump. The feed material is preheated in a heat exchanger before being fed to a batch reactor. Once the reactor is full, further heating takes place inside the reactor by passing steam into the reactor jacket before the reaction proceeds. During the later stages of the reaction, cooling water is applied to the reactor jacket. Once the reaction is complete, the reactor product is withdrawn using a pump. The reactor product is cooled in a heat exchanger before going to storage.

The first two steps, pumping for reactor filling and feed preheating, are both semi-continuous. The heating inside the reactor, the reaction itself and the cooling using the reactor jacket are all batch. The pumping to empty the reactor and the productcooling step are again semi-continuous. The hierarchy in batch process design is no different from that in continuous processes and the hierarchy illustrated in Figure 1.8 prevails for batch processes also. However, the time dimension brings constraints that do not present a problem in the design of continuous processes. For example, heat recovery might be considered for the process in Figure 1.9. The reactor effluent (which requires cooling) could be used to preheat the incoming feed to the reactor (which requires heating). Unfortunately, even if the reactor effluent is at a high enough temperature to allow this, the reactor feeding and emptying take place at different times, meaning that this will not be possible without some way to store the heat. Such heat storage is possible but usually uneconomic, especially for small-scale processes.

If a batch process manufactures only a single product, then the equipment can be designed and optimized for that product. The dynamic nature of the process creates additional challenges for design and optimization. It might be that the optimization calls for variations in the conditions during the batch through time, according to some *profile*. For example, the temperature in a batch reactor might need to be increased or decreased as the batch progresses.

Multiproduct batch processes, with a number of different products manufactured in the same equipment, present even bigger challenges for design and optimization (Biegler, Grossman and Westerberg, 1997). Different products will demand different designs, different operating conditions and, perhaps, different trajectories for the operating conditions through time. The design of equipment for multiproduct plants will thus require a compromise to be made across the requirements of a number of different products. The more flexible the equipment and the configuration of the equipment, the more it will be able to adapt to the optimum requirements of each product.

Batch processes:

- are economical for small volumes;
- are flexible in accommodating changes in product formulation;
- are flexible in changing the production rate by changing the number of batches made in any period of time;


- allow the use of standardized multipurpose equipment for the production of a variety of products from the same plant;
- are best if equipment needs regular cleaning because of fouling or needs regular sterilization;
- are amenable to direct scale-up from the laboratory and
- allow product identification. Each batch of product can be clearly identified in terms of when it was manufactured, the feeds involved and conditions of processing. This is particularly important in industries such as pharmaceuticals and foodstuffs. If a problem arises with a particular batch, then all the products from that batch can be identified and withdrawn from the market. Otherwise, all the products available in the market would have to be withdrawn.

One of the major problems with batch processing is batch-tobatch conformity. Minor changes to the operation can mean slight changes in the product from batch to batch. Fine and specialty chemicals are usually manufactured in batch processes. However, these products often have very tight tolerances for impurities in the final product and demand batch-to-batch variation to be minimized.

Batch processes will be considered in more detail in Chapter 16.

1.6 New Design and Retrofit

There are two situations that can be encountered in process design. The first is in the design of *new plant* or *grassroot* design. In the second, the design is carried out to modify an existing plant in *retrofit* or *revamp*. The motivation to retrofit an existing plant could be, for example, to increase capacity, allow for different feed or product specifications, reduce operating costs, improve safety or reduce environmental emissions. One of the most common motivations is to increase capacity. When carrying out a retrofit, whatever the motivation, it is desirable to try and make as effective use as possible of the existing equipment. The basic problem with this is that the design of the existing equipment might not be ideally suited to the new role that it will be put to. On the other hand, if equipment is reused, it will avoid unnecessary investment in new equipment, even if it is not ideally suited to the new duty.

When carrying out a retrofit, the connections between the items of equipment can be reconfigured, perhaps adding new equipment where necessary. Alternatively, if the existing equipment differs significantly from what is required in the retrofit, then in addition to reconfiguring the connections between the equipment, the equipment itself can be modified. Generally, the fewer the modifications to both the connections and the equipment, the better.

The most straightforward design situations are those of grassroot design as it has the most freedom to choose the design options and the size of equipment. In retrofit, the design must try to work within the constraints of existing equipment. Because of this, the ultimate goal of the retrofit design is often not clear. For example, a design objective might be given to increase the capacity of a plant by 50%. At the existing capacity limit of the plant, at least one item of equipment must be at its maximum capacity. Most items of equipment are likely to be below their maximum capacity. The differences in the spare capacity of different items of equipment in the existing design arises from different design allowances (or contingency) in the original design, changes to the operation of the plant relative to the original design, errors in the original design data, and so on. An item of equipment at its maximum capacity is the *bottleneck* to prevent increased capacity. Thus, to overcome the bottleneck or *debottleneck*, the item of equipment is modified, or replaced with new equipment with increased capacity, or a new item is placed in parallel or series with the existing item, or the connections between existing equipment are reconfigured, or a combination of all these actions is taken. As the capacity of the plant is increased, different items of equipment will reach their maximum capacity. Thus, there will be thresholds in the plant capacity created by the limits in different items of equipment. All equipment with capacity less than the threshold must be modified in some way, or the plant reconfigured, to overcome the threshold. To overcome each threshold requires capital investment. As capacity is increased from the existing limit, ultimately it is likely that it will be prohibitive for the investment to overcome one of the design thresholds. This is likely to become the design limit, as opposed to the original remit of a 50% increase in capacity in the example.

Another important issue in retrofit is the downtime required to make the modifications. The cost of lost production while the plant is shut down to be modified can be prohibitively expensive. Thus, one of the objectives for retrofit is to design modifications that require only a short shutdown. This often means designing modifications that allow the bulk of the work to be carried out while the process is still in operation. For example, new equipment can be installed with final piping connections made when the process is shut down. Decisions whether to replace a major process component completely, or to supplement with a new component working in series or parallel with the existing component, can be critical to the downtime required for retrofit.

1.7 Reliability, Availability and Maintainability

As already discussed, *availability* is often an important issue in process design. Unless the plant is operating in its intended way, it is not productive. Availability measures the portion of the total time that the process meets its production requirements. Availability is related to *reliability* and *maintainability*. *Reliability* is the probability of survival after the unit/system operates for a certain period of time (e.g. a unit has a 95% probability of survival after 8000 hours). Reliability defines the failure frequency and determines the *uptime* patterns. *Maintainability* describes how long it takes for the unit/system to be repaired, which determines the *downtime* patterns. Availability measures the percentage of uptime the process operates at its production requirements over the time horizon, and is determined by reliability and maintainability.

Availability can be improved in many ways. Maintenance policy has a direct influence. *Preventive maintenance* can be used to prevent unnecessary breakdowns. *Condition monitoring* of equipment using techniques such as monitoring vibration of rotating equipment like compressors can be used to detect mechanical problems early, and again prevent unnecessary breakdowns. In design, using *standby* components (sometimes referred to as *spare* or *redundant* components) is a common way to increase system availability. Instead of having one item of equipment on line and vulnerable to breakdown, there may be two, with one online and one off-line. These two items of equipment can be sized and operated in many ways:

- $2 \times 100\%$ one on-line, one off-line switched off;
- $2 \times 100\%$ one on-line, one off-line idling;
- 2×50% both on-line, with system capacity reduced to 50% if one fails;
- 2×75% both on-line operating at 2/3 capacity when both operating, but with system capacity 75% if one fails;
- and so on.

Over-sizing equipment, particularly rotating equipment like pumps and compressors, can make it more reliable in some cases. Determining the optimum policy for standby equipment involves complex trade-offs that need to consider capital cost, operating cost, maintenance costs and reliability.

1.8 Process Control

Once the basic process configuration has been fixed, a *control system* must be added. The control system compensates for the influence of external *disturbances* such as changes in feed flowrate, feed conditions, feed costs, product demand, product specifications, product prices, ambient temperature, and so on. Ensuring safe operation is the most important task of a control system. This is achieved by monitoring the process conditions and maintaining them within safe operating limits. While maintaining the operation within safe operating limits, the control system should optimize the process performance under the influence of external disturbances. This involves maintaining product specifications, meeting production targets and making efficient use of raw materials and utilities.

A control mechanism is introduced that makes changes to the process in order to cancel out the negative impact of disturbances. In order to achieve this, instruments must be installed to measure the operational performance of the plant. These *measured variables* could include temperature, pressure, flowrate, composition, level, pH, density and particle size. *Primary measurements* may be made to directly represent the control objectives (e.g. measuring the composition that needs to be controlled). If the control objectives are not measurable, then *secondary measurements* of other variables must be made and these secondary measurements related to the control objective. Having measured the variables that need to be controlled, other variables need to be *manipulated* in order to achieve the control objectives. A control system is then designed that responds to variations in the measured variables and manipulates other variables to control the process.

Having designed a process configuration for a continuous process and having optimized it to achieve some objective (e.g. maximize profit) at steady state, is the influence of the control system likely to render the previously optimized process to now be nonoptimal? Even for a continuous process, the process is always likely to be moving from one state to another in response to the influence of disturbances and control objectives. In the steadystate design and optimization of continuous processes, these different states can be allowed for by considering *multiple operating cases*. Each operating case is assumed to operate for a certain proportion of the year. The contribution of the operating case to the overall steady-state design and optimization is weighted according to the proportion of the time under which the plant operates at that state.

While this takes some account of operation under different conditions, it does not account for the dynamic transition from one state to another. Are these transitory states likely to have a significant influence on the optimality? If the transitory states were to have a significant effect on the overall process performance in terms of the objective function being optimized, then the process design and control system design would have to be carried out simultaneously. Simultaneous design of the process and the control system presents an extremely complex problem. It is interesting to note that where steady-state optimization for continuous processes has been compared with simultaneous optimization of the process and control system, the two process designs have been found to be almost identical (Bansal *et al.*, 2000a, 2000b; Kookos and Perkins, 2001).

Industrial practice is to first design and optimize the process configuration (taking into account multiple states, if necessary) and then to add the control system. However, there is no guarantee that design decisions made on the basis of steady-state conditions will not lead to control problems once process dynamics are considered. For example, an item of equipment might be oversized for contingency, because of uncertainty in design data or future debottlenecking prospects, based on steady-state considerations. Once the process dynamics are considered, this oversized equipment might make the process difficult to control, because of the large inventory of process materials in the oversized equipment. The approach to process control should adopt an approach that considers the control of the whole process, rather than just the control of the individual process steps in isolation (Luyben, Tyreus and Luyben, 1999).

The control system arrangement is shown in the *piping and* instrumentation diagram (P & I D) for the process (Sinnott and Towler, 2009). The piping and instrumentation diagram shows all of the process equipment, piping connections, valves, pipe fittings and the control system. All equipment and connections are shown. This includes not only the main items of equipment and connections but also standby equipment, equipment and piping used for start-up, shutdown, maintenance operations and abnormal operation. Figure 1.10a illustrates a very simple process flow diagram. This shows only the main items of equipment and the normal process flows. The information shown on such process flow diagrams, and their style, vary according to the practice of different companies. As an example, Figure 1.10a shows the component flowrates and stream temperatures and pressures. By contrast Figure 1.10b shows the corresponding piping and instrumentation diagram. This shows all of the equipment (including the standby pump in this case), all piping connections and fittings, including



Figure 1.10

Ð

Process flow diagrams (PFD) and piping and instrumentation diagrams (P&ID).

those used for start-up, shutdown, maintenance and abnormal operation. It also shows the layout of the control system. Additional information normally included would be identification numbers for the equipment, piping connections and control equipment. Information on materials of construction might also be included. But information on process flows and conditions would not normally be shown. As with process flow diagrams, the style of piping and instrumentation diagrams varies according to the practice of different companies.

This text will concentrate on the design and optimization of the process configuration and will not deal with process control. Process control demands expertise in different techniques and will be left to other sources of information (Luyben, Tyreus and Luyben, 1999). Thus, the text will describe how to develop a flowsheet or process flow diagram, but will not take the final step of adding the instrumentation, control and auxiliary pipes and valves required for the final engineering design in the piping and instrumentation diagram.

Batch processes are, by their nature, always in a transitory state. This requires the dynamics of the process to be optimized, and will be considered in Chapter 16. However, the control systems required to put this into practice will not be considered.

Approaches to 1.9 **Chemical Process Design and Integration**

In broad terms, there are three approaches to chemical process design and integration:

1) Creating an irreducible structure. The first approach follows the "onion logic", starting the design by choosing a reactor and then moving outward by adding a separation and recycle system, and so on. At each layer, decisions must be made on the basis of the information available at that stage. The ability to look ahead to the completed design might lead to different decisions. Unfortunately, this is not possible and, instead, decisions must be based on an incomplete picture.

This approach to creation of the design involves making a series of best local decisions. This might be based on the use of heuristics or rules of thumb developed from experience (Douglas, 1985) or on a more systematic approach. Equipment is added only if it can be justified economically on the basis

of the information available, albeit an incomplete picture. This keeps the structure *irreducible* and features that are technically or economically redundant are not included.

There are two drawbacks to this approach:

- a) Different decisions are possible at each stage of the design. To be sure that the best decisions have been made, the other options must be evaluated. However, each option cannot be evaluated properly without completing the design for that option and optimizing the operating conditions. This means that many designs must be completed and optimized in order to find the best.
- **b)** Completing and evaluating many options gives no guarantee of ultimately finding the best possible design, as the search is not exhaustive. Also, complex interactions can occur between different parts of a flowsheet. The effort to keep the system simple and not add features in the early stages of design may result in missing the benefit of interactions between different parts of the flowsheet in a more complex system.

The main advantage of this approach is that the design team can keep control of the basic decisions and interact as the design develops. By staying in control of the basic decisions, the intangibles of the design can be included in the decision making.

2) *Creating and optimizing a superstructure*. In this approach, a *reducible* structure, known as a *superstructure*, is first created that has embedded within it all feasible process options and all feasible interconnections that are candidates for an optimal design structure. Initially, redundant features are built into the superstructure. As an example, consider Figure 1.11 (Kocis and Grossmann, 1988). This shows one possible structure of a process for the manufacture of benzene from the reaction between toluene and hydrogen. In Figure 1.11, the hydrogen enters the process with a small amount of methane as an impurity. Thus, in Figure 1.11, the option of either purifying the hydrogen feed with a membrane or of passing it directly to the process is embedded. The hydrogen and toluene are mixed and preheated to reaction temperature. Only a furnace has been considered feasible in this case because of the high temperature



Figure 1.11

A superstructure for the manufacture of benzene from toluene and hydrogen incorporating some redundant features. (Reproduced from Kocis GR and Grossman IE, *Comp Chem Eng*, **13**: 797, with permission from Elsevier.)



Figure 1.12

Optimization discards many structural features leaving an optimized structure. (Reproduced from Kocis GR and Grossman IE, *Comp Chem Eng*, **13**: 797, with permission from Elsevier.)

required. Then the two alternative reactor options, isothermal and adiabatic reactors, are embedded, and so on. Redundant features have been included in an effort to ensure that all features that could be part of an optimal solution have been included.

The design problem is next formulated as a mathematical model. Some of the design features are continuous, describing the operation of each unit (e.g. flowrate, composition, temperature and pressure) and its size (e.g. volume, heat transfer area, etc.). Other features are discrete (e.g. whether a connection in the flowsheet is included or not, a membrane separator is included or not). Once the problem is formulated mathematically, its solution is carried out through the implementation of an optimization algorithm. An objective function is maximized or minimized (e.g. profit is maximized or cost is minimized) in a structural and parameter optimization. The optimization justifies the existence of structural features and deletes those features from the structure that cannot be justified economically. In this way, the structure is reduced in complexity. At the same time, the operating conditions and equipment sizes are also optimized. In effect, the discrete decision-making aspects of process design are replaced by a discrete/continuous optimization. Thus, the initial structure in Figure 1.11 is optimized to reduce the structure to the final design shown in Figure 1.12 (Kocis and Grossmann, 1988). In Figure 1.12, the membrane separator on the hydrogen feed has been removed

by optimization, as has the isothermal reactor and many other features of the initial structure shown in Figure 1.11.

There are a number of difficulties associated with this approach:

- a) The approach will fail to find the optimal structure if the initial structure does not have the optimal structure embedded somewhere within it. The more options included, the more likely it will be that the optimal structure has been included.
- **b**) If the individual unit operations are represented accurately, the resulting mathematical model will be extremely large and the objective function that must be optimized will be extremely irregular. The profile of the objective function can be like the terrain in a range of mountains with many peaks and valleys. If the objective function is to be maximized (e.g. maximize profit), each peak in the mountain range represents a local optimum in the objective function. The highest peak represents the global optimum. Optimization requires searching around the mountains in a thick fog to find the highest peak, without the benefit of a map and only a compass to tell the direction and an altimeter to show the height. On reaching the top of any peak, there is no way of knowing whether it is the highest peak because of the fog. All peaks must be searched to find the highest. There are crevasses to fall into that might be impossible to climb out of.

1

Such problems can be overcome in a number of ways. The first way is by changing the model such that the solution space becomes more regular, making the optimization simpler. This most often means simplifying the mathematical model. A second way is by repeating the search many times, but starting each new search from a different initial location. A third way exploits mathematical transformations and bounding techniques for some forms of mathematical expression to allow the global optimum to be found (Floudas, 2000). A fourth way is by allowing the optimization to search the solution space so as to allow the possibility of going downhill, away from an optimum point, as well as uphill. As the search proceeds, the ability of the algorithm to move downhill must be gradually taken away. These problems will be dealt with in more detail in Chapter 3.

c) The most serious drawback of this approach is that the design engineer is removed from the decision making. Thus, the many intangibles in design, such as safety and layout, which are difficult to include in the mathematical formulation, cannot be taken into account satisfactorily.

On the other hand, this approach has a number of advantages. Many different design options can be considered at the same time. The complex multiple trade-offs usually encountered in chemical process design can be handled by this approach. Also, the entire design procedure can be automated and is capable of producing designs quickly and efficiently.

3) Creating an initial design and evolving through structural and parameter optimization. The third approach is a variation on creating and optimizing a superstructure. In this approach, an initial design is first created. This is not necessarily intended to be an optimal design and does not necessarily have redundant features, but is simply a starting point. The initial design is then subjected to evolution through structural and parameter optimization, using an optimization algorithm (see Chapter 3). As with the superstructure approach, the design problem is formulated as a mathematical model. The design is then evolved one step at a time. Each step is known as a move. Each move in the evolution might change one of the continuous variables in the flowsheet (e.g. flowrate, composition, temperature or pressure) or might change the flowsheet structure. Changing the flowsheet structure might mean adding or deleting equipment (together with the appropriate connections), or adding new connections or deleting existing connections.

At each move, the objective function is evaluated (e.g. profit or cost). New moves are then carried out with the aim of improving the objective function. Rules must be created to specify the moves. The same structural options might be allowed as included in the superstructure approach, but in this evolutionary approach the structural moves are carried out to add or delete structural features one at a time. In addition to the structural moves, continuous moves are also carried out to optimize the flowsheet conditions.

The difficulties associated with this approach are similar to those for the creation and optimization of a superstructure:

a) The approach will fail to find the optimal structure if the structural moves have not been defined such that the

sequence of moves can lead to the optimal structure from the initial design.

- b) The objective function that must be optimized will be extremely irregular. Thus, again there are difficulties finding the global optimum. The optimization methods normally adopted for this approach allow the search to proceed even if the objective function deteriorates after a move. The ability of the algorithm to accept the deteriorating objective function is gradually taken away as the optimization proceeds. This approach will be dealt with in more detail in Chapter 3.
- c) Again, this approach has the disadvantage that the design engineer is removed from the decision making.

In summary, the three general approaches to chemical process design have advantages and disadvantages. However, whichever is used in practice, there is no substitute for understanding the problem.

This text concentrates on developing an understanding of the concepts required at each stage of the design. Such an understanding is a vital part of chemical process design and integration, whichever approach is followed.

1.10 The Nature of Chemical Process Design and Integration – Summary

Chemical products can be divided into three broad classes: commodity, fine and specialty chemicals. Commodity chemicals are manufactured in large volumes with low added value. Fine and specialty chemicals tend to be manufactured in low volumes with high added value. The priorities in the design of processes for the manufacture of these three classes of chemical products will differ.

The original design problem posed to the design team is often ill defined, even if it appears on the surface to be well defined. The design team must formulate well-defined design options from the original ill-defined problem, and these must be compared on the basis of consistent criteria.

Design starts by synthesizing design options, followed by simulation and evaluation. The simulation allows further detail to be added to the design. Care should be taken when carrying out simulation to choose appropriate physical property correlations and data. Safety should be considered as early as possible in the development of the design to make the design inherently safe. Both structural and parameter optimization can be carried out to improve the evaluation. Design options should be left open as far as practicable to avoid potentially attractive options being eliminated inappropriately on the basis of uncertain data.

The design might be new or retrofit of an existing process. If the design is a retrofit, then one of the objectives should be to maximize the use of existing equipment, even if it is not ideally suited to its new purpose. Another objective is to minimize the downtime required to carry out the modification.

Both continuous and batch process operations can be used. Batch processes are generally preferred for small-scale and specialty chemicals production.

When developing a chemical process design, there are two basic problems:

- Can all possible structures be identified?
- Can each structure be optimized such that all structures can be compared on a valid basis?

Design starts at the reactor, because it is likely to be the only place in the process where raw materials are converted into the desired chemical products. The reactor design dictates the separation and recycle problem. Together, the reactor design and separation and recycle dictate the heating and cooling duties for the heat exchanger network. Those duties that cannot be satisfied by heat recovery dictate the need for external heating and cooling utilities. The process and the utility system both have a demand for water and create aqueous effluents, giving rise to the water system. This hierarchy is represented by the layers in the "onion diagram" of Figure 1.8. Both continuous and batch process design follow this hierarchy, even though the time dimension in batch processes brings additional constraints in process design.

There are three general approaches to chemical process design:

- creating an irreducible structure;
- creating and optimizing a superstructure;
- creating an initial design and evolving through structural and parameter optimization.

Each of these approaches have advantages and disadvantages.

References

- Azapagic A (2014) Sustainability Considerations for Integrated Biorefineries, Trends in Biotechnology, 32: 1.
- Bansal V, Perkins JD, Pistikopoulos EN, Ross R and van Shijnedel JMG (2000a) Simultaneous Design and Control Optimisation Under Uncertainty, *Comp Comput Chem Eng*, 24: 261.
- Bansal V, Ross R, Perkins JD, Pistikopoulos EN (2000b) The Interactions of Design and Control: Double Effect Distillation, J Process Control, 10: 219.
- Biegler LT, Grossmann IE and Westerberg AW (1997), Systematic Methods of Chemical Process Design, Prentice Hall.
- Brennan D (1998) Process Industry Economics, IChemE, UK.
- Cussler EL and Moggridge GD (2011) Chemical Product Design, 2nd Edition, Cambridge University Press.
- Douglas JM (1985) A Hierarchical Decision Procedure for Process Synthesis, AIChE J, 31: 353.
- Floudas CA (2000) Deterministic Global Optimization: Theory, Methods and Applications, Kluwer Academic Publishers.
- Kocis GR and Grossmann IE (1988) A Modelling and Decomposition Strategy for the MINLP Optimization of Process Flowsheets, *Comp Comput Chem Eng*, 13: 797.
- Kookos IK and Perkins JD (2001) An Algorithm for Simultaneous Process Design and Control, *Ind Eng Chem Res*, 40: 4079.
- Linnhoff B, Townsend DW, Boland D, Hewitt GF, Thomas BEA, Guy AR and Marsland RH (1982) A User Guide on Process Integration for the Efficient Use of Energy, IChemE, Rugby, UK.
- Luyben WL, Tyreus BD and Luyben ML (1999) *Plant-wide Process Control*, McGraw-Hill.
- Seider WD, Seader JD, Lewin DR and Widagdo S (2010) Product and Process Design Principles, 3rd Edition, John Wiley & Sons, Inc.
- Sharratt PN (1997) Handbook of Batch Process Design, Blackie Academic and Professional.
- Sinnott R and Towler G (2009) *Chemical Engineering Design*, 5th Edition, Butterworth-Heinemann, Oxford, UK.
- Smith R and Linnhoff B (1988) The Design of Separators in the Context of Overall Processes, *Trans IChemE ChERD*, 66: 195.

Chapter 2

Process Economics

2.1 The Role of Process Economics

The purpose of chemical processes is to make money. An understanding of process economics is therefore critical in process design. Process economics has three basic roles in process design:

- 1) *Evaluation of design options*. Costs are required to evaluate process design options; for example, should a membrane or an adsorption process be used for a purification?
- 2) Process optimization. The settings of some process variables can have a major influence on the decision making in developing the flowsheet and on the overall profitability of the process. Optimization of such variables is usually required.
- **3)** *Overall project profitability.* The economics of the overall project should be evaluated at different stages during the design to assess whether the project is economically viable.

Before discussing how to use process economics for decision making, the most important costs that will be needed to compare options must first be reviewed.

2.2 Capital Cost for New Design

The total investment required for a new design can be broken down into four main parts that add up to the total investment:

- Battery limits investment
- Utility investment
- Off-site investment
- Working capital.
- Battery limits investment. The battery limit is a geographic boundary that defines the manufacturing area of the process. This is that part of the manufacturing system that converts raw

materials into products. It includes process equipment and buildings or structures to house it but excludes boiler-house facilities, site storage, pollution control, site infrastructure, and so on. The term battery limit is sometimes used to define the boundary of responsibility for a given project, especially in retrofit projects.

The battery limits investment required is the purchase of the individual plant items and their installation to form a working process. Equipment costs may be obtained from equipment vendors or from published cost data. Care should be taken as to the basis of such cost data. What is required for cost estimates is delivered cost, but cost is often quoted as FOB (free on board). Free on board means the manufacturer pays for loading charges on to a shipping truck, railcar, barge or ship, but not freight or unloading charges. To obtain a delivered cost requires typically 5 to 10% to be added to the FOB cost. The delivery cost depends on location of the equipment supplier, location of site to be delivered to, size of the equipment, and so on.

The cost of a specific item of equipment will be a function of:

- size,
- materials of construction,
- design pressure,
- design temperature.

Cost data are often presented as cost versus capacity charts, or expressed as a power law of capacity:

$$C_E = C_B \left(\frac{Q}{Q_B}\right)^M \tag{2.1}$$

where C_E = equipment cost with capacity Q

- C_B = known base cost for equipment with capacity Q_B
- M = constant depending on equipment type

A number of sources of such data are available in the open literature (Guthrie, 1969; Anson, 1977; Hall, Matley and McNaughton, 1982; Ulrich, 1984; Hall, Vatavuk and Matley, 1988; Remer and Chai, 1990; Gerrard, 2000; Peters, Timmerhaus and West, 2003). Published data are often old, sometimes from a variety of sources, with different ages. Such

Chemical Process Design and Integration, Second Edition. Robin Smith. © 2016 John Wiley & Sons, Ltd. Published 2016 by John Wiley & Sons, Ltd. Companion Website: www.wiley.com/go/smith/chemical2e

data can be brought up-to-date and put on a common basis using cost indexes:

$$\frac{C_1}{C_2} = \frac{INDEX_1}{INDEX_2} \tag{2.2}$$

where C_1 = equipment cost in year 1 C_2 = equipment cost in year 2 $INDEX_1$ = cost index in year 1 $INDEX_2$ = cost index in year 2

Commonly used indexes are the Chemical Engineering Indexes (1957–1959 index = 100) and Marshall and Swift (1926 index = 100), published in *Chemical Engineering* magazine, and the Nelson–Farrar Cost Indexes for refinery construction (1946 index = 100), published in the *Oil and Gas Journal*. The Chemical Engineering (CE) Indexes are particularly useful. CE Indexes are available for equipment covering:

- Heat Exchangers and Tanks
- Pipes, Valves and Fittings
- Process Instruments
- Pumps and Compressors
- Electrical Equipment
- Structural Supports and Miscellaneous.

A combined CE Index of Equipment is available. CE Indexes are also available for:

- Construction and Labor Index
- Buildings Index
- Engineering and Supervision Index.

All of the above indexes are combined to produce a CE Composite Index.

Table 2.1 presents data for a number of equipment items on the basis of January 2000 costs (Gerrard, 2000) (CE Composite Index = 391.1, CE Index of Equipment = 435.8). A guide to the selection of materials of construction can be found in Appendix B.

Cost correlations for vessels are normally expressed in terms of the mass of the vessel. This means that not only a preliminary sizing of the vessel is required but also a preliminary assessment of the mechanical design (Mulet, Corripio and Evans, 1981a, 1981b).

Materials of construction have a significant influence on the capital cost of equipment. Table 2.2 gives some approximate average factors to relate the different materials of construction for equipment capital cost.

It should be emphasized that the factors in Table 2.2 are average and only approximate and will vary, amongst other things, according to the type of equipment. As an example, consider the effect of materials of construction on the capital cost of distillation columns. Table 2.3 gives materials of construction cost factors for distillation columns.

The cost factors for shell-and-tube heat exchangers are made more complex by the ability to construct the different components from different materials of construction. Table 2.4 gives typical materials of construction factors for shell-and-tube heat exchangers. The operating pressure also influences equipment capital cost as a result of thicker vessel walls to withstand increased pressure. Table 2.5 presents typical factors to account for the pressure rating.

As with materials of construction correction factors, the pressure correction factors in Table 2.5 are average and only approximate and will vary, amongst other things, according to the type of equipment. Finally, its operating temperature also influences equipment capital cost. This is caused by, amongst other factors, a decrease in the allowable stress for materials of construction as the temperature increases. Table 2.6 presents typical factors to account for the operating temperature.

Thus, for a base cost for carbon steel equipment at moderate pressure and temperature, the actual cost can be estimated from:

$$C_E = C_B \left(\frac{Q}{Q_B}\right)^M f_M f_P f_T \tag{2.3}$$

where C_E = equipment cost for carbon steel at moderate

pressure and temperature with capacity Q

 C_B = known base cost for equipment with capacity Q_B

M = constant depending on equipment type

 f_M = correction factor for materials of construction

 f_P = correction factor for design pressure

 f_T = correction factor for design temperature

In addition to the purchased cost of the equipment, investment is required to install the equipment. Installation costs include:

- cost of installation,
- piping and valves,
- control systems,
- foundations,
- structures,
- insulation,
- fire proofing,
- electrical,
- painting,
- engineering fees,
- contingency.

The total capital cost of the installed battery limits equipment will normally be two to four times the purchased cost of the equipment (Lang, 1947; Hand, 1958).

In addition to the investment within the battery limits, investment is also required for the structures, equipment and services outside the battery limits that are required to make the process function.

- Utility investment. Capital investment in utility plant could include equipment for:
 - electricity generation,
 - electricity distribution,

Table 2.1

Typical equipment capacity delivered capital cost correlations.

Equipment	Material of construction	Capacity measure	Base size $Q_{\rm B}$	Base cost C _B (\$)	Size range	Cost exponent M
Agitated reactor	CS	Volume (m ³)	1	1.15×10^{4}	1–50	0.45
Pressure vessel	SS	Mass (t)	6	9.84×10^{4}	6–100	0.82
Distillation column (empty shell)	CS	Mass (t)	8	6.56×10^4	8–300	0.89
Sieve trays (10 trays)	CS	Column diameter (m)	0.5	6.56×10^{3}	0.5–4.0	0.91
Valve trays (10 trays)	CS	Column diameter (m)	0.5	1.80×10^{4}	0.5–4.0	0.97
Structured packing (5 m height)	SS (low grade)	Column diameter (m)	0.5	1.80×10^{4}	0.5–4.0	1.70
Scrubber (including random packing)	SS (low grade)	Volume (m ³)	0.1	4.92×10^{3}	0.1–20	0.53
Cyclone	CS	Diameter (m)	0.4	1.64×10^{3}	0.4–3.0	1.20
Vacuum filter	CS	Filter area (m ²)	10	8.36×10^{4}	10–25	0.49
Dryer	SS (low grade)	Evaporation rate $(kg H_2O \cdot h^{-1})$	700	2.30×10^{5}	700–3000	0.65
Shell-and-tube heat exchanger	CS	Heat transfer area (m ²)	80	3.28×10^{4}	80–4000	0.68
Air-cooled heat exchanger	CS	Plain tube heat transfer area (m ²)	200	1.56×10^{5}	200–2000	0.89
Centrifugal pump (small, including motor)	SS (high grade)	Power (kW)	1	1.97×10^{3}	1–10	0.35
Centrifugal pump (large, including motor)	CS	Power (kW)	4	9.84×10^{3}	4–700	0.55
Compressor (including motor)		Power (kW)	250	9.84×10^{4}	250-10,000	0.46
Fan (including motor)	CS	Power (kW)	50	1.23×10^{4}	50-200	0.76
Vacuum pump (including motor)	CS	Power (kW)	10	1.10×10^{4}	10–45	0.44
Electric motor		Power (kW)	10	1.48×10^{3}	10–150	0.85
Storage tank (small atmospheric)	SS (low grade)	Volume (m ³)	0.1	3.28×10^{3}	0.1–20	0.57
Storage tank (large atmospheric)	CS	Volume (m ³)	5	1.15×10^{4}	5-200	0.53
Silo	CS	Volume (m ³)	60	1.72×10^{4}	60–150	0.70
Package steam boiler (fire-tube boiler)	CS	Steam generation $(kg \cdot h^{-1})$	50,000	4.64×10^{5}	50,000-350,000	0.96
Field erected steam boiler (water-tube boiler)	CS	Steam generation $(kg \cdot h^{-1})$	20,000	3.28×10^{5}	10,000-800,000	0.81
Cooling tower (forced draft)		Water flowrate $(m^3 \cdot h^{-1})$	10	4.43×10^{3}	1040	0.63

CS = carbon steel; SS (low grade) = low-grade stainless steel, for example, type 304; SS (high grade) = high-grade stainless steel, for example, type 316.

Table 2.2

Typical average equipment materials of construction capital cost factors.

Material	Correction factor f_M
Carbon steel	1.0
Aluminum	1.3
Stainless steel (low grades)	2.4
Stainless steel (high grades)	3.4
Hastelloy C	3.6
Monel	4.1
Nickel and inconel	4.4
Titanium	5.8

Table 2.3

Typical materials of construction capital cost factors for pressure vessels and distillation columns (Mulet, Corripio and Evans, 1981a, 1981).

Material	Correction factor f_M
Carbon steel	1.0
Stainless steel (low grades)	2.1
Stainless steel (high grades)	3.2
Monel	3.6
Inconel	3.9
Nickel	5.4
Titanium	7.7

- steam generation,
- steam distribution,
- process water,
- cooling water,
- firewater,
- effluent treatment,
- refrigeration,
- compressed air,
- inert gas (nitrogen).

The cost of utilities is considered from their sources within the site to the battery limits of the chemical process served.

- 3) Off-site investment. Off-site investment includes:
 - auxiliary buildings such as offices, medical, personnel, locker rooms, guardhouses, warehouses and maintenance shops,
 - roads and paths,
 - railroads,

Table 2.4

Typical materials of construction capital cost factors for shell-and-tube heat exchangers (Anson, 1977).

Material	Correction factor f_M
CS shell and tubes	1.0
CS shell, aluminum tubes	1.3
CS shell, monel tubes	2.1
CS shell, SS (low grade) tubes	1.7
SS (low grade) shell and tubes	2.9

Table 2.5

Typical equipment pressure capital cost factors.

Design pressure (bar absolute)	Correction factor f_P
0.01	2.0
0.1	1.3
0.5–7	1.0
50	1.5
100	1.9

Table 2.6

Typical equipment temperature capital cost factors.

Design temperature (°C)	Correction factor f_T
0–100	1.0
300	1.6
500	2.1

- fire protection systems,
- communication systems,
- waste disposal systems,
- storage facilities for end product, water and fuel not directly connected with the process,
- plant service vehicles, loading and weighing devices.

The cost of the utilities and off-sites (together sometimes referred to as *services*) ranges typically from 20 to 40% of the total installed cost of the battery limits plant (Bauman, 1955). In general terms, the larger the plant, the larger will tend to be the fraction of the total project cost that goes to utilities and off-sites. In other words, a small project will require typically 20% of the total installed cost for utilities

and off-sites. For a large project, the figure can be typically up to 40%.

- **4)** *Working capital*, Working capital is what must be invested to get the plant into productive operation. This is money invested before there is a product to sell and includes:
 - raw materials for plant start-up (including wasted raw materials),
 - raw materials, intermediate and product inventories,
 - cost of transportation of materials for start-up,
 - money to carry accounts receivable (i.e. credit extended to customers) less accounts payable (i.e. credit extended by suppliers),
 - money to meet payroll when starting up.

Theoretically, in contrast to fixed investment, this money is not lost but can be recovered when the plant is closed down.

Stocks of raw materials, intermediate and product inventories often have a key influence on the working capital and are under the influence of the designer. Issues relating to storage will be discussed in more detail in Chapters 14 and 16. For an estimate of the working capital requirements, take either (Holland, Watson and Wilkinson, 1983):

- a) 30% of annual sales or
- b) 15% of total capital investment.
- 5) *Total capital cost.* The total capital cost of the process, services and working capital can be obtained by applying multiplying factors or *installation factors* to the purchase cost of individual items of equipment (Lang, 1947; Hand, 1958):

$$C_F = \sum_i f_i C_{E,i} \tag{2.4}$$

where C_F = fixed capital cost for the complete system

 $C_{E,i} = \text{cost of equipment } i$

 f_i = installation factor for equipment *i*

If an average installation factor for all types of equipment is to be used (Lang, 1947),

$$C_F = f_I \sum_{i} C_{E,i} \tag{2.5}$$

where f_I = overall installation factor for the complete system.

The overall installation factor for new design is broken down in Table 2.7 into component parts according to the dominant phase being processed. The cost of the installation will depend on the balance of gas and liquid processing versus solids processing. If the plant handles only gases and liquids, it can be characterized as fluid processing. A plant can be characterized as solids processing if the bulk of the material handling is solid phase. For example, a solid processing plant could be a coal or an ore preparation plant. Between the two extremes of fluid processing and solids processing are processes that handle a significant amount of both solids and fluids. For example, a shale oil plant involves preparation of the shale oil followed by extraction of fluids from the shale oil and then separation and processing of the fluids. For these types of plant, the contributions to the capital cost can be estimated

Table 2.7

Typical factors for capital cost based on delivered equipment costs.

	Type of process			
Item	Fluid processing	Solid processing		
Direct costs				
Equipment delivered cost	1	1		
Equipment erection, f_{ER}	0.4	0.5		
Piping (installed), f_{PIP}	0.7	0.2		
Instrumentation and controls (installed), f_{INST}	0.2	0.1		
Electrical (installed), f_{ELEC}	0.1	0.1		
Utilities, f_{UTIL}	0.5	0.2		
Off-sites, fos	0.2	0.2		
Buildings (including services), <i>f</i> _{BUILD}	0.2	0.3		
Site preparation, f_{SP}	0.1	0.1		
Total capital cost of installed equipment	3.4	2.7		
Indirect costs				
Design, engineering and construction, f_{DEC}	1.0	0.8		
Contingency (about 10% of fixed capital costs), f_{CONT}	0.4	0.3		
Total fixed capital cost	4.8	3.8		
Working capital				
Working capital (15% of total capital cost), f_{WC}	0.7	0.6		
Total capital cost, f_I	5.8	4.4		

from the two extreme values in Table 2.7 by interpolation in proportion of the ratio of major processing steps that can be characterized as fluid processing and solid processing.

A number of points should be noted about the various contributions to the capital cost in Table 2.7. The values are:

- based on carbon steel, moderate operating pressure and temperature;
- average values for all types of equipment, whereas in practice the values will vary according to the type of equipment;
- only guidelines and the individual components will vary from project to project;
- applicable to new design only.

When equipment uses materials of construction other than carbon steel, or operating temperatures are extreme, the capital cost needs to be adjusted accordingly. Whilst the equipment cost and its associated pipework will be changed, the other installation costs will be largely unchanged, whether the equipment is manufactured from carbon steel or exotic materials of construction. Thus, the application of the factors from Tables 2.2 to 2.6 should only be applied to the equipment and pipework:

$$C_{F} = \sum_{i} [f_{M}f_{P}f_{T}(1+f_{PIP})]_{i}C_{E,i} + (f_{ER}+f_{INST}+f_{ELEC}+f_{UTIL}+f_{OS}+f_{BUILD} + f_{SP}+f_{DEC}+f_{CONT}+f_{WS})\sum_{i}C_{E,i}$$
(2.6)

Thus, to estimate the fixed capital cost:

- 1) List the main plant items and estimate their size.
- 2) Estimate the equipment cost of the main plant items.
- **3)** Adjust the equipment costs to a common time basis using a cost index.
- 4) Convert the cost of the main plant items to carbon steel, moderate pressure and moderate temperature.
- 5) Select the appropriate installation subfactors from Table 2.7 and adjust for individual circumstances.
- 6) Select the appropriate materials of construction, operating pressure and operating temperature correction factors for each of the main plant items.
- 7) Apply Equation 2.6 to estimate the total fixed capital cost.

Equipment cost data used in the early stages of a design will by necessity normally be based on capacity, materials of construction, operating pressure and operating temperature. However, in reality, the equipment cost will depend also on a number of factors that are difficult to quantify (Brennan, 1998):

- multiple purchase discounts,
- buyer-seller relationships,
- capacity utilization in the fabrication shop (i.e. how busy the fabrication shop is),
- required delivery time,
- availability of materials and fabrication labor,
- special terms and conditions of purchase, and so on.

Care should also be taken to the geographic location. Costs to build the same plant can differ significantly between different locations, even within the same country. Such differences will result from variations in climate and its effect on the design requirements and construction conditions, transportation costs, local regulations, local taxes, availability and productivity of construction labor, and so on (Gary and Handwerk, 2001). For example, in the United States of America, Gulf Coast costs tend to be the lowest, with costs in other areas typically 20 to 50% higher, and those in Alaska two or three times higher than the US Gulf Coast (Gary and Handwerk, 2001). In Australia, costs tend to be the lowest in the region of Sydney and the other metropolitan cities, with costs in remote areas such as North Queensland typically 40 to 80% higher (Brennan, 1998). Costs also differ from country to country. For example, relative to costs for a plant located in the US Gulf Coast, costs in India might be expected to be 20% cheaper, in Indonesia 30% cheaper, but in the United Kingdom 15% more expensive, because of labor costs, cost of land, and so on (Brennan, 1998).

It should be emphasized that capital cost estimates using installation factors are at best crude and at worst highly misleading. When preparing such an estimate, the designer spends most of the time on the equipment costs, which represents typically 20 to 40% of the total installed cost. The bulk costs (civil engineering, labor, etc.) are factored costs that lack definition. At best, this type of estimate can be expected to be accurate to $\pm 30\%$. To obtain greater accuracy requires detailed examination of all aspects of the investment. Thus, for example, to estimate the erection cost accurately requires knowledge of how much concrete will be used for foundations, how much structural steelwork is required, and so on. Such detail can only be included from access to a large database of cost information.

The shortcomings of capital cost estimates using installation factors are less serious in preliminary process design if used to compare options on a common basis. If used to compare options, the errors will tend to be less serious as the errors will tend to be consistent across the options.

Example 2.1 A new heat exchanger is to be installed as part of a large project. Preliminary sizing of the heat exchanger has estimated its heat transfer area to be 250 m^2 . Its material of construction is low-grade stainless steel and its pressure rating is 5 bar. Estimate the contribution of the heat exchanger to the total cost of the project (CE Index of Equipment = 682.0).

Solution From Equation 2.1 and Table 2.1, the capital cost of a carbon steel heat exchanger can be estimated from:

$$C_E = C_B \left(\frac{Q}{Q_B}\right)^M \\ = 3.28 \times 10^4 \left(\frac{250}{80}\right)^{0.68} \\ = \$7.12 \times 10^4$$

The cost can be adjusted to bring it up-to-date using the ratio of cost indexes:

$$C_E = 7.12 \times 10^4 \left(\frac{682.0}{435.8}\right)$$
$$= \$1.11 \times 10^5$$

The cost of a carbon steel heat exchanger needs to be adjusted for the material of construction. Because of the low pressure rating, no correction for pressure is required (Table 2.5), but the cost needs to be adjusted for the materials of construction. From Table 2.4, $f_M = 2.9$, and the total cost of the installed equipment can be estimated from Equation 2.6 and Table 2.7. If the project is a complete new plant, the contribution of the heat exchanger to the total cost can be estimated to be:

$$\begin{split} C_F &= f_M (1 + f_{PIP}) C_E + (f_{ER} + f_{INST} + f_{ELEC} + f_{UTIL} + f_{OS} \\ &+ f_{BUILD} + f_{SP} + f_{DEC} + f_{CONT} + f_{WS}) C_E \\ &= 2.9 (1 + 0.7) 1.11 \times 10^5 + (0.4 + 0.2 + 0.1 + 0.5 + 0.2) \\ &+ 0.2 + 0.1 + 1.0 + 0.4 + 0.7) 1.11 \times 10^5 \\ &= 8.73 \times 1.11 \times 10^5 \\ &= \$9.69 \times 10^5 \end{split}$$

Had the new heat exchanger been an addition to an existing plant that did not require investment in electrical services, utilities, offsites, buildings, site preparation or working capital, then the cost would be estimated from:

$$C_F = f_M (1 + f_{PIP}) C_E + (f_{ER} + f_{INST} + f_{DEC} + f_{CONT}) C_E$$

= 2.9(1 + 0.7)1.11 × 10⁵ + (0.4 + 0.2 + 1.0 + 0.4)1.11 × 10⁵
= 6.93 × 1.11 × 10⁵
- \$7.69 × 10⁵

Installing a new heat exchanger into an existing plant might require additional costs over and above those estimated here. Connecting new equipment to existing equipment, modifying or relocating existing equipment to accommodate the new equipment and downtime might all add to the costs.

2.3 Capital Cost for Retrofit

Estimating the capital cost of a retrofit project is much more difficult than for new design. In principle, the cost of individual items of new equipment will be the same, whether it is a grassroot design or a retrofit. By contrast, installation factors for equipment in retrofit can be completely different from grassroot design, and could be higher or lower. If the new equipment can take advantage of existing space, foundations, electrical cabling, and so on, the installation factor might in some cases be lower than in new design. This will especially be the case for small items of equipment. However, most often, retrofit installation factors will tend to be higher than in grassroot design and can be very much higher. This is because existing equipment might need to be modified or moved to allow installation of new equipment. Also, access to the area where the installation is required is likely to be much more restricted in retrofit than in the phased installation of new plant. Smaller projects (as the retrofit is likely to be) tend to bring higher cost of installation per unit of installed equipment than larger projects.

As an example, one very common retrofit situation is the replacement of distillation column internals to improve the performance of the column. The improvement in performance sought is often an increase in the throughput. This calls for existing internals to be removed and then to be replaced with the new internals. Table 2.8 gives typical factors for the removal of old internals and the installation of new ones (Bravo, 1997).

As far as utilities and off-sites are concerned, it is also difficult to generalize. Small retrofit projects are likely not to require any investment in utilities and off-sites. Larger-scale retrofit might

Table 2.8

Modification costs for distillation column retrofit (Bravo, 1997).

Column modification	Cost of modification (multiply factor by cost of new hardware)
Removal of trays to install new trays	0.1 for the same tray spacing 0.2 for different tray spacing
Removal of trays to install packing	0.1
Removal of packing to install new trays	0.07
Installation of new trays	1.0–1.4 for the same tray spacing 1.2–1.5 for different tray spacing 1.3–1.6 when replacing packing
Installation of new structured packing	0.5–0.8

demand a major revamp of the utilities and off-sites, which can be particularly expensive because existing equipment might need to be modified or removed to make way for new utilities and off-sites equipment.

Working capital is also difficult to generalize. Most often, there will be no significant working capital associated with a retrofit project. For example, if a few items of equipment are replaced to increase the capacity of a plant, this will not significantly change the raw materials and product inventories, money to carry accounts receivable, money to meet payroll, and so on. On the other hand, if the plant changes function completely, significant new storage capacity is added, and so on, there might be a significant element of working capital.

One of the biggest sources of cost associated with retrofit can be the downtime (the period during which the plant will not be productive) required to carry out the modifications. The cost of lost production can be the dominant feature of retrofit projects. The cost of lost production should be added to the capital cost of a retrofit project. To minimize the downtime and cost of lost production requires that as much preparation as possible is carried out whilst the plant is operating. The modifications requiring the plant to be shut down should be minimized. For example, it might be possible for new foundations to be installed and new equipment put into place while the plant is still operating, leaving the final pipework and electrical modifications for the shutdown. Retrofit projects are often arranged such that the preparation is carried out prior to a regular maintenance shutdown, with the final modifications coinciding with the planned maintenance shutdown. Such considerations often dominate the decisions made as to how to modify the process for retrofit.

Because of all of these uncertainties, it is difficult to provide general guidelines for the capital cost of retrofit projects. The basis of the capital cost estimate should be to start with the required investment in new equipment. Installation factors for the installation of equipment for grassroot design from Table 2.7 need to be adjusted according to circumstances (usually increased). If old equipment needs to be modified to take up a new role (e.g. move an existing heat exchanger to a new duty), then an installation cost must be applied without the equipment cost. In the absence of better information, the installation cost can be taken to be that for the equivalent piece of new equipment. Some elements of the total cost breakdown in Table 2.7 will not be relevant and should not be included. In general, for the estimation of capital cost for retrofit, a detailed examination of the individual features of retrofit projects is necessary.

Example 2.2 An existing heat exchanger is to be repiped to a new duty in a retrofit project without moving its location. The only significant investment is piping modifications. The heat transfer area of the existing heat exchanger is 250 m^2 . The material of construction is low-grade stainless steel and its design pressure is 5 bar. Estimate the cost of the project (CE Index of Equipment = 682.0).

Solution All retrofit projects have individual characteristics and it is impossible to generalize the costs. The only way to estimate costs with any certainty is to analyze the costs of all of the modifications in detail. However, in the absence of such detail, a very preliminary estimate can be obtained by estimating the retrofit costs from the appropriate installation costs for a new design. In this case, piping costs can be estimated from those for a new heat exchanger of the same specification, but excluding the equipment cost. For Example 2.1, the cost of a new stainless steel heat exchanger with an area of 250 m^2 was estimated to be $\$1.11 \times 10^5$. The piping costs (stainless steel) can therefore be estimated to be:

Piping cost =
$$f_M f_{PIP} C_E$$

= 2.9 × 0.7 × 1.11 × 10⁵
= 2.03 × 1.11 × 10⁵
= \$2.25 × 10⁵

This estimate should not be treated with any confidence. It will give an idea of the costs and might be used to compare retrofit options on a like-for-like basis, but could be very misleading.

Example 2.3 An existing distillation column is to be revamped to increase its capacity by replacing the existing sieve trays with stainless steel structured packing. The column shell is 46 m tall and 1.5 m diameter and currently fitted with 70 sieve trays with a spacing of 0.61 m. The existing trays are to be replaced with stainless steel structured packing with a total height of 30 m. Estimate the cost of the project (CE Index of Equipment = 682.0).

Solution First, estimate the purchase cost of the new structured packing from Equation 2.1 and Table 2.1, which gives costs for a 5 m height of packing:

$$C_E = C_B \left(\frac{Q}{Q_B}\right)^M$$

= 1.8 × 10⁴ × $\frac{30}{5} \left(\frac{1.5}{0.5}\right)^{1.5}$
= \$6.99 × 10⁵

Adjusting the cost to bring it up-to-date using the ratio of cost indexes:

$$C_E = 6.99 \times 10^5 \left(\frac{682.0}{435.8}\right)$$
$$= \$1.09 \times 10^6$$

From Table 2.8, the factor for removing the existing trays is 0.1 and that for installing the new packing is 0.5 to 0.8 (say 0.7). Estimated total cost of the project:

$$= (1+0.1+0.7)1.09 \times 10^{6}$$
$$= \$1.96 \times 10^{6}$$

2.4 Annualized Capital Cost

Capital for new installations may be obtained from:

- a) Loans from banks
- **b**) Issue by the company of common stock (ordinary shares), preferred stock (preference shares) or bonds
- c) Accumulated net cash flow arising from company profit over time.

Interest on loans from banks, preferred stock and bonds is paid at a fixed rate of interest. A share of the profit of the company is paid as a dividend on common stock and preferred stock (in addition to the interest paid on preferred stock).

The cost of the capital for a project thus depends on its source. The source of the capital often will not be known during the early stages of a project and yet there is a need to select between process options and carry out optimization on the basis of both capital and operating costs. This is difficult to do unless both capital and operating costs can be expressed on a common basis. Capital costs can be expressed on an annual basis if it is assumed that the capital has been borrowed over a fixed period (for large projects, usually 5 to 10 years) at a fixed rate of interest, in which case the capital cost can be annualized according to:

Annualized capital cost = Capital cost
$$\times \frac{i(1+i)^n}{(1+i)^n - 1}$$
 (2.7)

where i = fractional interest rate per year n = number of years

The derivation of Equation 2.7 is given in Appendix C.

As stated previously, the source of capital is often not known, and hence it is not known whether Equation 2.7 is appropriate to represent the cost of capital. Equation 2.7 is, strictly speaking, only appropriate if the money for capital expenditure is to be borrowed over a fixed period at a fixed rate of interest. Moreover, if Equation 2.7 is accepted, then the number of years over which

the capital is to be annualized is known, as is the rate of interest. However, the most important thing is that, even if the source of capital is not known, and uncertain assumptions are necessary, Equation 2.7 provides a common basis for the comparison of competing projects and design alternatives within a project.

Example 2.4 The purchased cost of a new distillation column installation is \$1 million. Calculate the annual cost of installed capital if the capital is to be annualized over a five-year period at a fixed rate of interest of 5%.

Solution First calculate the installed capital cost (Table 2.7):

$$C_F = f_i C_E$$

= 5.8 × (1,000,000)
= \$5,800,000
Annualization factor = $\frac{i(1+i)^n}{(1+i)^n - 1}$
= $\frac{0.05(1+0.05)^5}{(1+0.05)^5 - 1}$
= 0.2310
Annualized capital cost = 5,800,000 × 0.2310
= \$1,340,000 y^{-1}

A

When using annualized capital cost to carry out optimization, the designer should not lose sight of the uncertainties involved in the capital annualization. In particular, changing the annualization period can lead to very different results when, for example, carrying out a trade-off between energy and capital costs. When carrying out optimization, the sensitivity of the result to changes in the assumptions should be tested.

2.5 Operating Cost

1) Raw materials cost. In most processes, the largest individual operating cost is raw materials. The cost of raw materials and the product selling prices tend to have the largest influence on the economic performance of the process. The cost of raw materials and price of products depends on whether the materials in question are being bought and sold under a contractual arrangement (either within the same company or outside the company) or on the open market. Open market prices for some chemical products can fluctuate considerably with time. Raw materials might be purchased and products sold below or above the open market price when under a contractual arrangement, depending on the state of the market. Buying and selling on the open market may give the best purchase and selling prices but give rise to an uncertain economic environment. A long-term contractual agreement may reduce profit per unit of production but gives a degree of certainty over the project life.

The values of raw materials and products can be found in various on-line sources and trade journals. However, the values reported in such sources will be subject to short-term fluctuations, and long-term forecasts will be required for investment analysis.

2) Catalysts and chemicals consumed in manufacturing other than raw materials. Catalysts will need to be replaced or regenerated though the life of a process (see Chapter 6). The replacement of catalysts might be on a continuous basis if homogeneous catalysts are used (see Chapters 5 and 6). Heterogeneous catalysts might also be replaced continuously if they deteriorate rapidly, and regeneration cannot fully reinstate the catalyst activity. More often for heterogeneous catalysts, regeneration or replacement will be carried out on an intermittent basis, depending on the characteristics of the catalyst deactivation.

In addition to the cost of catalysts, there might be significant costs associated with chemicals consumed in manufacturing that do not form part of the final product. For example, acids and alkalis might be consumed to adjust the pH of streams. Such costs might be significant.

- **3)** *Utility operating cost.* Utility operating cost is usually the most significant variable operating cost after the cost of raw materials. This is especially the case for the production of commodity chemicals. Utility operating cost includes:
 - fuel,
 - electricity,
 - steam,
 - cooling water,
 - refrigeration,
 - compressed air,
 - inert gas.

Utility costs can vary enormously between different processing sites. This is especially true of fuel and power costs. Not only do fuel costs vary considerably between different fuels (coal, oil, natural gas) and geographic locations but costs also tend to be sensitive to market fluctuations. Contractual relationships also have a significant effect on fuel costs. The price paid for fuel depends very much on how much is purchased and the pattern of usage.

When electricity is bought from centralized power-generation companies under long-term contract, the price tends to be more stable than fuel costs, since power-generation companies tend to negotiate long-term contracts for fuel supply. However, purchased electricity prices (and sales price if excess electricity is generated and exported) are normally subject to tariff variations. Electricity tariffs can depend on the season of the year (e.g. winter versus summer), the time of the week (e.g. weekend versus weekday) and the time of day (e.g. night versus day). In hot countries, electricity is usually more expensive in the summer than in the winter because of the demand from air-conditioning systems. In cold countries, electricity is usually more expensive in the winter than in the summer because of the demand from space heating. The price structure for electricity can be complex, but should be predictable if based on contractual arrangements. If electricity is purchased from a spot market in those countries that have such arrangements, then prices can vary wildly.

Steam costs vary with the price of fuel and electricity. If steam is only generated at low pressure and not used for power generation in steam turbines, then the cost can be estimated from fuel costs assuming an efficiency of generation and distribution losses. The efficiency of generation depends on the boiler efficiency and the steam consumed in boiler feedwater production (see Chapter 23). Losses from the steam distribution system include heat losses from steam distribution and condensate return pipework to the environment, steam condensate lost to drain and not returned to the boilers and steam leaks. The efficiency of steam generation (including auxiliary boiler-house requirements, see Chapter 23) is typically around 85 to 90% and distribution losses of perhaps another 10%, giving an overall efficiency for steam generation and distribution of typically around 75 to 80% (based on the net calorific value of the fuel). Care should be exercised when considering boiler efficiency and the efficiency of steam generation. These figures are often quoted on the basis of gross calorific value of the fuel, which includes the latent heat of the water vapor from combustion. This latent heat is rarely recovered through condensation of the water vapor in the combustion gases. The net calorific value of the fuel assumes that the latent heat of the water vapor is not recovered and is therefore the most relevant value, yet figures are often quoted on the basis of gross calorific value.

If high-pressure steam mains are used, then the cost of steam should be related in some way to its capacity to generate power in a steam turbine rather than simply to its additional heating value. The high-pressure steam is generated in the utility boilers and the low-pressure steam is generated by reducing pressure through steam turbines to produce power. This will be discussed in more detail in Chapter 23. One simple way to cost steam is to calculate the cost of the fuel required to generate the high-pressure steam (including any losses), and this fuel cost is then the cost of the high-pressure steam. Although this neglects water costs, labor costs, and so on, it will be a reasonable estimate, as costs are normally dominated by fuel costs. Lowpressure mains have a value equal to that of the high-pressure mains minus the value of power generated by letting the steam down to the low pressure in a steam turbine. To calculate the cost of steam that has been expanded though a steam turbine, the power generated in such an expansion can be calculated. The simplest way to do this is on the basis of a comparison between an ideal and a real expansion though a steam turbine. Figure 2.1 shows a steam turbine expansion on an enthalpyentropy plot. In an ideal turbine, steam with an initial pressure P_1 and enthalpy H_1 expands isentropically to pressure P_2 and enthalpy $H_{2,IS}$. In such circumstances, the ideal work output is $(H_1 - H_{2,IS})$. Because of the frictional effects in the turbine nozzles and blade passages, the exit enthalpy is greater than it would be in an ideal turbine, and the work output is consequently less, given by H_2 in Figure 2.1. The actual work output is given by $(H_1 - H_2)$. The turbine isentropic efficiency η_{IS} measures the ratio of the actual to ideal work obtained:

$$\eta_{IS} = \frac{H_1 - H_2}{H_1 - H_{2,IS}} \tag{2.8}$$





The output from the turbine might be superheated or partially condensed, as is the case in Figure 2.1. The following example illustrates the approach.

Example 2.5 The pressures of three steam mains have been set to the conditions given in Table 2.9. High-pressure (HP) steam is generated in boilers at 41 barg and superheated to 400 °C. Medium-pressure (MP) and low-pressure (LP) steam are generated by expanding high-pressure steam through a steam turbine with an isentropic efficiency of 80%. The cost of fuel is 4.00 GJ^{-1} and the cost of electricity is $0.07 \text{ kW}^{-1} \cdot \text{h}^{-1}$. Boiler feedwater is available at 100 °C with a heat capacity of 4.2 kJ·kg⁻¹·K⁻¹. Assuming an efficiency of steam generation of 85% and distribution losses of 10% of the fuel fired, estimate the cost of steam for the three levels.

Table 2.9

Steam mains pressure settings.

Mains	Pressure (barg)
HP	41
MP	10
LP	3

Solution *Cost of 41 barg steam.* From steam tables, for 41 barg steam at 400 °C:

Enthalpy =
$$3212 \text{ kJ} \cdot \text{kg}^{-1}$$

For boiler feedwater:

ł

Enthalpy =
$$4.2(100-0)$$
 (relative to water at 0 °C)
= $420 \text{ kJ} \cdot \text{kg}^{-1}$

To generate 41 barg steam at 400 °C:

Heat duty =
$$3212 - 420 = 2792 \text{ kJ} \cdot \text{kg}^-$$

For 41 barg steam:

Cost =
$$4.00 \times 10^{-6} \times 2792 \times \frac{1}{0.75}$$

= $0.01489 \$ \cdot kg^{-1}$
= $14.89 \$ \cdot t^{-1}$

Cost of 10 barg steam. Here 41 barg steam is now expanded to 10 barg in a steam turbine.

From steam tables, inlet conditions at 41 barg and 400 °C are:

$$H_1 = 3212 \text{ kJ} \cdot \text{kg}^{-1}$$

 $S_1 = 6.747 \text{ kJ} \cdot \text{kg}^{-1} \cdot \text{K}^{-1}$

Turbine outlet conditions for isentropic expansion to 10 barg are:

- 1

$$H_{2,IS} = 2873 \text{ kJ} \cdot \text{kg}^{-1}$$

 $S_2 = 6.747 \text{ kJ} \cdot \text{kg}^{-1} \cdot \text{K}^{-1}$

For single-stage expansion with isentropic efficiency of 80%:

$$H_2 = H_1 - \eta_{IS}(H_1 - H_{2,IS})$$

= 3212 - 0.8(3212 - 2873)
= 2941 kJ \cdot kg^{-1}

From steam tables, the outlet temperature is 251 °C, which corresponds to a superheat of 67 °C. Although steam for process heating is preferred at saturated conditions, it is not desirable in this case to de-superheat by boiler feedwater injection to bring to saturation conditions. If saturated steam is fed to the main, then the heat losses from the main will cause a large amount of condensation in the main, which is undesirable. Hence, it is standard practice to feed steam to the main with a superheat of at least 10 °C to avoid condensation in the main.

Power generation =
$$3212 - 2941$$

= $271 \text{ kJ} \cdot \text{kg}^{-1}$
Value of power generation = $271 \times \frac{0.07}{3600}$
= $0.00527 \$ \cdot \text{kg}^{-1}$
Cost of 10 barg steam = $0.01489 - 0.00527$
= $0.00962 \$ \cdot \text{kg}^{-1}$
= $9.62\$ \cdot \text{t}^{-1}$

Cost of 3 barg steam. Here 10 barg steam from the exit of the first turbine is assumed to be expanded to 3 barg in another turbine. From steam tables, inlet conditions of 10 barg and $251 \,^{\circ}$ C are:

$$H_1 = 2941 \text{ kJ} \cdot \text{kg}^{-1}$$

$$S_1 = 6.880 \text{ kJ} \cdot \text{kg}^{-1} \cdot \text{K}^{-1}$$

1

Turbine outlet conditions for isentropic expansion to 3 barg are:

$$H_{2,IS} = 2732 \text{ kJ} \cdot \text{kg}^{-1}$$

$$S_2 = 6.880 \text{ kJ} \cdot \text{kg}^{-1} \cdot \text{K}^{-1}$$

For a single-stage expansion with isentropic efficiency of 80%:

$$H_2 = H_1 - \eta_{IS}(H_1 - H_{2,IS})$$

= 2941 - 0.8(2941 - 2732)
= 2774 kJ · kg⁻¹

From steam tables, the outlet temperature is $160 \,^{\circ}$ C, which is superheated by $16 \,^{\circ}$ C. Again, it is desirable to have some superheat for the steam fed to the low-pressure main.

Power generation =
$$2941 - 2774$$

= $167 \text{ kJ} \cdot \text{kg}^{-1}$
Value of power generation = $167 \times \frac{0.07}{3600}$
= $0.00325 \$ \cdot \text{kg}^{-1}$
Cost of 3 barg steam = $0.00962 - 0.00325$
= $0.00637 \$ \cdot \text{kg}^{-1}$
= $6.37 \$ \cdot \text{t}^{-1}$

If the steam generated in the boilers is at a very high pressure and/or the ratio of power to fuel costs is high, then the value of low-pressure steam can be extremely low or even negative (see Chapter 23). This simplistic approach to costing steam is often unsatisfactory, especially if the utility system already exists. Steam costs will be considered in more detail in Chapter 23.

The operating cost for cooling water tends to be low relative to the value of both fuel and electricity. The principal operating cost associated with the provision of cooling water is the cost of power to drive the cooling tower fans and cooling water circulation pumps. The cost of cooling duty provided by cooling water is in the order of 1% that of the cost of power. For example, if power costs $0.07 \text{ kW}^{-1} \cdot \text{h}^{-1}$, then cooling water will typically cost $0.07 \times 0.01/3600 = 0.19 \times 10^{-6} \text{ kJ}^{-1}$ or 0.19 GJ^{-1} . Cooling water systems will be discussed in more detail in Chapter 24.

The cost of power required for a refrigeration system can be estimated as a multiple of the power required for an ideal system:

$$\frac{W_{IDEAL}}{Q_C} = \frac{T_H - T_C}{T_C} \tag{2.9}$$

where W_{IDEAL} = ideal power required for the refrigeration cycle

- Q_C = the cooling duty
- T_C = temperature at which heat is taken into the refrigeration cycle (K)
- T_H = temperature at which heat is rejected from the refrigeration cycle (K)

The ratio of ideal to actual power is often around 0.6. Thus:

$$W = \frac{Q_C}{0.6} \left(\frac{T_H - T_C}{T_C} \right) \tag{2.10}$$

where W is the actual power required for the refrigeration cycle.

Example 2.6 A process requires 0.5 MW of cooling at -20 °C. A refrigeration cycle is required to remove this heat and reject it to cooling water supplied at 25 °C and returned at 30 °C. Assuming a minimum temperature difference (ΔT_{min}) of 5 °C and both vaporization and condensation of the refrigerant occur isothermally, estimate the annual operating cost of refrigeration for an electrically driven system operating 8000 hours per year. The cost of electricity is \$0.07 kW⁻¹·h⁻¹.

Solution

$$W = \frac{Q_C}{0.6} \left(\frac{T_H - T_C}{T_C} \right)$$

$$T_H = 30 + 5 = 35 \,^{\circ}\text{C} = 308 \,\text{K}$$

$$T_C = -20 - 5 = -25 \,^{\circ}\text{C} = 248 \,\text{K}$$

$$W = \frac{0.5}{0.6} \left(\frac{308 - 248}{248} \right)$$

$$= 0.202 \,\text{MW}$$

Cost of electricity =
$$0.202 \times 103 \times 0.07 \times 8000$$

= \$113, 120 y⁻¹

More accurate methods to calculate refrigeration costs will be discussed in Chapter 24.

4) Labor cost. The cost of labor is difficult to estimate. It depends on whether the process is batch or continuous, the level of automation, the number of processing steps and the level of production. When synthesizing a process, it is usually only necessary to screen process options that have the same basic character (e.g. continuous), have the same level of automation, have a similar number of processing steps and have the same level of production. In this case, labor costs will be common to all options and hence will not affect the comparison.

If, however, options are to be compared that are very different in nature, such as a comparison between batch and continuous operation, some allowance for the difference in the cost of labor must be made. Also, if the location of the plant has not been fixed, the differences in labor costs between different geographical locations can be important.

5) Maintenance. The cost of maintenance depends on whether processing materials are solids on the one hand or gas and liquid on the other. Handling solids tends to increase maintenance costs. Highly corrosive process fluids increase maintenance costs. Average maintenance costs tend to be around 6% of the fixed capital investment (Peters, Timmerhaus and West, 2003).

2.6 Simple Economic Criteria

To evaluate design options and carry out process optimization, simple economic criteria are needed. Consider what happens to the revenue from product sales after the plant has been commissioned. The sales revenue must pay for both fixed costs that are independent of the rate of production and variable costs, which do depend on the rate of production. After this, taxes are deducted to leave the net profit.

Fixed costs independent of the rate of production include:

- Capital cost repayments
- Routine maintenance
- Overheads (e.g. safety services, laboratories, personnel facilities, administrative services)
- Quality control
- Local taxes
- Labor
- Insurance.

Variable costs that depend on the rate of production include:

- Raw materials
- Catalysts and chemicals consumed in manufacturing (other than raw materials)
- Utilities (fuel, steam, electricity, cooling water, process water, compressed air, inert gases, etc.)
- Maintenance costs incurred by operation
- Royalties
- Transportation costs.

There can be an element of maintenance that is a fixed and an element that is variable. Fixed maintenance costs cover routine maintenance, such as statutory maintenance on safety equipment, that must be carried out irrespective of the rate of production. Variable maintenance costs result from certain items of equipment needing more maintenance as the production rate increases. Also, the royalties that cover the cost of purchasing another company's process technology may have different bases. Royalties may be a variable cost, since they can sometimes be paid in proportion to the rate of production or sales revenue. Alternatively, the royalty might be a single-sum payment at the beginning of the project. In this case, the single-sum payment will become part of the project capital investment. As such, it will be included in the annual capital repayment, and this becomes part of the fixed cost.

Two simple economic criteria are useful in process design: *Economic potential (EP)*:

$$EP =$$
Value of products – Fixed costs
– Variable costs – Taxes (2.11)

Total annual cost (TAC):

$$TAC =$$
 Fixed costs + Variable costs + Taxes (2.12)

When synthesizing a flowsheet, these criteria are applied at various stages when the picture is still incomplete. Hence, it is usually not possible to account for all the fixed and variable costs listed above during the early stages of a project. Also, there is little point in calculating taxes until a complete picture of operating costs and cash flows has been established.

The preceding definitions of economic potential and total annual cost can be simplified if it is accepted that they will be used to compare the relative merits of different structural options in the flowsheet and different settings of the operating parameters. Thus, items that will be common to the options being compared can be neglected.

2.7 Project Cash Flow and Economic Evaluation

As the design progresses, more information is accumulated. The best methods of assessing the profitability of alternatives are based on projections of the cash flows during the project life (Allen, 1980).

Figure 2.2 shows the cash flow pattern for a typical project. The cash flow is a cumulative cash flow. Consider Curve 1 in Figure 2.2. From the start of the project at Point A, cash is spent without any immediate return. The early stages of the project consist of development, design and other preliminary work, which causes the cumulative curve to dip to Point B. This is followed



Figure 2.2

Cash flow pattern for a typical project. (Reproduced from Allen DH, (1980) A Guide to the Economic Evaluation of Projects, by permission of the Institution of Chemical Engineers.)

by the main phase of capital investment in buildings, plant and equipment, and the curve drops more steeply to Point C. Working capital is spent to commission the plant between Points C and D. Production starts at D, where revenue from sales begins. Initially, the rate of production is likely to be below design conditions until full production is achieved at E. At F, the cumulative cash flow is again zero. This is the project breakeven point. Toward the end of the project's life at G, the net rate of cash flow may decrease owing to, for example, increasing maintenance costs, a fall in the market price for the product, and so on.

Ultimately, the plant might be permanently shut down or given a major revamp. This marks the end of the project, H. If the plant is shut down, working capital is recovered and there may be salvage value, which would create a final cash inflow at the end of the project.

The predicted cumulative cash flow curve for a project throughout its life forms the basis for more detailed evaluation. Many quantitative measures or indices have been proposed. In each case, important features of the cumulative cash flow curve are identified and transformed into a single numerical measure as an index.

- 1) *Payback time*. Payback time is the time that elapses from the start of the project (*A* in Figure 2.2) to the breakeven point (*F* in Figure 2.2). The shorter the payback time, the more attractive is the project. Payback time is often calculated as the time to recoup the capital investment based on the mean annual cash flow. In retrofit, payback time is usually calculated as the time to recoup the retrofit capital investment from the mean annual improvement in operating costs.
- 2) *Return on investment (ROI).* Return on investment (ROI) is usually defined as the ratio of average yearly income over the productive life of the project to the total initial investment, expressed as a percentage. Thus, from Figure 2.2:

$$ROI = \frac{KH}{KD} \times \frac{100}{LD}$$
% per year (2.13)

Payback and ROI select particular features of the project cumulative cash flow and ignore others. They take no account of the *pattern* of cash flow during a project. The other indices to be described, net present value and discounted cash flow return, are more comprehensive because they take account of the changing pattern of project net cash flow with time. They also take account of the *time value* of money.

3) Net present value (NPV). Since money can be invested to earn interest, money received now has a greater value than money if received at some time in the future. The net present value of a project is the sum of the present values of each individual cash flow. In this case, the *present* is taken to be the start of a project.

Time is taken into account by discounting the annual cash flow A_{CF} with the rate of interest to obtain the annual discounted cash flow A_{DCF} . Thus at the end of year 1:

$$A_{DCF1} = \frac{A_{CF1}}{(1+i)}$$

at the end of year 2:

$$A_{DCF2} = \frac{A_{CF2}}{\left(1+i\right)^2}$$

and at the end of year *n*:

$$A_{DCFn} = \frac{A_{CFn}}{\left(1+i\right)^n} \tag{2.14}$$

The sum of the annual discounted cash flows over *n* years Σ A_{DCF} is known as the *net present value* (NPV) of the project:

$$NPV = \sum A_{DCF} \tag{2.15}$$

The value of NPV is directly dependent on the choice of the fractional interest rate i and project lifetime n.

Returning to the cumulative cash flow curve for a project, the effect of discounting is shown in Figure 2.2. Curve 1 is the original curve with no discounting, that is, i=0, and the

Example 2.7 A company has the alternative of investing in one of two projects, *A* or *B*. The capital cost of both projects is \$10 million. The predicted annual cash flows for both projects are shown in Table 2.10. Capital is restricted, and a choice is to be made on the basis of the discounted cash flow rate of return, based on a five-year lifetime.

project *NPV* is equal to the final net cash position given by *H*. Curve 2 shows the effect of discounting at a fixed rate of interest and the corresponding project *NPV* is given by *J*. Curve 3 in Figure 2.2 shows a larger rate of interest, but it is chosen such that the *NPV* is zero at the end of the project.

The greater the positive *NPV* for a project, the more economically attractive it is. A project with a negative *NPV* is not a profitable proposition.

4) Discounted cash flow rate of return. The discounted cash flow rate of return is defined as the discount rate *i*, which makes the NPV of a project equal to zero (Curve 3 in Figure 2.2):

$$NPV = \sum A_{DCF} = 0 \tag{2.16}$$

The value of *i* given by this equation is known as the *discounted* cash flow rate of return (DCFRR). It may be found graphically or by trial and error.

Table 2.10

Predicted annual cash flows.

	Cash flows (\$10 ⁶)			
Year	Project A	Project B		
0	-10	-10		
1	1.6	6.5		
2	2.8	5.2		
3	4.0	4.0		
4	5.2	2.8		
5	6.4	1.6		

Solution Project A

Start with an initial guess for *DCFRR* of 20% and increase as detailed in Table 2.11.

Table 2.11

Calculation of DCFRR for Project A.

		DCF 20%		DCF 30%		DCF 25%	
Year	A _{CF}	A _{DCF}	ΣA_{DCF}	A _{DCF}	ΣA_{DCF}	A _{DCF}	ΣA_{DCF}
0	-10	-10	-10	-10	-10	-10	-10
1	1.6	1.33	-8.67	1.23	-8.77	1.28	-8.72
2	2.8	1.94	-6.73	1.66	-7.11	1.79	-6.93
3	4.0	2.31	-4.42	1.82	-5.29	2.05	-4.88
4	5.2	2.51	-1.91	1.82	-3.47	2.13	-2.75
5	6.4	2.57	0.66	1.72	-1.75	2.10	-0.65

Twenty percent is too low since ΣA_{DCF} is positive at the end of year 5. Thirty percent is too high since ΣA_{DCF} is negative at the end of year 5, as is the case with 25%. The answer must be between 20 and 25%. Interpolating on the basis of ΣA_{DCF} the *DCFRR* \approx 23%.

Project B

Again, start with an initial guess for *DCFRR* of 20% and increase as detailed in Table 2.12.

Table 2.12

Calculation of DCFRR for Project B.

		DCF 20%		DCF 40%		DCF 35%	
Year	A _{CF}	A _{DCF}	ΣA_{DCF}	A _{DCF}	ΣA_{DCF}	A _{DCF}	ΣA_{DCF}
0	-10	-10	-10	-10	-10	-10	-10
1	6.5	5.42	-4.58	4.64	-5.36	4.81	-5.19
2	5.2	3.61	-0.97	2.65	-2.71	2.85	-2.34
3	4.0	2.31	1.34	1.46	-1.25	1.63	-0.71
4	2.8	1.35	2.69	0.729	-0.521	0.843	0.133
5	1.6	0.643	3.33	0.297	-0.224	0.357	0.490

From ΣA_{DCF} at the end of year 5, 20% is to low, 40% too high and 35% also too low. Interpolating on the basis of ΣA_{DCF} , the *DCFRR* \approx 38%. Project *B* should therefore be chosen.

2.8 Investment Criteria

Economic analysis should be performed at all stages of an emerging project as more information and detail become available. The decision as to whether to proceed with a project will depend on many factors. There is most often stiff competition within companies for any capital available for investment in projects. However, in an efficient capital market there should be plenty of capital available, providing the returns are high enough. The decision as to where to spend capital on a particular project will, in the first instance, but not exclusively, depend on the economic criteria discussed in the previous section. Criteria that account for the timing of the cash flows (the NPV and DCFRR) should be the basis of the decision making. The higher the value of the NPV and DCFRR for a project, the more attractive it is. The absolute minimum acceptable value of the DCFRR is the market interest rate. If the DCFRR is lower than the market interest rate, it would be better to put the money in the bank. However, since the bank account is less risky than most technical projects, it is prudent to set a return target of at least 5 to 10% higher than the return on a bank account. The essential distinction between NPV and DCFRR is:

- Net present value measures profit but does not indicate how efficiently capital is being used.
- *DCFRR* measures the rate of return that a project might achieve, but gives no indication of how competitive the project would be.

When choosing between competing projects, these will have different cash flow patterns and different capital investments. If the goal is to maximize profit, then projects with the highest *NPV* should be selected. This does not necessarily equate with selecting the project with the highest *DCFRR*. A lower *DCFRR* on a larger investment could yield a higher *NPV* than a higher *DCFRR* on a smaller investment. *NPV* gives a direct cash measure of the attractiveness of a project. For competing projects with different capital investments, a quick way to compare between the projects is the *investment efficiency*:

Investment efficiency =
$$\frac{NPV}{Capital investment}$$
 (2.17)

Predicting future cash flows for a project is extremely difficult. There are many uncertainties, including the project life. Also, the appropriate interest rate will not be known with certainty. The acceptability of the rate of return will depend on the risks associated with the project and the company investment policy. For example, a *DCFRR* of 20% might be acceptable for a low-risk project. A higher return of, say, 30% might be demanded of a project with some risk, whereas a high-risk project with significant uncertainty might demand a 50% *DCFRR*.

The sensitivity of the economic analysis to the underlying assumptions should always be tested. A sensitivity analysis should be carried out to test the sensitivity of the economic analysis to:

• errors in the capital cost estimate,

- delays in the start-up of the project after the capital has been invested (particularly important for a high capital cost project),
- changes in the cost of raw materials,
- changes in the sales price of the product,
- reduction in the market demand for the product, and so on.

When carrying out an economic evaluation, the magnitude and timing of the cash flows, the project life and interest rate are not known with any certainty. However, providing that consistent assumptions are made for projections of cash flows and the assumed rate of interest, the economic analysis can be used to choose between competing projects. It is important to compare different projects and options within projects on the basis of consistent assumptions. Thus, even though the evaluation will be uncertain in an absolute sense, it can still be meaningful in a relative sense for choosing between options. Because of this, it is important to have a reference against which to judge any project or option within a project.

However, the final decision to proceed with a project will be influenced as much by business strategy as by the economic measures described above. The business strategy might be to gradually withdraw from a particular market, perhaps because of adverse long-term projections of excessive competition, even though there might be short-term attractive investment opportunities. The long-term business strategy might be to move into different business areas, thereby creating investment priorities. Priority might be given to increasing market share in a particular product to establish business dominance in the area and achieve long-term global economies of scale in the business.

2.9 Process Economics – Summary

Process economics is required to evaluate design options, carry out process optimization and evaluate overall project profitability. Two simple criteria can be used:

- economic potential,
- total annual cost.

These criteria can be used at various stages in the design without a complete picture of the process.

The dominant operating cost is usually raw materials. However, other significant operating costs involve catalysts and chemicals consumed other than raw materials, utility costs, labor costs and maintenance.

Capital costs can be estimated by applying installation factors to the purchase costs of individual items of equipment. However, there is considerable uncertainty associated with cost estimates obtained in this way, as equipment costs are typically only 20 to 40% of the total installed costs, with the remainder based on factors. Utility investment, off-site investment and working capital are also needed to complete the capital investment. The capital cost can be annualized by considering it as a loan over a fixed period at a fixed rate of interest. As a more complete picture of the project emerges, the cash flows through the project life can be projected. This allows more detailed evaluation of project profitability on the basis of cash flows. Net present value can be used to measure the profit, taking into account the time value of money. The discounted cash flow rate of return measures how efficiently the capital is being used.

Overall, there are always considerable uncertainties associated with an economic evaluation. In addition to the errors associated with the estimation of capital and operating costs, the project life or interest rates are not known with any certainty. The important thing is that different projects, and options within projects, are compared on the basis of consistent assumptions. Thus, even though the evaluation will be uncertain in an absolute sense, it will still be meaningful in a relative sense for choosing between options.

2.10 Exercises

- 1. A new agitated reactor with a new external shell-and-tube heat exchanger and new centrifugal pump are to be installed in an existing facility. The agitated reactor is to be glass-lined, which can be assumed to have an equipment cost of three times the cost of a carbon steel vessel. The heat exchanger, pump and associated piping are all high-grade stainless steel. The equipment is rated for moderate pressure. The reactor has a volume of 9 m³, the heat exchanger an area of 50 m² and the pump has a power of 5 kW. No significant investment is required in utilities, off-sites, buildings, site preparation or working capital. Using Equation 2.1 and Table 2.1 (extrapolating beyond the range of the correlation if necessary), estimate the cost of the project (CE Index of Equipment = 680.0).
- 2. Steam is distributed on a site via high-pressure and low-pressure steam mains. The high-pressure main is at 40 bar and 350 °C. The low-pressure main is at 4 bar. The high-pressure steam is generated in boilers. The overall efficiency of steam generation and distribution is 75%. The low-pressure steam is generated by expanding the high-pressure stream through steam turbines with an isentropic efficiency of 80%. The cost of fuel in the boilers is 3.5 \$•GJ⁻¹ and the cost of electricity is \$0.05 kW⁻¹⋅h⁻¹. The boiler feedwater is available at 100 °C with a heat capacity of 4.2 kJ⋅kg⁻¹⋅K⁻¹. Estimate the cost of the high-pressure and low-pressure steam.
- **3.** A refrigerated distillation condenser has a cooling duty of 0.75 MW. The condensing stream has a temperature of -10 °C. The heat from a refrigeration circuit can be rejected to cooling water at a temperature of 30 °C. Assuming a temperature difference in the distillation condenser of 5 °C and a temperature difference for heat rejection from refrigeration to cooling water of 10 °C, estimate the power requirements for the refrigeration.
- **4.** Acetone is to be produced by the dehydrogenation of an aqueous solution of isopropanol according to the reaction:

$$(CH_3)_2 CHOH \rightarrow CH_3 COCH_3 + H_2 \\ {}_{Isopropanol} Acetone + H_2 Hydrogen$$

Table 2.13

Data for Exercise 4.

Component	Flowrate in vapor (kmol·h ⁻¹)	Raw material value (\$·kmol ⁻¹)	Fuel value (\$·kmol ⁻¹)
Hydrogen	51.1	0	0.99
Acetone	13.5	34.8	6.85

The effluent from the reactor enters a phase separator that separates vapor from liquid. The liquid contains the bulk of the product and the vapor is a waste stream. The vapor stream is at a temperature of $30 \,^{\circ}$ C and an absolute pressure of 1.1 bar. The component flowrates in the vapor stream are given in Table 2.13, together with their raw material values and fuel values. Three options are to be considered:

- a) Burn the vapor in a furnace.
- b) Recover the acetone by absorption in water recycled from elsewhere in the process with the tail gas being burnt in a furnace. It is expected that 99% will be recovered by this method at a cost of 1.8 \$⋅kmol⁻¹ acetone recovered.
- c) Recover the acetone by condensation using refrigerated coolant with the tail gas being burnt in a furnace. It is anticipated that a temperature of -10° C will need to be achieved in the condenser. It can be assumed that the hydrogen is an inert gas that will not dissolve in the liquid acetone. The vapor pressure of acetone is given by

$$\ln P = 10.031 - \frac{2940.5}{T - 35.93}$$

where P = pressure (bara)

T = absolute temperature (K)

The cost of refrigerant is 11.5 GJ^{-1} , the mean molal heat capacity of the vapor is $40 \text{ kJ} \cdot \text{kmol}^{-1} \cdot \text{K}^{-1}$ and the latent heat of acetone is 29,100 kJ \cdot kmol^{-1}.

Calculate the economic potential of each option given the data in Table 2.13.

- 5. A process for the production of cellulose acetate fiber produces a waste stream containing mainly air but with a small quantity of acetone vapor. The flowrate of air is $300 \text{ kmol} \cdot \text{h}^{-1}$ and that of acetone is $4.5 \text{ kmol} \cdot \text{h}^{-1}$. It is proposed to recover the acetone from the air by absorption into water followed by distillation of the acetone–water mixture. The absorber requires a flow of water 2.8 times that of the air.
 - a) Assuming acetone costs 34.8 \$·kmol⁻¹, fresh water is to be used in the absorber at a cost of \$0.004 kmol⁻¹ and the process operates for $8000 \text{ h} \cdot \text{y}^{-1}$, calculate the maximum economic potential assuming complete recovery of the acetone.

Table 2.14

Cash flows for two competing projects.

	Cash flows \$1000		
Year	Project A	Project B	
0	-1000	-1000	
1	150	500	
2	250	450	
3	350	300	
4	400	200	
5	400	100	

- **b**) The absorber and the absorber and distillation column are both assumed to operate at 99% recovery of acetone. If the product acetone overhead from the distillation column must be 99% pure and there is assumed to be no loss of water in the air stream, sketch the flowsheet for the system and calculate the flows of acetone and water to and from the distillation column.
- c) Calculate the revised economic potential to allow for incomplete recovery in the absorption and distillation columns. In addition, the effluent from the bottom of the distillation column will cost \$50 for each kmol of acetone plus \$0.004 for each kmol of water to treat before it can be disposed of.
- **d)** How might the costs be decreased for the same recoveries and purities in the separations?
- **6.** A company has the option of investing in one of the two projects, *A* or *B*. The capital cost of both projects is \$1,000,000. The predicted annual cash flows for both projects are shown in Table 2.14. For each project, calculate:
 - a) The payback time for each project in terms of the average annual cash flow.
 - b) The return on investment.
 - c) The discounted cash flow rate of return.

What do you conclude from the result?

- 7. A company has the option of investing in one of the two projects, *A* or *B*. The capital cost of both projects is \$1,000,000. The predicted annual cash flows for both projects are shown in Table 2.15. For each project, calculate:
 - a) The payback time for each project in terms of the average annual cash flow.
 - b) The return on investment.
 - c) The discounted cash flow rate of return.

What do you conclude from the result?

8. A process has been developed for a new product for which the market is uncertain. A plant to produce $50,000 \text{ t} \cdot \text{y}^{-1}$ requires

Table 2.15

Cash flows for two competing projects.

	Cash flows \$1000		
Year	Project A	Project B	
0	-1000	-1000	
1	150	500	
2	250	450	
3	350	300	
4	400	200	
5	400	100	

an investment of \$10,000,000 and the expected project life is five years. Fixed operating costs are expected to be $750,000 \cdot y^{-1}$ and variable operating costs (excluding raw materials) are expected to be $40 \cdot t^{-1}$ product. The stoichiometric raw material costs are $80 \cdot t^{-1}$ product. The yield of product per ton of raw material is 80%. Tax is paid in the same year as the relevant profit is made at a rate of 20%. Calculate the selling price of the product to give a minimum acceptable discounted cash flowrate of return of 15% year.

- **9.** How can the concept of simple payback be improved to give a more meaningful measure of project profitability?
- 10. It is proposed to build a plant to produce $170,000 \text{ ty}^{-1}$ of a commodity chemical. A study of the supply and demand projections for the product indicates that current installed capacity in the industry is $6.8 \times 10^6 \text{ ty}^{-1}$, whereas total production is running at $5.0 \times 10^6 \text{ ty}^{-1}$. Maximum plant utilization is thought to be around 90%. If the demand for

the product is expected to grow at 8% per year and it will take 3 years to commission a new plant from the start of a project, what do you conclude about the prospect for the proposed project?

References

- Allen DH (1980) A Guide to the Economic Evaluation of Projects, IChemE, Rugby, UK.
- Anson HA (1977) A New Guide to Capital Cost Estimating, IChemE, UK.
- Bauman HC (1955) Estimating Costs of Process Auxiliaries, *Chem Eng Prog*, 51: 45.
- Bravo JL (1997) Select Structured Packings or Trays? *Chem Eng Prog*, July: 36.
- Brennan D (1998) Process Industry Economics, IChemE, UK.
- Gary JH and Handwerk GE (2001) Petroleum Refining Technology and Economics, 4th Edition, Marcel Dekker.
- Gerrard AM (2000) *Guide to Capital Cost Estimating*, 4th Edition, IChemE, UK.
- Guthrie KM (1969) Data and Techniques for Preliminary Capital Cost Estimating, *Chem Eng*, **76**: 114.
- Hall RS, Matley J and McNaughton KJ (1982) Current Costs of Process Equipment, *Chem Eng*, **89**: 80.
- Hall R.S, Vatavuk WM and Matley J (1988) Estimating Process Equipment Costs, Chem Eng, 95: 66.
- Hand WE (1958) From Flowsheet to Cost Estimate, Petrol Refiner, 37: 331.
- Holland FA, Watson FA and Wilkinson JK (1983) Introduction to Process Economics, 2nd Edition, John Wiley & Sons, Inc., New York.
- Lang HJ (1947) Cost Relationships in Preliminary Cost Estimation, *Chem Eng*, 54: 117.
- Mulet A, Corripio AB and Evans LB (1981a) Estimate Costs of Pressure Vessels Via Correlations, *Chem Eng*, Oct: 145.
- Mulet A, Corripio AB and Evans LB (1981b) Estimate Costs of Distillation and Absorption Towers Via Correlations, *Chem Eng*, **Dec**: 77.
- Peters MS, Timmerhaus KD and West RE (2003) *Plant Design and Economics* for Chemical Engineers, 5th Edition, McGraw-Hill, New York.
- Remer DS and Chai LH (1990) Design Cost Factors for Scaling-up Engineering Equipment, *Chem Eng Prog*, **86**: 77.
- Ulrich GD (1984) A Guide to Chemical Engineering Process Design and Economics, John Wiley & Sons, Inc., New York.

Chapter 3

Optimization

3.1 Objective Functions

Optimization will almost always be required at some stage in a process design. It is usually not necessary for a designer to construct an optimization algorithm in order to carry out an optimization, as general-purpose software is usually available for this. However, it is necessary for the designer to have some understanding of how optimization works in order to avoid the pitfalls that can occur. More detailed accounts of optimization can be found elsewhere (Floudas, 1995; Biegler, Grossmann and Westerberg, 1997; Edgar, Himmelblau and Lasdon, 2001).

Optimization problems in process design are usually concerned with maximizing or minimizing an *objective function*. The objective function is a measure of the quality of the solution and might typically cause economic potential to be maximized or cost to be minimized. For example, consider the recovery of heat from a hot waste stream. A heat exchanger could be installed to recover the waste heat. The heat recovery is illustrated in Figure 3.1a as a plot of temperature versus enthalpy. There is heat available in the hot stream to be recovered to preheat the cold stream. However, how much heat should be recovered? Expressions can be written for the recovered heat as:

$$Q_{REC} = m_H C_{P,H} (T_{H,in} - T_{H,out})$$
(3.1)

$$Q_{REC} = m_C C_{P,C} (T_{C,out} - T_{C,in})$$
(3.2)

where Q_{REC} = recovered heat m_H, m_C = mass flowrates of the hot and cold streams $C_{P,H}, C_{P,C}$ = specific heat capacity of the hot and cold streams $T_{H,in}, T_{H,out}$ = hot stream inlet and outlet temperatures $T_{C,in}, T_{C,out}$ = cold stream inlet and outlet temperatures The effect of the heat recovery is to decrease the energy requirements. Hence, the energy cost of the process:

$$Energy \ cost = (Q_H - Q_{REC})C_{Energy} \tag{3.3}$$

where Q_H = process hot utility requirement prior to heat recovery from the waste stream C_{Energy} = unit cost of energy

There is no change in cost associated with cooling as the hot stream is a waste stream being sent to the environment. An expression can also be written for the heat transfer area of the recovery exchanger (see Chapter 12):

$$A = \frac{Q_{REC}}{U \Delta T_{LM}} \tag{3.4}$$

where A = heat transfer area

U = overall heat transfer coefficient

 ΔT_{LM} = logarithmic mean temperature difference

$$=\frac{(T_{H,in} - T_{C,out}) - (T_{H,out} - T_{C,in})}{\ln\left[\frac{T_{H,in} - T_{C,out}}{T_{H,out} - T_{C,in}}\right]}$$

In turn, the area of the exchanger can be used to estimate the annualized capital cost:

Annualized capital cost =
$$(a + bA^c)AF$$
 (3.5)

where a, b, c = cost coefficients

AF = annualization factor (see Chapter 2)

Suppose that the mass flowrates, heat capacities and inlet temperatures of both streams are fixed and the current hot utility requirement, unit cost of energy, overall heat transfer coefficient, cost coefficients and annualization factor are known. The five Equations 3.1 to 3.5, together with the specifications for the 13 variables m_H , m_C , $C_{P,H}$, $C_{P,C}$, $T_{H,in}$, $T_{C,in}$, U, a, b, c, AF, Q_H and C_{Energy} , constitute 18 *equality constraints*. In addition to these 13 variables, there are a further six unknown variables Q_{REC} , $T_{H,out}$, $T_{C,out}$, *energy cost, annualized capital cost* and A. Thus, there are 18 equality constraints and 19 variables, and the problem cannot be

Chemical Process Design and Integration, Second Edition. Robin Smith. © 2016 John Wiley & Sons, Ltd. Published 2016 by John Wiley & Sons, Ltd. Companion Website: www.wiley.com/go/smith/chemical2e



Figure 3.1

Recovery of heat from a waste steam involves a trade-off between reduced energy cost and increased capital cost of heat exchanger.

solved. For the system of equations (equality constraints) to be solved, the number of variables must be equal to the number of equations (equality constraints). It is underspecified. Another specification (equality constraint) is required to solve the problem; there is one *degree of freedom*. This degree of freedom can be optimized. In this case, it is the sum of the annualized energy and capital costs (i.e. the total cost), as shown in Figure 3.1b.

If the mass flowrate of the cold stream through the exchanger had not been fixed, there would have been one fewer equality constraint, and this would have provided an additional degree of freedom and the optimization would have been a two-dimensional optimization. Each degree of freedom provides an opportunity for optimization.

Figure 3.1b illustrates how the investment in the heat exchanger is optimized. As the amount of recovered heat increases, the cost of energy for the system decreases. On the other hand, the size and capital cost of the heat exchange equipment increase. The increase in size of heat exchanger results from the greater amount of heat to be transferred. Also, because the conditions of the waste stream are fixed, as more heat is recovered from the waste stream, the temperature differences in the heat exchanger become lower, causing a sharp rise in the capital cost. In theory, if the recovery was taken to its limit of zero temperature difference, an infinitely large heat exchanger with infinitely large capital cost would be needed. The costs of energy and capital cost can be expressed on an annual basis, as explained in Chapter 2, and combined as shown in Figure 3.1b to obtain the total cost. The total cost shows a minimum, indicating the optimum size of the heat exchanger.

Given a mathematical model for the objective function in Figure 3.1b, finding the minimum point should be straightforward and various strategies could be adopted for this purpose. The objective function is continuous. Also, there is only one extreme point. Functions with only one extreme point (maximum or minimum) are termed unimodal. By contrast, consider the objective functions in Figure 3.2. Figure 3.2a shows an objective function to be minimized that is discontinuous. If the search for the minimum is started at point x_1 , it could be easily concluded that the optimum point is at x_2 , whereas the true optimum is at x_3 . Discontinuities can present problems to optimization algorithms searching for the optimum. Figure 3.2b shows an objective function that has a number of points where the gradient is zero. These points where the gradient is zero are known as *stationary* points. Functions that exhibit a number of stationary points are known as multimodal. If the function in Figure 3.2b is to be minimized, it has a *local optimum* at x_2 . If the search is started at x_1 , it could be concluded that the global optimum is at x_2 . However, the global optimum is at x_3 . The gradient is zero at x_4 , which is a saddle point. By considering the gradient only, this could be confused with a maximum or minimum. Thus, there is potentially another local optimum at x_4 . Like discontinuities, multimodality presents problems to optimization algorithms. It is clear that it is not sufficient to find a point of zero gradient in the objective function in order to ensure that the optimum has been reached. Reaching a point of zero gradient is a necessary condition for optimality but not a sufficient condition.

To establish whether an optimal point is a local or a global optimum, the concepts of *convexity* and *concavity* must be introduced. Figure 3.3a shows a function to be minimized. In Figure 3.3a, if a straight line is drawn between any two points on the function, then the function is *convex* if all the values of this line lie above the curve. Similarly, if an objective function is to be maximized, as shown in Figure 3.3b, and a straight line drawn between any two points on the function, then if all values on this





line lie below the curve, the function is concave. A convex or concave function provides a single optimum. Thus, if a minimum is found for a function that is to be minimized and is known to be convex, then it is the global optimum. Similarly, if a maximum is found for a function that is to be maximized and known to be concave, then it is the global optimum. On the other hand, a nonconvex or nonconcave function may have multiple local optima.

Searching for the nonlinear optimum in Figures 3.1 to 3.3 constitutes a one-dimensional search. If the optimization involves two variables, say x_1 and x_2 corresponding to the function $f(x_1, x_2)$, it can be represented as a contour plot, as shown in Figure 3.4. Figure 3.4 shows a function $f(x_1, x_2)$ to be minimized. The contours represent lines of uniform values of the objective function. The objective function in Figure 3.4 is multimodal, involving a local optimum and a global optimum. The concepts of convexity and concavity can be extended to problems with more than one dimension (Edgar, Himmelblau and Lasdon, 2001). The objective function in Figure 3.4 is nonconvex. A straight line cannot be drawn between any two points on the surface represented by the contours to ensure that all points on the straight line are below the surface. These concepts can be expressed in a formal way mathematically, but is outside the scope of this text (Floudas, 1995; Biegler, Grossmann and Westerberg, 1997; Edgar, Himmelblau and Lasdon, 2001).

The optimization might well involve more than two variables in a multivariable optimization, but in this case it is difficult to visualize the problem.

 x_2



(b) A concave function to be maximized.

Figure 3.3 Convex and concave functions.



Figure 3.4

A contour plot of a multimodal function to be minimized.

3.2 Single-Variable Optimization

Searching for the optimum (minimum) for the objective function in Figure 3.1b involves a one-dimensional search across a single variable. In the case of Figure 3.1b, a search is made for the amount of heat to be recovered. An example of a method for a single-variable search is *region elimination*. The function is assumed to be unimodal. Figure 3.5 illustrates the approach. In Figure 3.5 the search region is for convenience set between 0 and 1. This is not a restriction on the approach, but for simplicity of explanation. The approach in Figure 3.5 has been to divide the solution space into four equal intervals and to evaluate the function at both the boundaries and the internal points, leading in Figure 3.5a to the minimum amongst these five evaluations at x_1 . It cannot yet be determined exactly where the minimum point is. However, if the function is assumed to be unimodal, the minimum point must lie somewhere between 0 and x_2 . The region between x_2 and 1 can therefore be eliminated. Figure 3.5b shows a different example with a different outcome from the five function evaluations in which the minimum point is at x_2 . In this case, because the function is assumed to be unimodal, the minimum point must lie between x_1 and x_3 , with the regions between 0 and x_1 and between x_3 and 1 eliminated. A third example with a different outcome is shown in Figure 3.5c, in which the minimum is located at x_3 . In this case, the region between 0 and x_2 can be eliminated. In each of the cases in Figure 3.5, the search region has been halved by the evaluation of the function at five points.

The simple strategy adopted for region elimination in Figure 3.5 can in some cases eliminate more than half of the search region. Figure 3.6a shows a different example in which the minimum of the five points is on the lower boundary. In this case, the minimum must lie somewhere between 0 and x_1 and the region between x_1 and 1 can be eliminated. Figure 3.6b shows an example where the minimum of the five points is on the upper boundary. In this case, the minimum must lie somewhere between 0 and x_3 can be eliminated. In Figure 3.6c the example shows two of the function evaluations being equal and minimum, in this case at x_1 and x_2 . This might be an unusual circumstance, but if it happens the minimum can be located between x_1 and x_2 . Thus, in the cases shown in Figure 3.6, three-quarters of the search regions have been eliminated by the evaluation of the function at five points.

The strategy shown in Figures 3.5 and 3.6 has allowed part of the search region to be eliminated and the location of the optimum point narrowed down. The location of the optimal point can be narrowed down further by repeating the strategy in the remaining region that has not been eliminated. The location of the optimum is identified once the region has been narrowed down to within the desired tolerance.

For Figures 3.5 and 3.6, intervals were chosen to be equally spaced. In most cases, this allowed the search region to be halved. In special circumstances a greater part of the search region could be eliminated. However, a symmetrical location of the search points leads to a more efficient search method known as *Golden Section*.



Figure 3.5 Region elimination for the optimization of a single variable.



Figure 3.6 Region elimination can in some cases eliminate much greater ranges.

Consider Figure 3.7a. Again the search region is for convenience set between 0 and 1. In Figure 3.7a two search points have been located in such a way that the ratio of the whole region [l + (1 - l)] to the larger subregion *l* is the same as the ratio of the ratio of the larger subregion *l* to the smaller region (1 - I). Thus:

$$\frac{l+(1-l)}{l} = \frac{l}{1-l}$$
(3.6)

Since:

$$l + (1 - l) = 1 \tag{3.7}$$

substituting and rearranging gives:

$$l^2 + l - 1 = 0 \tag{3.8}$$

This quadratic equation can be solved to give l=0.618. In Figure 3.7b, the function has been evaluated at these two search points. If the function is assumed to be unimodal, then in this case the region between x_2 and 1 can be eliminated. The remaining subregion of length l now has a search point located interior to it such that the symmetry of the search pattern is maintained. Hence in Figure 3.7c the new search point is located at $x_3=0.382 \times 0.618=0.236$. Evaluation of the function at this search point allows the region between x_1 and x_2 to be further eliminated, assuming a unimodal function. This process can then be repeated in the remaining region in order to identify more precisely the location of the optimum. Each new search point reduces the region of the search space by 0.618. For most problems this is more efficient than the five-point method illustrated in Figure 3.5.

A number of other methods for single-variable optimization is also possible (Edgar, Himmelblau and Lasdon, 2001).



Figure 3.7 Golden section region elimination.

3.3 Multivariable Optimization

The problem of multivariable optimization is illustrated in Figure 3.4. Search methods used for multivariable optimization can be classified as *deterministic* and *stochastic*. Deterministic methods follow a predetermined search pattern and do not involve any guessed or random steps. On the other hand, stochastic search methods use random choice to guide the search. Stochastic search methods generate a randomized path to the solution on the basis of probabilities.

- 1) *Deterministic methods.* Deterministic methods can be further classified into *direct* and *indirect* search methods. Direct search methods do not require derivatives (gradients) of the function. Indirect methods use derivatives, even though the derivatives might be obtained numerically rather than analytically.
 - a) Direct search methods. An example of a direct search method is a *univariate search*, as illustrated in Figure 3.8. All of the variables except one are fixed and the remaining variable is optimized. Once a minimum or maximum point has been reached, this variable is fixed and another variable optimized, with the remaining variables being fixed. This is repeated until there is no further improvement in the objective function. Figure 3.8 illustrates a two-dimensional search in which x_1 is first fixed and x_2 optimized. Then x_2 is fixed and x_1 optimized, and so on until no further improvement in the objective function is obtained. In Figure 3.8, the univariate search is able to locate the global optimum. It is easy to see that if the starting point for the search in



A univariate search.

Figure 3.8 had been at a lower value of x_I , then the search would have located the local optimum, rather than the global optimum. For searching multivariable optimization problems, often the only way to ensure that the global optimum has been reached is to start the optimization from different initial points.

Another example of a direct search is a sequential simplex search. The method uses a regular geometric shape (a *simplex*) to generate search directions. In two dimensions, the simplest shape is an equilateral triangle. In three dimensions, it is a regular tetrahedron. The objective function is evaluated at the vertices of the simplex, as illustrated in Figure 3.9. The objective function must first be evaluated at the Vertices A, B and C. The general direction of search is projected away from the worst vertex (in this case Vertex A) through the centroid of the remaining vertices (B and C) (Figure 3.9a). A new simplex is formed by replacing the worst vertex by a new point that is the mirror image of the simplex (Vertex D), as shown in Figure 3.9a. Then Vertex D replaces Vertex A, as Vertex A is an inferior point. The simplex vertices for the next step are B, C and D. This process is repeated for successive moves in a zigzag fashion, as shown in Figure 3.9b. The direction of search can change, as illustrated in Figure 3.9c. When the simplex is close to the optimum, there may be some repetition of simplexes, with the search going around in circles. If this is the case, then the size of the simplex should be reduced.

b) Indirect search methods. Indirect search methods use derivatives (gradients) of the objective function. The derivatives may be obtained analytically or numerically. Many methods are available for indirect search. An example is the method of steepest descent in a minimization problem. The direction of steepest descent is the search direction that gives the maximum rate of change for the objective function from the current point. The method is illustrated in Figure 3.10. One problem with this search method is that the appropriate step size is not known and this is under circumstances when the gradient might change significantly during the search. Another problem is that the search can slow down significantly as it reaches the optimum point. If a search is made for the maximum in an objective function, then the analogous search is one of steepest ascent.

One fundamental practical difficulty with both the direct and indirect search methods is that, depending on the shape of the solution space, the search can locate local optima, rather than the global optimum. Often, the only way to ensure that the global optimum has been reached is to repeat the optimization starting from different initial points and repeat the process.

2) Stochastic search methods. In all of the optimization methods discussed so far, the algorithm searches the objective function seeking to improve the objective function at each step, using information such as gradients. Unfortunately, as already noted, this process can mean that the search is attracted towards a local



Figure 3.9

The simplex search.

optimum. On the other hand, stochastic search methods use random choice to guide the search and can allow deterioration of the objective function during the search. It is important to recognize that a randomized search does not mean a directionless search. Stochastic search methods generate a randomized path to the solution on the basis of probabilities. Improvement in the objective function becomes the ultimate rather than the immediate goal, and some deterioration in the objective





function is tolerated, especially during the early stages of the search. In searching for a minimum in the objective function, rather than the search always attempting to go downhill, stochastic search methods allow the search to also sometimes go uphill. Similarly, if the objective function is to be maximized, stochastic search methods allow the search to sometimes go downhill. This helps to reduce the problem of being trapped in a local optimum.

Stochastic search methods do not need auxiliary information, such as derivatives, in order to progress. They only require an objective function for the search. This means that stochastic search methods can handle problems in which the calculation of the derivatives would be complex and cause deterministic methods to fail.

Two of the most popular stochastic search methods are *simulated annealing* and *genetic algorithms*.

a) Simulated annealing. Simulated annealing emulates the physical process of annealing of metals (Metropolis *et al.*, 1953; Kirkpatrick, Gelatt and Vecchi, 1983). In the physical process, at high temperatures, the molecules of the liquid metal move freely with respect to one another. If the liquid is cooled slowly, thermal mobility is lost. The atoms are able to line themselves up and form perfect crystals. This crystal state is one of minimum energy for the system. If the liquid metal is cooled quickly, it does not reach this state but rather ends up in a polycrystalline or amorphous state having higher energy. Therefore the essence of the process is slow cooling, allowing ample time for redistribution of the atoms as they lose mobility to reach a state of minimum energy.

With this physical process in mind, an algorithm can be suggested in which a system *moves* from one point to another. A *move* might be a change in the temperature,

pressure, flowrate, etc., in the problem. Rules must be defined to create the moves, for example the size of step change in temperature. The moves are then selected randomly. If the objective function from iteration *i* to (i + 1) changes from E_i to E_{i+1} and the objective function is to be minimized, then a move in which $(E_{i+1} - E_i)$ is negative (i.e. the objective function improves) is accepted. If $(E_{i+1} - E_i)$ is positive, the objective function deteriorates, but this does not mean that the move will be rejected. The probability of accepting a change in which the objective function deteriorates can be assumed to follow a relationship similar to the Boltzmann probability distribution, maintaining the analogy with physical annealing (Metropolis *et al.*, 1953):

$$P = \exp[-(E_{i+1} - E_i)/T]$$
(3.9)

where P is the probability, E is the objective function (the analogy of energy) and T is a control parameter (the analogy of annealing temperature). To determine whether a move is accepted or not according to Equation 3.9, a random number generator creates random numbers between zero and unity. If Equation 3.9 predicts a probability greater than the random number generator, then the move is accepted. If it is less, then the move is rejected and another move is attempted instead. It is evident from Equation 3.9 that when T is a large value, virtually all modifications made to the system are accepted. When T is close to zero, virtually all modifications that yield $(E_{i+1}-E_i) > 0$ are rejected. Thus, Equation 3.9 dictates whether a move is accepted or rejected. In this way, the method will always accept a downhill step when minimizing an objective function, while sometimes it will take an uphill step.

The algorithm starts with a high value of T, allowing a high probability of moves to be accepted that cause deterioration in the objective function. The value of Tis gradually decreased as the search progresses. At any value of T a series of random moves is made that constitutes a *Markov chain*. In a Markov chain, the algorithm moves from one state to another in a chain-like manner in a random process in which the next state depends only on the current state and not on the past. It is necessary to specify the length of the Markov chain. A short Markov chain length reduces the likelihood of achieving equilibrium at each value of T. A long Markov chain would make the computation excessively expensive. The appropriate Markov chain length depends on the type of problem to be solved.

A *cooling schedule* needs to be proposed. Aarts and Van Laarhoven (1985) suggested a cooling schedule expressed in the form:

$$T_{k+1} = T_k \left(1 + \frac{\ln(1+\theta)T_k}{3\sigma(T_k)} \right)$$
(3.10)

- where T_k = annealing temperature setting k for a Markov chain
 - σ = standard deviation of the objective functions generated at various Markov moves at annealing temperature T_k
 - θ = cooling parameter that controls how fast the annealing temperature is decreased

The bigger the value of θ , the faster the cooling process, but the greater the likelihood of being trapped into a local optimum. The performance of the algorithm is greatly affected by the initial value of the annealing temperature T_0 . Too high a temperature will unnecessarily increase the time required for the algorithm to converge. On the other hand, too low a temperature will limit the number and magnitude of the uphill moves accepted, thus losing the ability of the algorithm to escape from local optima. The optimal initial annealing temperature depends on the nature of the problem and the scale of the objective function.

Thus, the way the algorithm works is to set an initial value for the annealing temperature. At this setting, a series of random moves are made. Equation 3.9 dictates whether an individual move is accepted or rejected. The annealing temperature is lowered and a new series of random moves is made, and so on. As the annealing temperature is lowered, the probability of accepting deterioration in the objective function, as dictated by Equation 3.9, decreases. In this way, the acceptability for the search to move uphill in a minimization or downhill during maximization is gradually withdrawn.

Whilst simulated annealing can be extremely powerful in solving difficult optimization problems with many local optima, it has a number of disadvantages. Initial and final values of the annealing temperature, an annealing schedule and the number of random moves for each Markov chain must be specified, which depend on the class of problem.

b) Genetic algorithms. Genetic algorithms draw their inspiration from biological evolution (Goldberg, 1989). Unlike all of the optimization methods discussed so far, which move from one point to another, a genetic algorithm moves from one set of points (termed a *population*) to another set of points. Populations of strings called *chromosomes* are created to represent an underlying set of parameters (e.g. temperatures, pressures or concentrations). A simple genetic algorithm exploits three basic operators: reproduction (selection), crossover and mutation.

Reproduction or *selection* is a process in which individual members of a population are copied according to the value of the objective function in order to generate new population sets. The operator is inspired by "natural selection" and the "survival of the fittest". The easiest way to understand selection is to make an analogy with a roulette wheel. In a roulette wheel, the probability of selection is proportional to the area of the slots in the wheel. In a genetic algorithm, the probability of selection is proportional to the *fittness* (the objective function). Although the selection procedure is stochastic, fitter chromosomes (with better values of the objective function) are given a better chance of selection (survival). The selection operator can be implemented in a genetic algorithm in many ways (Goldberg, 1989).

Crossover involves the combination of genetic material from two successful *parents* to form two *offspring* (children). Crossover involves cutting two parent chromosomes at random points and combining them differently to form new offspring. The crossover point is generated randomly. Crossover spreads good properties amongst the population. The fraction of new population generated by crossover is generally large (as observed in nature) and is controlled stochastically. The crossover operator can be implemented in a genetic algorithm in many ways (Goldberg, 1989).

Mutation creates new chromosomes by randomly changing (mutating) parts of chromosomes, but (as with nature) with a low probability of occurring. A random change is made in one of the genes in order to preserve diversity. Mutation creates a new solution in the neighborhood of a point undergoing mutation. However, the probability of mutation is usually set to be low. If the probability of mutation is set too high, the search will turn into a primitive random search.

Thus, a genetic algorithm works by first generating an initial population randomly. The population is evaluated according to its fitness (value of the objective function). A selection operator then provides an intermediate population using stochastic selection but biased towards survival of the fittest. Crossover and mutation operators are then applied to the intermediate population to create a new generation of population. The new population is evaluated according to its fitness and the search is continued with further selection, crossover and mutation until the population meets the required convergence criterion (e.g. maximum number of generations or the difference between the average and maximum fitness value).

Consider the example of a function f(x, y) for which x and y must be manipulated to minimize the function, subject to $0.5 \le x \le 7.5$ and $20.3 \le y \le 80.0$. Values of x and y can be normalized such that $0 \le x \le 1$ and $0 \le y \le 1$; for example, for, say, (x, y) = (3.2, 52.8). Normalizing to four digits gives x = (3.2 - 0.5)/(7.5 - 0.5) = 0.3857 and y = (52.8 - 20.3)/(7.5 - 0.5) = 0.3857(80.0-20.3) = 0.5444. The two normalized values can then be encoded as genes [3857] and [5444] into the chromosome as {38575444}. This chromosome encodes (x, y) = (3.2, 52.8). The genetic algorithm starts by creating an initial population that is randomly generated, for example {27066372}, {83876194}, {48473693}, and so on. Although this initial population is normally generated randomly, chromosomes can be filtered out, representing undesirable pairs of values. On the other hand, values can be forced to be included in the chromosomes. This initial population is then ranked by fitness according to the function f(x, y). Fit chromosomes, which in this case means low values of the objective function, might be selected for reproduction by copying to the next generation according to, for example, the roulette wheel approach. Fit parents are also subject to crossover that attempts to combine portions of good individuals to possibly create even better individuals. Crossover sites along the chromosomes can be chosen in different ways. For example, the two chromosomes $\{27|0663|72\}$ and $\{48|4736|93\}$ can be subject to crossover to give {27|4736|72} and {48|0663|93}. The offspring of these two parents are decoded, evaluated and possibly allowed into the next generation. Mutation alters one or more values in the chromosome from its initial state and helps to prevent the population from stagnating at a local optimum. For example, mutation of the chromosome $\{270|6|6372\}$ might be to $\{270|9|6372\}$. After mutation, chromosomes are decoded, evaluated and possibly allowed into the next generation. The genetic algorithm searches over a specified number of generations to improve the population.

The major strengths of stochastic search methods are that they can tackle the most difficult optimization problems with a high probability of locating solutions in the region of the global optimal. Another major advantage is that, if the solution space is highly irregular, they can produce a range of solutions with close to optimal performance, rather than a single optimal point. This opens up a range of solutions to the designer, rather than having just one option. However, there are also significant disadvantages with stochastic search optimization methods. They can be very slow in converging. The various operations in the stochastic search methods require parameters to be set. The most appropriate values for these parameters vary between different classes of problem and usually need to be adapted to solve particular problems. This means that the methods need tailoring to suit different applications.

3.4 Constrained Optimization

Most optimization problems involve constraints. For example, it might be necessary for a maximum temperature or maximum flowrate not to be exceeded. Thus, the general form of an optimization problem involves three basic elements:

- An objective function to be optimized (e.g. minimize total cost, maximize economic potential, etc.).
- 2) Equality constraints, which are equations describing the model of the process or equipment.
- Inequality constraints, expressing minimum or maximum limits on various parameters.

These three elements of the general optimization problem can be expressed mathematically as:

minimize
$$f(x_1, x_2, ..., x_n)$$

subject to $h_i(x_1, x_2, ..., x_n) = 0$ $(i = 1, p)$
 $g_j(x_1, x_2, ..., x_n) \le 0$ $(j = 1, q)$
(3.11)

In this case, there are n design variables, with p equality constraints and q inequality constraints. The existence of such constraints can simplify the optimization problem by reducing the size of the problem to be searched or avoiding problematic regions of the objective function. In general, though, the existence of the constraints complicates the problem relative to the problem with no constraints.

Now consider the influence of the inequality constraints on the optimization problem. The effect of inequality constraints is to reduce the size of the solution space that must be searched. However, the way in which the constraints bound the feasible region is important. Figure 3.11 illustrates the concept of convex and nonconvex regions. Figure 3.11a shows a convex region. In a convex region, a straight line can be drawn between any two points A and B located within the feasible region and all points on this straight line will also be located within the feasible region. By contrast, Figure 3.11b shows a nonconvex region. This time, when a straight line is drawn between two points A and B located within the region, some of the points on the straight line can fall outside the feasible region. Figure 3.11c shows a region that is constrained by a set of linear inequality constraints. The region shown in Figure 3.10c is convex, but it is worth noting that a set of linear inequality constraints will always provide a convex region (Edgar, Himmelblau and Lasdon, 2001). These concepts can be represented mathematically for the general problem (Floudas, 1995; Biegler, Grossmann and Westerberg, 1997; Edgar, Himmelblau and Lasdon, 2001).

Now superimpose the constraints on to the objective function. Figure 3.12a shows a contour diagram that has a set of inequality constraints imposed. The feasible region is convex and an appropriate search algorithm should be able to locate the unconstrained optimum. The unconstrained optimum lies inside the feasible region in Figure 3.12a. At the optimum none of the constraints are active. By contrast, consider Figure 3.12b. In this case, the region is also convex, but the unconstrained optimum lies outside the feasible region. This time, the optimum point is on the edge of the feasible region and one of the constraints is active. When the inequality constraint is satisfied, as in Figure 3.12b, it





(b) A non-convex region



(c) A set of linear constraints always provides a convex region.



(a) Unconstrained optimum can be reached.



(b) Unconstrained maximum cannot be reached.



(c) A non-convex region might prevent the global optimum from being reached

Figure 3.12

Effects of constraint and optimization

Figure 3.11 Convex and nonconvex regions.
becomes an equality constraint. Figure 3.12c illustrates the potential problem of having a nonconvex region. If the search is initiated in the right-hand part of the diagram at high values of x_1 , then it is likely that the search will find the global optimum. However, if the search is initiated to the left of the diagram at low values of x_1 , it is likely that the search will locate the local optimum (Edgar, Himmelblau and Lasdon, 2001).

It should be noted that it is not sufficient to simply have a convex region in order to ensure that a search can locate the global optimum. The objective function must also be convex if it is to be minimized or concave if it is to be maximized.

The stochastic search optimization methods described previously are readily adapted to the inclusion of constraints. For example, in simulated annealing, if a move suggested at random takes the solution outside the feasible region, then the algorithm can be constrained to prevent this by simply setting the probability of that move to zero. Other methods in stochastic search optimization introduce penalties into the objective function if the constraints are violated.

3.5 Linear Programming

An important class of the constrained optimization problems is one in which the objective function, equality constraints and inequality constraints are all linear. A linear function is one in which the dependent variables appear only to the first power. For example, a linear function of two variables x_1 and x_2 would be of the general form:

$$f(x_1, x_2) = a_0 + a_1 x_1 + a_2 x_2 \tag{3.12}$$

where a_0 , a_1 and a_2 are constants. Search methods for such problems are well developed in *linear programming* (LP). Solving such linear optimization problems is best explained through a simple example.

Example 3.1 A company manufactures two products (Product 1 and Product 2) in a batch plant involving two steps (Step I and Step II). The value of Product 1 is 3-kg^{-1} and that of Product 2 is 2-kg^{-1} . Each batch has the same capacity of 1000 kg per batch but batch cycle times differ between products. These are given in Table 3.1.

Step I has a maximum operating time of $5000 \text{ h} \cdot \text{y}^{-1}$ and Step II $6000 \text{ h} \cdot \text{y}^{-1}$. Determine the operation of the plant to obtain the maximum annual revenue.

Table 3.1

Times for different steps in the batch process.

	Step I (h)	Step II (h)
Product 1	25	10
Product 2	10	20

Solution For Step I, the maximum operating time dictates that:

$$25n_1 + 10n_2 \le 5000$$

where n_1 and n_2 are the number of batches per year manufacturing Products 1 and 2 on Step I. For Step II, the corresponding equation is:

$$10n_1 + 20n_2 \le 6000$$

The feasible solution space can be represented graphically by plotting the above inequality constraints as equality constraints:

$$25n_1 + 10n_2 = 5000$$
$$10n_1 + 20n_2 = 6000$$

This is shown in Figure 3.13. The feasible solution space in Figure 3.13 is given by ABCD.



Figure 3.13

Graphical representation of the linear optimization problem from Example 3.1.

The total annual revenue A is given by:

$$A = 3000n_1 + 2000n_2$$

On a plot of n_1 versus n_2 as shown in Figure 3.13, lines of constant annual revenue will follow a straight line given by:

$$n_2 = -\frac{3}{2}n_1 + \frac{A}{2000}$$

Lines of constant annual revenue are shown as dotted lines in Figure 3.13, with revenue increasing with increasing distance from the origin. It is clear from Figure 3.13 that the optimum point corresponds with the extreme point at the intersection of the two equality constraints at Point C.

A

At the intersection of the two constraints:

$$n_1 = 100$$

 $n_2 = 250$

As the problem involves discrete batches, it is apparently fortunate that the answer turns out to be two whole numbers. However, had the answer turned out not to be a whole number, then the solution would still have been valid because, even though part batches might not be able to be processed, the remaining part of a batch can be processed the following year.

Thus the maximum annual revenue is given by:

$$= 3000 \times 100 + 2000 \times 250 = \$800,000 \$ \cdot y^{-1}$$

Whilst Example 3.1 is an extremely simple example, it illustrates a number of important points. If the optimization problem is completely linear, the solution space is convex and a global optimum solution can be generated. The optimum always occurs at an extreme point, as is illustrated in Figure 3.13. The optimum cannot occur inside the feasible region; it must always be at the boundary. For linear functions, running up the gradient can always increase the objective function until a boundary wall is hit.

Whilst simple two-variable problems like the one in Example 3.1 can be solved graphically, more complex problems require a more formal nongraphical approach. This is illustrated by returning to Example 3.1 to solve it in a nongraphical way.

Example 3.2 Solve the problem in Example 3.1 using an analytical approach.

Solution The problem in Example 3.1 was expressed as:

$$A = 3000n_1 + 2000n_2$$
$$25n_1 + 10n_2 \le 5000$$
$$10n_1 + 20n_2 \le 6000$$

To solve these equations algebraically, the inequality signs must first be removed by introducing *slack variables* S_1 and S_2 such that:

$$25n_1 + 10n_2 + S_1 = 5000$$
$$10n_1 + 20n_2 + S_2 = 6000$$

In other words, these equations show that if the production of both products does not absorb the full capacities of both steps, then the slack capacities of these two processes can be represented by the variables S_1 and S_2 . Since slack capacity means that a certain amount of process capacity remains unused, it follows that the economic value of slack capacity is zero. Realizing that negative production rates and negative slack variables are infeasible, the problem can be formulated as:

$$3000n_1 + 2000n_2 + 0S_1 + 0S_2 = A \tag{3.13}$$

$$25n_1 + 10n_2 + 1S_1 + 0S_2 = 5000 \tag{3.14}$$

$$10n_1 + 20n_2 + 0S_1 + 1S_2 = 6000 \tag{3.15}$$

where $n_1, n_2, S_1, S_2 \ge 0$

Equations 3.14 and 3.15 involve four variables and can therefore not be solved simultaneously. At this stage, the solution can lie anywhere within the feasible area marked *ABCD* in Figure 3.13. However, providing the values of these variables is not restricted to integer values; two of the four variables will assume zero values at the optimum. In this example, n_1 , n_2 , S_1 and S_2 are treated as real and not integer variables.

The problem is started with an initial feasible solution that is then improved by a stepwise procedure. The search will be started at the worst possible solution when n_1 and n_2 are both zero. From Equations 3.14 and 3.15:

$$S_1 = 5000 - 25n_1 - 10n_2 \tag{3.16}$$

$$S_2 = 6000 - 10n_1 - 20n_2 \tag{3.17}$$

When n_1 and n_2 are zero:

$$S_1 = 5000$$

 $S_2 = 6000$

Substituting in Equation 3.13:

$$A = 3000 \times 0 + 2000 \times 0 + 0 \times 5000 + 0 \times 6000$$

= 0 (3.18)

This is Point *A* in Figure 3.13. To improve this initial solution, the value of n_1 and/or the value of n_2 must be increased, because these are the only variables that possess positive coefficients to increase the annual revenue in Equation 3.13. However, which variable, n_1 or n_2 , should be increased first? The obvious strategy is to increase the variable that makes the greatest increase in the annual revenue, which is n_1 . According to Equation 3.17, n_1 can be increased by 6000/10 = 600 before S_2 becomes negative. If n_1 is assumed to be 600 in Equation 3.16, then S_1 would be negative. Since negative slack variables are infeasible, Equation 3.16 is the dominant constraint on n_1 and it follows that its maximum value is 5000/25 = 200. Rearranging Equation 3.16:

$$n_1 = 200 - 0.4n_2 - 0.04S_1 \tag{3.19}$$

which would give a maximum when n_1 and S_1 are zero. Substituting the expression for n_1 in the objective function, Equation 3.13 gives:

$$A = 600,000 + 800n_2 - 120S_1 \tag{3.20}$$

Since n_2 is initially zero, the greatest improvement in the objective function results from making S_1 zero. This is equivalent to making n_1 equal to 200 from Equation 3.19, given that n_2 is initially zero. For $n_1 = 200$ and $n_2 = 0$, the annual revenue A = 600,000. This corresponds with Point *B* in Figure 3.13. However, Equation 3.20 also shows that the profit can be improved further by increasing the value of n_2 . Substituting n_1 from Equation 3.19 in Equation 3.17 gives:

$$n_2 = 250 + 0.025S_1 - 0.0625S_2 \tag{3.21}$$

This means that n_2 takes a value of 250 if both S_1 and S_2 are zero. Substituting n_2 from Equation 3.21 in Equation 3.19 gives:

$$n_1 = 100 - 0.05S_1 + 0.025S_2 \tag{3.22}$$

This means that n_1 takes a value of 100 if both S_1 and S_2 are zero. Finally, substituting the expression for n_1 and n_2 in the objective function, Equation 3.13 gives:

$$A = 800,000 - 100S_1 - 50S_2 \tag{3.23}$$

Equations 3.21 to 3.23 show that the maximum annual revenue is $800,000 \text{ y}^{-1}$ when $n_1 = 100$ and $n_2 = 250$. This corresponds with Point *C* in Figure 3.13.

It is also interesting to note that Equation 3.23 provides some insight into the sensitivity of the solution. The annual revenue would decrease by \$100 for each hour of production lost through poor utilization of Step I. The corresponding effect for Step II would be a reduction of \$50 for each hour of production lost. These values are known as *shadow prices*. If S_1 and S_2 are set to their availabilities of 5000 and 6000 hours respectively, then the revenue from Equation 3.23 becomes zero.

While the method used for the solution of Example 3.1 is not suitable for automation, it gives some insights into the way linear programming problems can be automated. The solution is started by turning the inequality constraints into equality constraints by the use of slack variables. Then the equations are solved to obtain an initial feasible solution. This is improved in steps by searching the extreme points of the solution space. It is not necessary to explore all the extreme points in order to identify the optimum. The method usually used to automate the solution of such linear programming problems is the simplex algorithm (Biegler, Grossmann and Westerberg, 1997; Edgar, Himmelblau and Lasdon, 2001). Note, however, that the simplex algorithm for linear programming should not be confused with the simplex search described previously, which is quite different. Here the term simplex is used to describe the shape of the solution space, which is a convex polyhedron, or simplex.

If the linear programming problem is not formulated properly, it might not have a unique solution, or even any solution at all. Such linear programming problems are termed *degenerate* (Edgar, Himmelblau and Lasdon, 2001). Figure 3.14 illustrates some degenerate linear programming problems (Edgar, Himmelblau and Lasdon, 2001). In Figure 3.14a, the objective function contours are parallel with one of the boundary constraints. Here there is no unique solution that maximizes the objective function within the feasible region. Figure 3.14b shows a problem in which the feasible region is unbounded. Hence the objective function can increase without bound. A third example is shown in Figure 13.14c, in which there is no feasible region according to the specified constraints.

3.6 Nonlinear Programming

When the objective function, equality or inequality constraints of Equation 3.11 are nonlinear, the optimization becomes a nonlinear programming (NLP) problem. The worst case is when all three are nonlinear. Direct and indirect methods that can be used for nonlinear optimization have previously been discussed. Whilst it is possible to include some types of constraints, the methods discussed are not well suited to the inclusion of complex sets of constraints. The stochastic search methods discussed previously can readily handle constraints by restricting moves to infeasible solutions, for example in simulated annealing by setting their probability to 0. The other methods discussed are not well suited to the inclusion of complex sets of constraints. It has already been observed in Figure 3.12 that, unlike the linear optimization problem, for the nonlinear optimization problem the optimum may or may not lie on the edge of the feasible region and can, in principle, be anywhere within the feasible region.

One approach that has been adopted for solving the general nonlinear programming problem is *successive linear programming*. These methods linearize the problem and successively apply the linear programming techniques described in the previous section. The procedures involve initializing the problem and linearizing the objective function and all of the constraints about



Figure 3.14 Degenerate linear programming problems.

the initial point, so as to fit the linear programming format. Linear programming is then applied to solve the problem. An improved solution is obtained and the procedure repeated. At each successive improved feasible solution, the objective function and constraints are linearized and the linear programming solution repeated, until the objective function does not show any significant improvement. If the solution to the linear programming problem moves to an infeasible point, then the nearest feasible point is located and the procedure applied at this new point.

Another method for solving nonlinear programming problems is based on *quadratic programming* (QP) (Edgar, Himmelblau and Lasdon, 2001). Quadratic programming is an optimization procedure that minimizes a quadratic objective function subject to linear inequality or equality (or both types of) constraints. For example, a quadratic function of two variables x_1 and x_2 would be of the general form:

$$f(x_1, x_2) = a_0 + a_1 x_1 + a_2 x_2 + a_{11} x_1^2 + a_{22} x_2^2 + a_{12} x_1 x_2$$
(3.24)

where a_{ii} are constants. Quadratic programming problems are the simplest form of nonlinear programming with inequality constraints. The techniques used for the solution of quadratic programming problems have many similarities with those used for solving linear programming problems (Edgar, Himmelblau and Lasdon, 2001). Each inequality constraint must either be satisfied as an equality or it is not involved in the solution of the problem. The quadratic programming technique can thus be reduced to a vertex searching procedure, similar to linear programming (Edgar, Himmelblau and Lasdon, 2001). In order to solve the general nonlinear programming problem, quadratic programming can be applied successively, in a similar way to that for successive linear programming, in successive (or sequential) quadratic programming (SQP). In this case, the objective function is approximated locally as a quadratic function. For a function of two variables, the function would be approximated by Equation 3.24. By approximating the function as a quadratic and linearizing the constraints, this takes the form of a quadratic programming problem that is solved in each iteration (Edgar, Himmelblau and Lasdon, 2001). In general, successive quadratic programming tends to perform better than successive linear programming, because a quadratic rather than a linear approximation is used for the objective function.

It is important to note that neither successive linear nor successive quadratic programming are guaranteed to find the global optimum in a general nonlinear programming problem. The fact that the problem is being turned into a linear or quadratic problem, for which global optimality can be guaranteed, does not change the underlying problem that is being optimized. All of the problems associated with local optima are still a feature of the background problem. When using these methods for the general nonlinear programming problem, it is important to recognize this and to test the optimality of the solution by starting the optimization from different initial conditions.

Stochastic search optimization methods described previously, such as simulated annealing, can also be used to solve the general nonlinear programming problem. These have the advantage that the search is sometimes allowed to move uphill in a minimization problem, rather than always searching for a downhill move. Alternatively, in a maximization problem, the search is sometimes allowed to move downhill, rather than always searching for an uphill move. In this way, the technique is less vulnerable to the problems associated with local optima.

3.7 Structural Optimization

1) Mixed integer linear programming. In Chapter 1, different ways were discussed that can be used to develop the structure of a flowsheet. In the first way, an irreducible structure is built by successively adding new features if these can be justified technically and economically. The second way to develop the structure of a flowsheet is to first create a superstructure. This superstructure involves redundant features but includes the structural options that should be considered. This superstructure is then subjected to optimization. The optimization varies the settings of the process parameters (e.g. temperature, flowrate) and also optimizes the structural features. Thus, to adopt this approach, both structural and parameter optimization must be carried out. So far the discussion of optimization has been restricted to parameter optimization. Consider now how structural optimization can be carried out.

The methods discussed for linear and nonlinear programming can be adapted to deal with structural optimization by introducing integer (binary) variables that identify whether a feature exists or not. If a feature exists, its binary variable takes the value 1. If the feature does not exist, then it is set to 0. Consider how different kinds of decisions can be formulated using binary variables (Biegler, Grossman and Westerberg, 1997).

a) Multiple choice constraints. It might be required to select only one item from a number of options. This can be represented mathematically by a constraint:

$$\sum_{j=1}^{J} y_j = 1 \tag{3.25}$$

where y_j is the binary variable to be set to 0 or 1 and the number of options is *J*. More generally, it might be required to select only *m* items from a number of options. This can be represented by:

$$\sum_{j=1}^{J} y_j = m \tag{3.26}$$

Alternatively, it might be required to select at most *m* items from a number of options, in which case the constraint can be represented by:

$$\sum_{j=1}^{J} y_j \le m \tag{3.27}$$

On the other hand, it might be required to select at least m items from a number of options. The constraint can be represented by:

$$\sum_{j=1}^{J} y_j \ge m \tag{3.28}$$

b) Implication constraints. Another type of logical constraint might be that if Item k is selected, Item j must be selected, but not vice versa, then this is represented by the constraint:

$$y_k - y_j \le 0 \tag{3.29}$$

A binary variable can be used to set a continuous variable to 0. If a binary variable y is 0, the associated continuous variable x must also be 0 if a constraint is applied such that:

$$x - Uy \le 0, x \ge 0 \tag{3.30}$$

where U is an upper limit to x.

c) Either–or constraints. Binary variables can also be applied to either–or constraints, known as *disjunctive constraints*. For example, either constraint $g_1(x) \le 0$ or constraint $g_2(x) \le 0$ must hold:

$$g_1(x) - M \, y \le 0 \tag{3.31}$$

$$g_2(x) - M(1 - y) \le 0 \tag{3.32}$$

where *M* is a large (arbitrary) value that represents an upper limit $g_1(x)$ and $g_2(x)$. If y=0, then $g_1(x) \le 0$ must be imposed from Equation 3.31. However, if y=0, the lefthand side of Equation 3.32 becomes a large negative number whatever the value of $g_2(x)$ and, as a result, Equation 3.32 is always satisfied. If y=1, then the lefthand side of Equation 3.31 is a large negative number whatever the value of $g_1(x)$ and Equation 3.31 is always satisfied. However, now $g_2(x) \le 0$ must be imposed from Equation 3.32.

Some simple examples can be used to illustrate the application of these principles.

Example 3.3 A gaseous waste stream from a process contains valuable hydrogen that can be recovered by separating the hydrogen from the impurities using pressure swing adsorption (PSA), a membrane separator (MS) or a cryogenic condensation (CC). The pressure swing adsorption and membrane separator can in principle be used either individually or in combination. Write a set of integer equations that would allow one from the three options of pressure swing adsorption, membrane separator or cryogenic condensation to be chosen, but also allow the pressure swing adsorption and membrane separator to be chosen in combination.

Solution Let y_{PSA} represent the selection of pressure swing adsorption, y_{MS} the selection of the membrane separator and y_{CC}

the selection of cryogenic condensation. First restrict the choice between pressure swing adsorption and cryogenic condensation:

$$y_{PSA} + y_{CC} < 1$$

Now restrict the choice between membrane separator and cryogenic condensation:

$$y_{MS} + y_{CC} < 1$$

These two equations restrict the choices, but still allow the pressure swing adsorption and membrane separator to be chosen together.

Example 3.4 The temperature difference in a heat exchanger between the inlet temperature of the hot stream $T_{H, in}$ and the outlet of the cold stream $T_{C, out}$ is to be restricted to be greater than a practical minimum value of ΔT_{min} , but only if the option of having the heat exchanger is chosen. Write a disjunctive constraint in the form of an integer equation to represent this constraint.

Solution The temperature approach constraint can be written as:

$$T_{H,in} - T_{C,out} \ge \Delta T_{min}$$

However, this should apply only if the heat exchanger is selected. Let y_{HX} represent the option of choosing the heat exchanger:

$$T_{H,in} - T_{C,out} + M(1 - y_{HX}) \ge \Delta T_{\min}$$

where *M* is an arbitrary large number. If $y_{HX}=0$ (i.e. the heat exchanger is not chosen), then the left-hand side of this equation is bound to be greater than ΔT_{min} no matter what the values of $T_{H, in}$ and $T_{C, out}$ are. If $y_{HX}=1$ (i.e. the heat exchanger is chosen), then the equation becomes $(T_{H, in} - T_{C, out}) \ge \Delta T_{min}$ and the constraint must apply.

When a linear programming problem is extended to include integer (binary) variables, it becomes a *mixed integer linear programming problem* (MILP). Correspondingly, when a nonlinear programming problem is extended to include integer (binary) variables, it becomes a *mixed integer nonlinear programming problem* (MINLP).

First consider the general strategy for solving an MILP problem. Initially, the binary variables can be treated as continuous variables, such that $0 \le y_i \le 1$. The problem can then be solved as an LP. The solution is known as a *relaxed solution*. The most likely outcome is that some of the binary variables will exhibit noninteger values at the optimum LP solution. Because the relaxed solution is less constrained than the true mixed integer values, it will in general give a better value for the objective function than the true mixed integer solution. In general, the noninteger values of the binary variables cannot simply be rounded to the nearest integer value, either because the region) or because the rounding may render the solution nonoptimal (not at the edge of the feasible





Setting the binary variables to zero or one creates a tree structure. (Reproduced from Floudas CA, 1995, Nonlinear and Mixed-Integer Optimization, by permission of Oxford University Press.)

region). However, this relaxed LP solution is useful in providing a lower bound to the true mixed integer solution to a minimization problem. For maximization problems, the relaxed LP solutions form the upper bound to the solution. The noninteger values can then be set to either 0 or 1 and the LP solution repeated. The setting of the binary variables to be either 0 or 1 creates a solution space in the form of a tree, as shown in Figure 3.15 (Floudas, 1995). As the solution is stepped through, the number of possibilities increases by virtue of the fact that each binary variable can take a value of 0 or 1 (Figure 3.15). At each point in the search, the best relaxed LP solution provides a lower bound to the optimum of a minimization problem. Correspondingly, the best true mixed integer solution provides an upper bound. For maximization problems the best relaxed LP solution forms an upper bound to the optimum and the best true mixed integer solution provides the lower bound. A popular method of solving MILP problems is to use a branch and bound search (Mehta and Kokossis, 1988). This will be illustrated by a simple example from Edgar, Himmelblau and Lasdon (2001).

Example 3.5 A problem involving three binary variables y_1 , y_2 and y_3 has an objective function to be maximized (Edgar, Himmelblau and Lasdon, 2001):

maximize :
$$f = 86y_1 + 4y_2 + 40y_3$$

subject to : $774y_1 + 76y_2 + 42y_3 \le 875$
 $67y_1 + 27y_2 + 53y_3 \le 875$
 $y_1, y_2, y_3 = 0, 1$
(3.33)

Solution The solution strategy is illustrated in Figure 3.16a. First the LP problem is solved to obtain the relaxed solution, allowing y_1 , y_2 and y_3 to vary continuously between 0 and 1. Both $y_1 = 1$ and $y_3 = 1$ at the optimum of the relaxed solution but the value of y_2 is

0.776, Node 1 in Figure 3.16a. The objective function for this relaxed solution at Node 1 is f = 129.1. From this point, y_2 can be set to be either 0 or 1. Various strategies can be adopted to decide which one to choose. A very simple strategy will be adopted here of picking the closest integer to the real number. Given that $y_2 = 0.776$ is closer to 1 than 0, set $y_2 = 1$ and solve the LP at Node 2 in Figure 3.16a. Now $y_2 = 1$ and $y_3 = 1$ at the optimum of the relaxed solution but the value of y_1 is 0.978, Node 2 in Figure 3.16a. Given that $y_1 = 0.978$ is closer to 1 than 0, set $y_1 = 1$ and solve the LP at Node 3. This time y_1 and y_2 are integers, but $y_3 = 0.595$ is a noninteger. Setting $y_3 = 1$ yields an infeasible solution at Node 4 in Figure 3.16a as it violates the first inequality constraint in Equation 3.33. Backtracking to Node 3 and setting $y_3 = 0$ yields the first feasible integer solution at Node 5 for which $y_1 = 1$, $y_2 = 1$, $y_3 = 0$ and f = 90.0. There is no point in searching further from Node 4 as it is an infeasible solution or from Node 5 as it is a valid integer solution. When the search is terminated at a node for either reason, it is deemed to be *fathomed*. The search now backtracks to Node 2 and sets $y_1 = 0$. This yields the second feasible integer solution at Node 6 for which $y_1 = 0$, $y_2 = 1$, $y_3 = 1$ and f = 44.0. Finally, backtrack to Node 1 and set $y_2 = 0$. This yields the third feasible integer solution at Node 7 for which $y_1 = 1$, $y_2 = 0$, $y_3 = 1$ and f = 126.0. Since the objective function is being maximized, Node 7 is the optimum for the problem.

Searching the tree in the way done in Figure 3.16a is known as a depth first or backtracking approach. At each node, the branch was followed that appeared to be more promising to solve. Rather than using a depth first approach, a breadth first or jumptracking approach can be used, as illustrated in Figure 3.16b. Again start at Node 1 and solve the relaxed problem in Figure 3.16b. This gives an upper bound for the maximization problem of f = 129.1. However, this time the search goes across the tree with the initial setting of $y_2 = 0$. This yields a valid integer solution at Node 2 with $y_1 = 1$, $y_2 = 0$, $y_3 = 1$ and f = 126.0. Node 2 now forms a lower bound and is fathomed because it is an integer solution. In this approach, the search now backtracks to Node 1 whether Node 2 is fathomed or not. From Node 1 now set $y_2 = 1$. The solution at Node 3 in Figure 3.16b gives $y_1 = 0.978$, $y_2 = 1$, $y_3 = 1$ and f = 128.1, which is the new upper bound. Setting $y_1 = 0$ and branching to Node 4 gives the second valid integer solution. Now backtrack to Node 3 and set $y_1 = 1$. The solution at Node 5 has $y_1 = 1$, $y_2 = 1$, $y_3 = 0.595$ and f = 113.8. At this point, Node 5 is fathomed, even though it is neither infeasible nor a valid integer solution. The upper bound of this branch at Node 5 has a value of the objective function lower than that of the integer solution at Node 2. Setting the values to be integers from Node 5 can only result in an inferior solution. In this way, the search is bounded.

In this case, the breadth first search yields the optimum with a fewer number of nodes to be searched. Different search strategies than the ones used here can readily be used (Taha, 1975). It is likely that different problems would be suited to different search strategies.

Thus, the solution of the MILP problem is started by solving the first relaxed LP problem. If integer values are obtained for the binary variables, the problem has been solved. However, if integer values are not obtained, the use of bounds is examined to avoid parts of the tree that are known to be suboptimal. The node with the best noninteger solution provides a lower bound for minimization problems and the node with the best feasible



mixed integer solution provides an upper bound. In the case of maximization problems, the node with the best noninteger solution provides an upper bound and the node with the best feasible mixed integer solution provides a lower bound. Nodes with noninteger solutions are fathomed when the value of the objective function is inferior to the best integer solution (the lower bound). The tree can be searched by following a depth first approach or a breadth first approach, or a combination of the two. Given a more complex problem than Example 3.5, the search could, for example, set the values of the noninteger variables to be 0 and 1 in turn and carry out an evaluation of the objective function (rather than an optimization). This would then indicate the best direction in which to go for the next optimization. Many strategies are possible (Taha, 1975).

The series of LP solutions required for MILP problems can be solved efficiently by using one LP to initialize the next. An important point to note is that, in principle, a global optimum solution can be guaranteed in the same way as with LP problems.

2) *Mixed Integer Nonlinear Programming.* The general strategy for solving mixed integer nonlinear programming problems is very similar to that for linear problems (Floudas, 1995). The major difference is that each node requires the solution of a nonlinear program, rather than the solution of a linear program. Unfortunately, searching the tree with a succession of nonlinear optimizations can be extremely expensive in terms of the computation time required, as information cannot be readily carried from one NLP to the next, as can be done for LP. Another major problem is that, because a series of nonlinear optimizations is being carried out, there is no guarantee that the optimum will even be close to the global optimum, unless the NLP problem being solved at each node

is convex. Of course, different initial points can be tried to overcome this problem, but there can still be no guarantee of global optimality for the general problem.

Another way to deal with such nonlinear problems is to first approximate the solution to be linear and apply MILP, and then apply NLP to the problem. The method then iterates between MILP and NLP (Biegler, Grossmann and Westerberg, 1997).

In some cases, the nonlinearity in a problem can be isolated in a small number of the functions. If this is the case, then one simple way to solve the problem is to linearize the nonlinear function by a series of straight-line segments. Integer logic can then be used to ensure that only one of the straight-line segments is chosen at a time and MILP used to carry out the optimization. For some forms of nonlinear mathematical expressions, deterministic optimization methods can be tailored to find the global optimum through the application of mathematical transformations and bounding techniques (Floudas, 2000).

3) Stochastic Search Optimization. Stochastic search optimization can be extremely effective for structural optimization if the optimization is nonlinear in character. Consider again the approach to process design introduced in Chapter 1 in which an initial design is developed and then evolved through structural and parameter optimization. This evolutionary approach is suited to stochastic search optimization. This does not require either a superstructure or a search through a tree as branch and bound methods require. For example, when using simulated annealing, at each setting of the annealing temperature, a series of random moves is performed. These moves can be either step changes to continuous variables or can be changes in structure (either the addition or removal of a structural feature). Thus, the

approach does not require a superstructure to be created, as the approach can add features as well as remove features. However, structural moves must be defined and these must ensure that those moves can somehow create all the structural options that might be candidates for the global optimum structure. In theory, the initialization of the structure does not matter and any initial feasible structure will do to start the optimization, as long as the stochastic search optimization parameters have been set to appropriate values. In practice, it tends to be better for most problems to create an initial design with at least some redundant features, rather like a superstructure, but not necessarily including all possible structural options.

Because stochastic search optimization allows some deterioration in the objective function, it is not as prone to being trapped by a local optimum as MINLP. However, when optimizing a problem involving both continuous variables (e.g. temperature and pressure) and structural changes, stochastic search optimization algorithms take finite steps for the continuous variables and do not necessarily find the exact value for the optimum setting. Because of this, a deterministic method (e.g. SQP) can be applied after stochastic search optimization to fine-tune the answer. This uses the stochastic search method to provide a good initialization for the deterministic method.

Also, it is possible to combine stochastic and deterministic methods as *hybrid* methods. For example, a stochastic search method can be used to control the structural changes and a deterministic method to control the changes in the continuous variables. This can be useful if the problem involves a large number of integer variables, because, for such problems, the tree required for branch and bound methods explodes in size.

3.8 Solution of Equations Using Optimization

It is sometimes convenient to use optimization to solve equations, or sets of simultaneous equations. This arises from the availability of general-purpose optimization software, such as that available in spreadsheets. *Root finding* or *equation solving* is a special case for optimization, where the objective is to reach a value of 0. For example, suppose it is necessary to solve a function f(x) for the value x that satisfies f(x) = a, where a is a constant. As illustrated in Figure 3.17a, this can be solved for:

$$f(x) - a = 0 \tag{3.34}$$

The objective of the optimization would be for the equality constraint given by Equation 3.34 to be satisfied as nearly as possible. There are a number of ways in which this objective can be made specific for optimization. Three possibilities are (Williams, 1997):

(i) minimize |f(x) - a| (3.35)

(ii) minimize
$$[f(x) - a]^2$$
 (3.36)

(iii) minimize
$$(S_1 + S_2)$$
 (3.37)

subject to

$$f(x) - a + S_1 - S_2 = 0 \tag{3.38}$$

$$S_1, S_2 \ge 0 \tag{3.39}$$

where S_1 and S_2 are slack variables

Which of these objectives would be the best to use depends on the nature of the problem, the optimization algorithm being used and the initial point for the solution. For example, minimizing Objective (i) in Equation 3.35 can present problems to optimization methods as a result of the gradient being discontinuous, as illustrated in Figure 3.17b. However, the problem can be transformed such that the objective function for the optimization has no discontinuities in the gradient. One possible transformation is Objective (ii) in Equation 3.36. As illustrated in Figure 3.17c,



(a) An equation to be solved.



(b) Transforming the problem into an optimization problem.



(c) Transformation of the objective function to make the gradient continuous.

Figure 3.17

Solving equations using optimization.

this now has a continuous gradient. However, a fundamental disadvantage in using the transformation in Equation 3.36 is that if the equations to be solved are linear, then the objective function is transformed from linear to nonlinear. Also, if the equations to be solved are nonlinear, then such transformations will increase the nonlinearity. On the other hand, Objective (iii) in Equation 3.37 avoids both an increase in the nonlinearity and discontinuities in the gradient, but at the expense of introducing slack variables. Note that two slack variables are needed. Slack variable S_1 in Equation 3.38 for $S_1 \ge 0$ ensures that:

$$f(x) - a \le 0 \tag{3.40}$$

whereas slack variable S_2 in Equation 3.38 for $S_2 \ge 0$ ensures that:

$$f(x) - a \ge 0 \tag{3.41}$$

Equations 3.40 and 3.41 are only satisfied simultaneously by Equation 3.34. Thus, x, S_1 and S_2 can be varied simultaneously to solve Equations 3.37 to 3.39 without increasing the nonlinearity.

The approach can be extended to solve sets of simultaneous equations. For example, suppose a solution is required for x_1 and x_2 such that:

$$f_1(x_1, x_2) = a_1$$
 and $f_2(x_1, x_2) = a_2$ (3.42)

Three possible ways to formulate the objective for optimization are:

(i) minimize {
$$|f_1(x_1, x_2) - a_1| + |f_2(x_1, x_2) - a_2|$$
 } (3.43)

(ii) minimize
$$\left\{ [f_1(x_1, x_2) - a_1]^2 + [f_2(x_1, x_2) - a_2]^2 \right\}$$
 (3.44)

(iii) minimize
$$(S_{11} + S_{12} + S_{21} + S_{22})$$

subject to (3.45)

$$f_1(x_1, x_2) - a_1 + S_{11} - S_{12} = 0 \tag{3.46}$$

$$f_2(x_1, x_2) - a_2 + S_{21} - S_{22} = 0 \tag{3.47}$$

$$S_{11}, S_{12}, S_{21}, S_{22} \ge 0 \tag{3.48}$$

where S_{11} , S_{12} , S_{21} and S_{22} are slack variables

As will be seen later, these techniques will prove to be useful when solving design problems in general-purpose software, such as spreadsheets. Many of the numerical problems associated with optimization can be avoided by appropriate formulation of the model. Further details of model building can be found elsewhere (Williams, 1997).

3.9 The Search for Global Optimality

From the discussion in this chapter, it is clear that the difficulties associated with optimizing nonlinear problems are far greater than those for optimizing linear problems. For linear problems, finding the global optimum can, in principle, be guaranteed. Unfortunately, when optimization is applied to most relatively large design problems, the problem usually involves solving nonlinear optimization. In such situations, standard deterministic optimization methods will find only the first local optimum encountered. Starting from different initializations can allow the optimization to explore different routes through the solution space and might help identify alternative solutions. However, there is no guarantee of finding the global optimum. For some forms of nonlinear mathematical expressions, deterministic optimization methods can be tailored to find the global optimum through the application of mathematical transformations and bounding techniques (Floudas, 2000).

Alternatively, stochastic search optimization (e.g. simulation annealing or genetic algorithms) can be used. These have the advantage of, in principle, being able to locate the global optimum for the most general nonlinear optimization problems. They do not require good initialization and do not require gradients to be defined. However, they involve parameters that are system-dependent and might need to be adjusted for problems that are different in character. Another disadvantage is that they can be extremely slow in solving large complex optimization problems. The relative advantages of deterministic and stochastic search methods can be combined using hybrid methods by using stochastic search methods to provide a good initial point for a deterministic method. As mentioned previously, stochastic and deterministic methods can also be combined to solve structural optimization problems.

In Chapter 1, an objective function for a nonlinear optimization was likened to the terrain in a range of mountains. If the objective function is to be maximized, each peak in the mountain range represents a local optimum in the objective function. The highest peak represents the global optimum. Optimization requires searching around the mountains in a thick fog to find the highest peak, without the benefit of a map and only a compass to tell direction and an altimeter to show height. On reaching the top of any peak, there is no way of knowing whether it is the highest peak because of the fog.

When solving such nonlinear optimization problems, it is not desirable to terminate the search at a peak that is grossly inferior to the highest peak. The solution can be checked by repeating the search but starting from a different initial point.

However, the shape of the optimum for most optimization problems bears a greater resemblance to Table Mountain in South Africa rather than to Mount Everest. In other words, for most optimization problems, the region around the optimum is fairly flat. Although on one hand a grossly inferior solution should be avoided, on the other hand the designer should not be preoccupied with improving the solution by tiny amounts in an attempt to locate exactly the global optimum. There will be uncertainty in the design data, especially economic data. Also, there are many issues to be considered other than simply maximizing economic potential or minimizing cost. There could be many reasons why the solution at the exact location of the global optimum might not be preferred, while a slightly suboptimal solution might be preferred for other reasons, such as safety, ease of control, and so on. Different solutions in the region of the optimum should be examined, rather than considerable effort

being expended on finding the solution at the exact location of the global optimum and considering only that solution. In this respect, stochastic search optimization has advantages, as it can provide a range of solutions in the region of the optimum.

3.10 Optimization – Summary

Most design problems will require optimization to be carried out at some stage. The quality of the design is characterized by an objective function to be maximized (e.g. if economic potential is being maximized) or minimized (e.g. if cost is being minimized). The shape of the objective function is critical in determining the optimization strategy. If the objective function is convex in a minimization problem or concave in a maximization problem, then there is a single optimum point. If this is not the case, there can be local optima as well as the global optimum.

Various search strategies can be used to locate the optimum. Indirect search strategies do not use information on gradients, whereas direct search strategies require this information. These methods always seek to improve the objective function in each step in a search. On the other hand, stochastic search methods, such as simulated annealing and genetic algorithms, allow some deterioration in the objective function, especially during the early stages of the search, in order to reduce the danger of being attracted to a local optimum rather than the global optimum. However, stochastic search optimization can be very slow in converging and usually needs to be adapted to solve particular problems.

The addition of inequality constraints complicates the optimization. These inequality constraints can form convex or nonconvex regions. If the region is nonconvex, this means that the search can be attracted to a local optimum, even if the objective function is convex in the case of a minimization problem or concave in the case of a maximization problem. In the case where a set of inequality constraints is linear, the resulting region is always convex.

The general case of optimization in which the objective function, the equality and inequality constraints are all linear can be solved as a linear programming problem. This can be solved efficiently with, in principle, a guarantee of global optimality. However, the corresponding nonlinear programming problem cannot, in general, be solved efficiently and with a guarantee of global optimality. Such problems are solved by successive linear or successive quadratic programming. Stochastic search optimization methods can be very effective in solving nonlinear optimization, because they are less prone to be stuck in a local optimum than deterministic methods.

One of the approaches that can be used in design is to carry out structural and parameter optimization of a superstructure. The structural optimization required can be carried out using mixed integer linear programming in the case of a linear problem or mixed integer nonlinear programming in the case of a nonlinear problem. Stochastic search optimization can also be very effective for structural optimization problems.

3.11 Exercises

- 1. The cost of closed atmospheric cylindrical storage vessels can be considered to be proportional to the mass of steel required. Derive a simple expression for the dimensions of such a storage tank to give minimum capital cost. Assume the top and bottom are both flat. What are the dimensions for the minimum capital cost if the tank has an open top?
- The overhead of vapor of a distillation column is to be condensed in a heat exchanger using cooling water. There is a trade-off involving the flowrate of cooling water and the size of the condenser. As the flowrate of cooling water increases, its cost increases. However, as the flowrate increases, the return temperature of the cooling water to the cooling tower decreases. This decreases the temperature differences in the condenser and increases its heat transfer area, and hence its capital cost. The condenser has a duty of 4.1 MW and the vapor condenses at a constant temperature of 80 °C. Cooling water is available at 20 °C with a cost of \$ $0.02 t^{-1}$. The overall heat transfer coefficient can be assumed to be $500 \text{ W} \cdot \text{m}^{-1} \cdot \text{K}^{-1}$. The cost of the condenser can be assumed to be $$2500 \text{ m}^{-2}$ with an installation factor of 3.5. Annual capital charges can be assumed to be 20% of the capital costs. The heat capacity of the cooling water can be assumed constant at $4.2 \text{ kJ} \cdot \text{kg}^{-1} \cdot \text{K}^{-1}$. The distillation column operates for $8000 \text{ h} \cdot \text{y}^{-1}$. Set up an equation for the heat transfer area of the condenser, and hence the annual capital cost of the condenser, in terms of the cooling water return temperature. Using this equation, carry out a tradeoff between the cost of the cooling water and the cost of the condenser to determine approximately the optimum cooling water return temperature. The maximum return temperature should be 50 °C. The heat exchange area required by the condenser is given by:

$$A = \frac{Q}{U\,\Delta T_{LM}}$$

where A = heat transfer area (m²)

Q = heat duty (W)

U = overall heat transfer coefficient $(W \cdot m^{-1} \cdot K^{-1})$

 ΔT_{LM} = logarithmic mean temperature difference

$$=\frac{(T_{COND} - T_{CW2}) - (T_{COND} - T_{CW1})}{\ln\left(\frac{T_{COND} - T_{CW2}}{T_{COND} - T_{CW1}}\right)}$$

 T_{COND} = condenser temperature (°C)

 T_{CW1} = inlet cooling water temperature (°C)

 T_{CW2} = outlet cooling water temperature (°C)

3. A vapor stream leaving a styrene production process contains hydrogen, methane, ethylene, benzene, water, toluene, styrene and ethylbenzene and is to be burnt in a furnace. It is proposed to recover as much of the benzene, toluene, styrene and ethylbenzene as possible from the vapor using low-temperature condensation. The low-temperature condensation requires refrigeration. However, the optimum temperature

Table 3.2

Stream flowrates and component values.

Component	Flowrate (kmol·s ⁻¹)	Value (\$·kmol ⁻¹)
Hydrogen	146.0	0
Methane	3.7	0
Ethylene	3.7	0
Benzene	0.67	21.4
Water	9.4	0
Toluene	0.16	12.2
Ethylbenzene	1.6	40.6
Styrene	2.4	60.1
Total	167.63	

for the condensation needs to be determined. This involves a trade-off in which the amount and value of material recovered increases as the temperature decreases, but the cost of the refrigeration increases as the temperature decreases. The material flows in the vapor leaving the flash drum are given in Table 3.2, together with their values.

The low-temperature condensation requires refrigeration, for which the cost is given by:

Refrigeration cost =
$$0.033Q_{COND} \left(\frac{40 - T_{COND}}{T_{COND} + 268} \right)$$

where Q_{COND} = condenser duty (MW) T_{COND} = condenser temperature (°C)

The fraction of benzene, toluene, styrene and ethylbenzene condensed can be determined from phase equilibrium calculations. The percent of the various components entering the condenser that leave with the vapor are given in Table 3.3 as a function of temperature. The total enthalpy of the flash

Table 3.4

Stream enthalpy data.

Temperature (°C)	Stream enthalpy (MJ·kmol ⁻¹)
40	0.45
30	-1.44
20	-2.69
10	-3.56
0	-4.21
-10	-4.72
-20	-5.16
-30	-5.56
-40	-5.93
-50	-6.29
-60	-6.63

drum vapor stream as a function of temperature is given in Table 3.4.

Calculate the optimum condenser temperature. What practical difficulties do you foresee in using very low temperatures?

4. A tank containing 1500 m^3 of naphtha is to be blended with two other hydrocarbon streams to meet the specifications for gasoline. The final product must have a minimum research octane number (RON) of 95, a maximum Reid vapor pressure (RVP) of 0.6 bar, a maximum benzene content of 2% vol and maximum total aromatics of 25% vol. The properties and costs of the three streams are given in the Table 3.5.

Assuming that the properties of the mixture blend are in proportion to the volume of stream used, how much reformate and alkylate should be blended to minimize cost?

5. A petroleum refinery has two crude oil feeds available. The first crude (Crude 1) is high-quality feed and costs \$30 per

	Percent of component entering condenser that leaves with the vapor (%)										
		Condensation temperature (°C)									
	40	30	20	10	0	- 10	- 20	- 30	- 40	- 50	- 60
Benzene	100	93	84	72	58	42	27	15	8	3	1
Toluene	100	80	60	41	25	14	7	3	1	1	0
Ethylbenzene	100	59	33	18	9	4	2	1	0	0	0
Styrene	100	54	29	15	8	4	2	1	0	0	0

Table 3.3

Condenser performance.

Table 3.5

Blending streams.

	RON	RVP (bar)	Benzene (% vol)	Total aromatics (% vol)	Cost \$⋅m ⁻³
Naphtha	92	0.80	1.5	15	275
Reformate	98	0.15	15	50	270
Alkylate	97.5	0.30	0	0	350

Table 3.6

Refinery data.

	Yield (%	volume)	Value of product	Maximum production
	Crude 1	Crude 2	(\$•bbl ⁻¹)	(bbl·day ⁻¹)
Gasoline	80	47	75	120,000
Jet fuel	4	8	55	8,000
Diesel	10	30	40	30,000
Fuel oil	6	15	30	-
Processing cost (\$ bb1 ⁻¹)	1.5	3		

barrel (1 barrel = 42 US gallons). The second crude (Crude 2) is a low-quality feed and costs \$20 per barrel. The crude oil is separated into gasoline, diesel, jet fuel and fuel oil. The percent yield of each of these products that can be obtained from Crude 1 and Crude 2 are listed in Table 3.6, together with maximum allowable production flowrates of the products in barrels per day and processing costs.

The economic potential can be taken to be the difference between the selling price of the products and the cost of the crude oil feedstocks. Determine the optimum feed flowrate of the two crude oils from a linear optimization solved graphically.

- **6.** Add a constraint to the specifications for Exercise 5 above such that the production of fuel oil must be greater than 15,000 bbl·day⁻¹. What happens to the problem? How would you describe the characteristics of the modified linear programming problem?
- 7. Devise a superstructure for a distillation design involving a single feed, two products, a reboiler and a condenser that will allow the number of plates in the column itself to be varied between 3 and 10 and at the same time vary the location of the feed tray.
- **8.** A reaction is required to be carried out between a gas and a liquid. Two different types of reactor are to be considered: an agitated vessel (AV) and a packed column (PC). Devise a superstructure that will allow one of the two options to be

chosen. Then describe this as integer constraints for the gas and liquid feeds and products.

References

- Aarts EHL and Van Laarhoven PGM (1985) Statistical Cooling: A General Approach to Combinatorial Optimization Problems, *Philips J Res*, 40: 193.
- Biegler LT, Grossmann IE and Westerberg AW (1997) Systematic Methods of Chemical Process Design, Prentice Hall.
- Edgar TF, Himmelblau DM and Lasdon LS (2001) Optimization of Chemical Processes, 2nd Edition, McGraw-Hill.
- Floudas CA (1995) Nonlinear and Mixed-Integer Optimization, Oxford University Press.
- Floudas CA (2000) Deterministic Global Optimization: Theory, Methods and Applications, Kluwer Academic Publishers.
- Goldberg DE (1989) Genetic Algorithms in Search Optimization and Machine Learning, Addison-Wesley.
- Kirkpatrick S, Gelatt CD and Vecchi MP (1983) Optimization by Simulated Annealing, *Science*, 220: 671.
- Mehta VL and Kokossis AC (1988) New Generation Tools for Multiphase Reaction Systems: A Validated Systematic Methodology for Novelty and Design Automation, *Comput Chem Eng*, **22S**: 5119.
- Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH and Teller E (1953) Equation of State Calculations by Fast Computing Machines, *J Chem Phys*, **21**: 1087.
- Taha HA (1975) Integer Programming Theory and Applications, Academic Press.
- Williams HP (1997) Model Building in Mathematical Programming, 3rd Edition, John Wiley & Sons.

Chapter 4

Chemical Reactors I – Reactor Performance

S ince process design starts with the reactor, the first decisions are those that lead to the choice of reactor. These decisions are amongst the most important in the whole process design. Good reactor performance is of paramount importance in determining the economic viability of the overall design and fundamentally important to the environmental impact of the process. In addition to the desired products, reactors produce unwanted byproducts. These unwanted byproducts not only lead to a loss of revenue but can also create environmental problems. As will be discussed later, the best solution to environmental problems is not to employ elaborate treatment methods, but to not produce waste in the first place.

Once the product specifications have been fixed, some decisions need to be made regarding the *reaction path*. There are sometimes different paths to the same product. For example, suppose ethanol is to be manufactured. Ethylene could be used as a raw material and reacted with water to produce ethanol. An alternative would be to start with methanol as a raw material and react it with synthesis gas (a mixture of carbon monoxide and hydrogen) to produce the same product. These two paths employ chemical reactor technology. A third path could employ a *biochemical* reaction (or *fermentation*) that exploits the metabolic processes of *microorganisms* in a biochemical reactor. Ethanol could therefore also be manufactured by fermentation of a carbohydrate.

Reactors can be broadly classified as chemical or biochemical. Most reactors, whether chemical or biochemical, are catalyzed. The strategy will be to choose the catalyst, if one is to be used, and the ideal characteristics and operating conditions needed for the reaction system. The issues that must be addressed for reactor design include:

- Reactor type
- Catalyst

- Size (volume)
- Operating conditions (temperature and pressure)
- Phase
- Feed conditions (concentration, temperature and pressure).

Once basic decisions have been made regarding these issues, a practical reactor is selected, approaching as nearly as possible the ideal in order that the design can proceed. However, the reactor design cannot be fixed at this stage since, as will be seen later, it interacts strongly with the rest of the flowsheet. The focus here will be on the choice of reactor and not its detailed sizing. For further details of sizing, see, for example, Levenspiel (1999), Denbigh and Turner (1984) and Rase (1977).

4.1 Reaction Path

As already noted, given that the objective is to manufacture a certain product, there are often a number of alternative reaction paths to that product. Reaction paths that use the cheapest raw materials and produce the smallest quantities of byproducts are to be preferred. Reaction paths that produce significant quantities of unwanted byproducts should especially be avoided, since they can create significant environmental problems.

However, there are many other factors to be considered in the choice of reaction path. Some are commercial, such as uncertainties regarding future prices of raw materials and byproducts. Others are technical, such as safety and energy consumption.

The lack of suitable catalysts is the most common reason preventing the exploitation of novel reaction paths. At the first stage of design, it is impossible to look ahead and see all of the consequences of choosing one reaction path or another, but some things are clear even at this stage. Consider the following example.

Chemical Process Design and Integration, Second Edition. Robin Smith. © 2016 John Wiley & Sons, Ltd. Published 2016 by John Wiley & Sons, Ltd. Companion Website: www.wiley.com/go/smith/chemical2e

Example 4.1 Given that the objective is to manufacture vinyl chloride, there are at least three reaction paths that can be readily exploited (Rudd, Powers and Siirola, 1973).

Path 1

$$\begin{array}{c} C_2H_2 + HCl \\ acetylene \\ hydrogen \\ chloride \\ chloride \\ chloride \end{array} \rightarrow \begin{array}{c} C_2H_3Cl \\ vinyl \\ chloride \\ chlori$$

 $C_2H_4Cl_2$

dichloroethane

Path 2

$$C_2H_4 + Cl_2 -$$

ethylene chlorine

$$\begin{array}{c} C_2H_4Cl_2 \xrightarrow{heat} C_2H_3Cl + HCl_2 \\ \hline dichloroethane & vinyl \\ chloride & chloride \end{array}$$

Path 3

$$\begin{array}{c} C_2H_4 &+ \frac{1/_2O_2}{\text{oxygen}} + \frac{2\text{HCl}}{\text{hydrogen}} \longrightarrow \begin{array}{c} C_2H_4\text{Cl}_2 &+ H_2\text{O}\\ \text{dichloroethane} & & \text{water} \end{array}$$

The market values and molar masses of the materials involved are given in Table 4.1.

Table 4.1

Molar masses and values of materials in Example 4.1.

Material	Molar mass (kg·kmol ⁻¹)	Value (\$·kg ⁻¹)
Acetylene	26	1.0
Chlorine	71	0.23
Ethylene	28	0.58
Hydrogen chloride	36	0.39
Vinyl chloride	62	0.46

Example 4.2 Devise a process from the three reaction paths in Example 4.1 that uses ethylene and chlorine as raw materials and produces no byproducts other than water (Rudd, Powers and Siirola, 1973). Does the process look economically attractive?

Solution A study of the stoichiometry of the three paths shows that this can be achieved by combining Path 2 and Path 3 to obtain a fourth path.

Paths 2 and 3

$$\begin{array}{c} C_2H_4 \ + \ Cl_2 \ \longrightarrow \ C_2H_4Cl_2 \\ ethylene \ + \ l_2O_2 \ + \ 2HCl \\ ethylene \ & oxygen \ + \ 2HCl \\ hydrogen \\ chloride \ & dichloroethane \ + \ H_2O \\ dichloroethane \ + \ Water \ & dichloroethane \ + \ H_2O \\ ethylene \ & oxygen \ + \ 2C_2H_4Cl_2 \ + \ H_2O \\ dichloroethane \ & dichloroethane \ + \ H_2O \\ ethylene \ & biorde \ & biorde$$

Oxygen is considered to be free at this stage, coming from the atmosphere. Which reaction path makes most sense on the basis of raw material costs, product and byproduct values?

Solution Decisions can be made on the basis of the economic potential of the process. At this stage, the best that can be done is to define the economic potential (*EP*) as (see Chapter 2):

EP = (Value of products) - (Raw materials costs)

Path 1

$$EP = (62 \times 0.46) - (26 \times 1.0 + 36 \times 0.39)$$

= -11.52 \$\cdot kmol^{-1} vinyl chloride product

Path 2

$$EP = (62 \times 0.46 + 36 \times 0.39) - (28 \times 0.58 + 71 \times 0.23)$$

= 9.99 \\$ \cdot kmol^{-1} vinyl chloride product

This assumes the sale of the byproduct HCl. If it cannot be sold, then:

$$EP = (62 \times 0.46) - (28 \times 0.58 + 71 \times 0.23)$$

= -4.05 \$\cdot kmol^{-1} vinyl chloride product

Path 3

$$EP = (62 \times 0.46) - (28 \times 0.58 + 2 \times 36 \times 0.39)$$

= -15.8 \$\cdot kmol^{-1} vinyl chloride product

Paths 1 and 3 are clearly not viable. Only Path 2 shows a positive economic potential when the byproduct HCl can be sold. In practice, this might be quite difficult, since the market for HCl tends to be limited. In general, projects should not be justified on the basis of the byproduct value.

The preference is for a process based on ethylene rather than the more expensive acetylene, and chlorine, rather than the more expensive hydrogen chloride. Electrolytic cells are a much more convenient and cheaper source of chlorine than hydrogen chloride. In addition, it is preferred to produce no byproducts.

These three reactions can be added to obtain the overall stoichiometry.

Path 4

$$\begin{array}{c} 2C_{2}H_{4} + Cl_{2} & + l/_{2}O_{2} \longrightarrow 2C_{2}H_{3}Cl + H_{2}O \\ \text{ethylene} & \text{chlorine} & \text{oxygen} & \text{vinyl} \\ \text{chloride} & \text{water} \end{array}$$

or

$$\begin{array}{c} C_2H_4 \\ \text{ethylene} \end{array} + \begin{array}{c} 1/_2Cl_2 \\ \text{chlorine} \end{array} + \begin{array}{c} 1/_4O_2 \\ \text{oxygen} \end{array} \longrightarrow \begin{array}{c} C_2H_3Cl \\ \text{vinyl} \\ \text{water} \end{array}$$

Now the economic potential is given by:

$$EP = (62 \times 0.46) - (28 \times 0.58 + 1/2 \times 71 \times 0.23)$$

= 4.12 \$\cdot kmol^{-1} vinvl chloride product

In summary, Path 2 from Example 4.1 is the most attractive reaction path if there is a large market for hydrogen chloride. In practice, it tends to be difficult to sell the large quantities of hydrogen chloride produced by such processes. Path 4 is the usual commercial route to vinyl chloride.

4.2 Types of Reaction Systems

Having made a choice of the reaction path, a choice of reactor type must be made, together with some assessment of the conditions in the reactor. This allows assessment of the reactor performance for the chosen reaction path in order for the design to proceed.

Before proceeding to the choice of reactor and operating conditions, some general classifications must be made regarding the types of reaction systems likely to be encountered. Reaction systems can be classified into six broad types:

 Single reactions. Most reaction systems involve multiple reactions. In practice, the *secondary* reactions can sometimes be neglected, leaving a single *primary* reaction to consider. Single reactions are of the type:

$$FEED \longrightarrow PRODUCT \tag{4.1}$$

or

$$FEED \longrightarrow PRODUCT + BYPRODUCT \qquad (4.2)$$

or

$$FEED \ 1 + FEED \ 2 \longrightarrow PRODUCT \tag{4.3}$$

and so on.

An example of this type of reaction that does not produce a byproduct is *isomerization* (the reaction of a feed to a product with the same chemical formula but a different molecular structure). For example, allyl alcohol can be produced from propylene oxide (Waddams, 1978):

$$\begin{array}{c} CH_{3}HC - CH_{2} \longrightarrow CH_{2} = CHCH_{2}OH \\ O \\ \end{array}$$
propylene oxide allyl alcohol

An example of a reaction that does produce a byproduct is the production of acetone from isopropyl alcohol, which produces a hydrogen byproduct:

$$(CH_3)_2CHOH \longrightarrow CH_3COCH_3 + H_2$$

isopropyl acetone

2) *Multiple reactions in parallel producing byproducts.* Rather than a single reaction, a system may involve secondary reactions producing (additional) byproducts in *parallel* with the primary reaction. Multiple reactions in parallel are of the type:

$$FEED \longrightarrow PRODUCT$$

$$FEED \longrightarrow BYPRODUCT$$

$$(4.4)$$

or

$$FEED \longrightarrow PRODUCT + BYPRODUCT 1$$

FEED \longrightarrow BYPRODUCT 2 + BYPRODUCT 3 (4.5)

or

$$FEED 1 + FEED 2 \longrightarrow PRODUCT$$

$$FEED 1 + FEED 2 \longrightarrow BYPRODUCT$$
(4.6)

and so on.

An example of a parallel reaction system occurs in the production of ethylene oxide (Waddams, 1978):

$$CH_2 = CH_2 + \frac{1}{2}O_2 \longrightarrow H_2C \longrightarrow CH_2$$

ethylene oxygen ethylene oxide

with the parallel reaction:

$$CH_2 = CH_2 + 3O_2 \longrightarrow 2CO_2 + 2H_2O$$
ethylene carbon dioxide water

Multiple reactions might not only lead to a loss of materials and useful product but might also lead to byproducts being deposited on, or poisoning, catalysts (see Chapters 5 and 6).

3) *Multiple reactions in series producing byproducts.* Rather than the primary and secondary reactions being in parallel, they can be in *series.* Multiple reactions in series are of the type:

$$FEED \longrightarrow PRODUCT$$

$$PRODUCT \longrightarrow BYPRODUCT$$

$$(4.7)$$

or

$$FEED \longrightarrow PRODUCT + BYPRODUCT1$$

$$PRODUCT \longrightarrow BYPRODUCT2 + BYPRODUCT3$$

$$(4.8)$$

or

$$FEED 1 + FEED 2 \longrightarrow PRODUCT$$

$$PRODUCT \longrightarrow BYPRODUCT 1 + BYPRODUCT 2$$

$$(4.9)$$

and so on.

An example of a series reaction system is the production of formaldehyde from methanol:

$$\begin{array}{c} CH_{3}OH + \frac{1}{2}O_{2} \longrightarrow HCHO \\ methanol \\ oxygen \end{array} \rightarrow \begin{array}{c} HCHO \\ formaldehyde \end{array} + \begin{array}{c} H_{2}O \\ water \end{array}$$

A series reaction of the formaldehyde occurs:

$$\begin{array}{ccc} HCHO & \longrightarrow & CO & + & H_2 \\ formaldehyde & & carbon & hydrogen \\ monoxide & & \end{array}$$

As with parallel reactions, series reactions might not only lead to a loss of materials and useful products but might also lead to byproducts being deposited on, or poisoning, catalysts (see Chapters 5 and 6).

4) Mixed parallel and series reactions producing byproducts. In more complex reaction systems, both parallel and series reactions can occur together. Mixed parallel and series 4

reactions are of the type:

$$FEED \longrightarrow PRODUCT$$

$$FEED \longrightarrow BYPRODUCT \qquad (4.10)$$

$$PRODUCT \longrightarrow BYPRODUCT$$

or

$$FEED \longrightarrow PRODUCT$$

$$FEED \longrightarrow BYPRODUCT 1 \qquad (4.11)$$

$$PRODUCT \longrightarrow BYPRODUCT 2$$

or

$$FEED 1 + FEED 2 \longrightarrow PRODUCT$$

$$FEED 1 + FEED 2 \longrightarrow BYPRODUCT 1$$

$$PRODUCT \longrightarrow BYPRODUCT 2 + BYPRODUCT 3$$

$$(4.12)$$

An example of mixed parallel and series reactions is the production of ethanolamines by reaction between ethylene oxide and ammonia (Waddams, 1978):

$$\begin{array}{ccc} H_2C - CH_2 &+ & NH_3 \longrightarrow & NH_2CH_2CH_2OH \\ & & & & \\ O & & & \\ \end{array}$$

ethylene oxide ammonia monoethanolamine

$$NH_{2}CH_{2}CH_{2}OH + H_{2}C - CH_{2} \longrightarrow NH(CH_{2}CH_{2}OH)_{2}$$

monoethanolamine ethylene oxide diethanolamine

$$\begin{array}{rcl} \mathrm{NH}(\mathrm{CH}_{2}\mathrm{CH}_{2}\mathrm{OH})_{2} &+ & \mathrm{H}_{2}\mathrm{C}-\mathrm{CH}_{2} \longrightarrow & \mathrm{N}(\mathrm{CH}_{2}\mathrm{CH}_{2}\mathrm{OH})_{3} \\ & & & & \\ & & & \\ & & & & \\ & & & & \\ & & & \\ & & & & \\ & & & \\ & & & & \\ & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & \\ & & & & \\ & & & & \\ & & & \\ & & & & \\ & & & \\ & & & \\ & & & \\ & & & & \\ & & & \\ &$$

Here the ethylene oxide undergoes parallel reactions, whereas the monoethanolamine undergoes a series reaction to diethanolamine and triethanolamine.

5) Polymerization reactions. In polymerization reactions, monomer molecules are reacted together to produce a high molar mass polymer. Depending on the mechanical properties required of the polymer, a mixture of monomers might be reacted together to produce a high molar mass copolymer. There are two broad types of polymerization reactions, those that involve a termination step and those that do not (Denbigh and Turner, 1984). An example that involves a termination step is free-radical polymerization of an alkene molecule, known as addition polymerization. A free radical is a free atom or fragment of a stable molecule that contains one or more unpaired electrons. The polymerization requires a free radical from an initiator compound such as a peroxide. The initiator breaks down to form a free radical (e.g. $\bullet CH_3$ or $\bullet OH$), which attaches to a molecule of alkene and in so doing generates another free radical. Consider the polymerization of vinyl chloride from a free-radical initiator •R. An *initiation step* first occurs:

$$\overset{\bullet}{\operatorname{R}} + \overset{\bullet}{\operatorname{CH}}_{2} = \overset{\bullet}{\operatorname{CHCl}} \longrightarrow \overset{\bullet}{\operatorname{RCH}}_{2} - \overset{\bullet}{\operatorname{CHCl}} \overset{\bullet}{\operatorname{initiator}}$$

$$\overset{\operatorname{vinyl}}{\operatorname{chloride}} \overset{\circ}{\operatorname{free radical}}$$

A *propagation step* involving growth around an *active center* follows:

$$RCH_2 - \dot{C}HCl + CH_2 = CHCl \longrightarrow$$
$$RCH_2 - CHCl - CH_2 - \dot{C}HCl$$

and so on, leading to molecules of the structure:

$$R - (CH_2 - CHCl)_n - CH_2 - CHCl$$

Eventually, the chain is terminated by steps such as the union of two radicals that consume but do not generate radicals:

$$R - (CH_2 - CHCl)_n - CH_2 - CHCl + CHCl - CH_2$$
$$- (CHCl - CH_2)_m - R \longrightarrow$$
$$R - (CH_2 - CHCl)_n - CH_2 - CHCl - CHCl$$
$$- CH_2 - (CHCl - CH_2)_m - R$$

This termination step stops the subsequent growth of the polymer chain. The period during which the chain length grows, that is before termination, is known as the *active life* of the polymer. Other termination steps are possible.

The orientation of the groups along the carbon chain, its *stereochemistry*, is critical to the properties of the product. The stereochemistry of addition polymerization can be controlled by the use of catalysts. A polymer where repeating units have the same relative orientation is termed *stereoregular*.

An example of a polymerization without a termination step is *polycondensation* (Denbigh and Turner, 1984).

$$HO - (CH_2)_n - COOH + HO - (CH_2)_n - COOH$$
$$\longrightarrow HO - (CH_2)_n - COO - (CH_2)_n - COOH + H_2O,$$
etc.

Here the polymer grows by successive esterification with elimination of water and no termination step. Polymers formed by linking monomers with carboxylic acid groups and those that have alcohol groups are known as *polyesters*. Polymers of this type are widely used for the manufacture of artificial fibers. For example, the esterification of terephthalic acid with ethylene glycol produces polyethylene terephthalate.

6) *Biochemical reactions*. Biochemical reactions, often referred to as *fermentations*, can be divided into two broad types. In the first type, the reaction exploits the metabolic pathways in selected microorganisms (especially bacteria, yeasts, moulds and algae) to convert feed material (often called *substrate* in biochemical reactor design) to the required product. The general form of such reactions is:

$$FEED \xrightarrow{\text{microorganisms}} PRODUCT + [\text{More microorganisms}]$$
(4.13)

or

$$FEED 1 + FEED 2 \xrightarrow{\text{microorganisms}} PRODUCT$$

$$+ [More microorganisms]$$
(4.14)

and so on.

In such reactions, the microorganisms reproduce themselves. In addition to the feed material, it is likely that nutrients (e.g. a mixture containing phosphorus, magnesium, potassium, etc.) will need to be added for the survival of the microorganisms. Reactions involving microorganisms include:

- hydrolysis,
- oxidation,
- esterification,
- reduction.

An example of an oxidation reaction is the production of citric acid from glucose:

$$C_{6}H_{12}O_{6} + \frac{3}{2}O_{2} \longrightarrow HOOCCH_{2}COH(COOH)CH_{2}COOH$$

citric acid
+ 2H₂O

In the second group, the reaction is promoted by *enzymes*. Enzymes are the catalyst proteins produced by microorganisms that accelerate chemical reactions in microorganisms. The biochemical reactions employing enzymes are of the general form:

$$FEED \xrightarrow{\text{enzyme}} PRODUCT \qquad (4.15)$$

and so on.

Unlike reactions involving microorganisms, in enzyme reactions the catalytic agent (the enzyme) does not reproduce itself. An example in the use of enzymes is the isomerization of glucose to fructose:

$$\begin{array}{c} CH_2OH(CHOH)_4CHO \xrightarrow{enzyme} CH_2OHCO(CHOH)_3CH_2OH \\ glucose & fructose \end{array}$$

Although nature provides many useful enzymes, they can also be engineered for improved performance and new applications. Biochemical reactions have the advantage of operating under mild reaction conditions of temperature and pressure and are usually carried out in an aqueous medium rather than using an organic solvent.

4.3 Measures of Reactor Performance

Before exploring how reactor conditions can be chosen, some measure of reactor performance is required.

For polymerization reactors, the main concern is the characteristics of the product that relate to the mechanical properties. The distribution of molar masses in the polymer product, orientation of groups along the chain, cross-linking of the polymer chains, copolymerization with a mixture of monomers, and so on, are the main considerations. Ultimately, the main concern is the mechanical properties of the polymer product.

For biochemical reactions, the performance of the reactor will normally be dictated by laboratory results, because of the difficulty of predicting such reactions theoretically (Shuler and Kargi, 2002). There are likely to be constraints on the reactor performance dictated by the biochemical processes. For example, in the manufacture of ethanol using microorganisms, as the concentration of ethanol rises, the microorganisms multiply more slowly until at a concentration of around 12% it becomes toxic to the microorganisms.

For other types of reactors, three important parameters are used to describe their performance (Wells and Rose, 1986):

$$Conversion = \frac{(Reactant consumed in the reactor)}{(Reactant fed to the reactor)}$$
(4.16)

Selectivity =
$$\frac{\text{(Desired product produced)}}{\text{(Reactant consumed in the reactor)}}$$
 (4.17)
× Stoichiometric factor

Reactor yield =
$$\frac{\text{(Desired product produced)}}{(\text{Reactant fed to the reactor)}} \times \text{Stoichiometric factor}$$
 (4.18)

in which the stoichiometric factor is the stoichiometric moles of reactant required per mole of product. When more than one reactant is required (or more than one desired product produced), Equations 4.16 to 4.18 can be applied to each reactant (or product).

The following example will help clarify the distinctions among these three parameters.

Example 4.3 Benzene is to be produced from toluene according to the reaction (Douglas, 1985):

$$C_6H_5CH_3 + H_2 \longrightarrow C_6H_6 + CH_4$$

toluene hydrogen benzene methane

Some of the benzene formed undergoes a number of secondary reactions in series to unwanted byproducts that can be represented by the single reaction to diphenyl, according to the reaction:

Table 4.2 gives the compositions of the reactor feed and effluent streams.

Table 4.2

Reactor feed and effluent streams.

Component	Inlet flowrate (kmol·h ⁻¹)	Outlet flowrate (kmol·h ⁻¹)
H ₂	1858	1583
CH ₄	804	1083
C ₆ H ₆	13	282
C ₆ H ₅ CH ₃	372	93
$C_{12}H_{10}$	0	4

Calculate the conversion, selectivity and reactor yield with respect to the:

- a) toluene feed
- b) hydrogen feed.

Solution

a

)	Toluene conversion	=	(Toulene consumed in the reactor) (Toulene fed to the reactor) 372 - 93
	Stoichiometric factor	=	372 0.75 Stoichiometric moles of toluene required per mole of benzene produced
	Benzene selectivity from toluene	=	(Benzene produced in the reactor) (Toluene consumed in the reactor)
		=	× Stoichiometric factor $\frac{282 - 13}{372 - 93} \times 1$ 0.96
	Reactor yield of benzene from toluene	=	$\frac{(\text{Benzene produced in the reactor})}{(\text{Toluene fed to the reactor})}$
		=	× Stoichiometric factor $\frac{282 - 13}{372} \times 1$ 0.72
b)	Hydrogen conversion	=	$\frac{(\text{Hydrogen consumed in the reactor})}{(\text{Hydrogen fed to the reactor})}$ $\frac{1858 - 1583}{1858}$ 0.15
	Stoichiometric factor	=	Stoichiometric moles of hydrogen required per mole of benzene produced
	Benzene selectivity from hydrogen	=	(Benzene produced in the reactor) (Hydrogen consumed in the reactor)
		=	× Stoichiometric factor $\frac{282 - 13}{1858 - 1583} \times 1$ 0.98
	Reactor yield of benzene from hydrogen	=	(Benzene produced in the reactor) (Hydrogen fed to the reactor)
		=	× Stoichiometric factor $\frac{282-13}{1858} \times 1$ 0.14

Because there are two feeds to this process, the reactor performance can be calculated with respect to both feeds. However, the principal concern is performance with respect to toluene, since it is more expensive than hydrogen. As will be discussed in the next chapter, if a reaction is reversible there is a maximum conversion, the *equilibrium conversion*, that can be achieved, which is less than 1.0.

In describing reactor performance, selectivity is often a more meaningful parameter than reactor yield. Reactor yield is based on the reactant fed to the reactor rather than on that which is consumed. Part of the reactant fed might be material that has been recycled rather than fresh feed. Reactor yield takes no account of the ability to separate and recycle unconverted raw materials. Reactor yield is only a meaningful parameter when it is not possible for one reason or another to recycle unconverted raw material to the reactor inlet. However, the yield of the overall process is an extremely important parameter when describing the performance of the overall plant, as will be discussed later.

4.4 Rate of Reaction

To define the rate of a reaction, one of the components must be selected and the rate defined in terms of that component. The rate of reaction is the number of moles formed with respect to time, per unit volume of reaction mixture:

$$r_i = \frac{1}{V} \left(\frac{dN_i}{dt} \right) \tag{4.19}$$

where r_i = rate of reaction of Component *i* (kmol·m⁻³·s⁻¹)

 N_i = moles of Component *i* formed (kmol)

V = reaction volume (m³) t = time (s)

If the volume of the reactor is constant (V = constant):

$$r_i = \frac{1}{V} \left(\frac{dN_i}{dt} \right) = \frac{dN_i/V}{dt} = \frac{dC_i}{dt}$$
(4.20)

where $C_i = \text{molar concentration of Component } i \text{ (kmol} \cdot \text{m}^{-3})$

The rate is negative if the component is a reactant and positive if it is a product. For example, for the general irreversible reaction:

$$bB + cC + \cdots \longrightarrow sS + tT + \cdots$$
 (4.21)

The rates of reaction are related by:

$$-\frac{r_B}{b} = -\frac{r_C}{c} = -\dots = \frac{r_S}{s} = \frac{r_T}{t} = \dots$$
 (4.22)

If the rate-controlling step in the reaction is the collision of the reacting molecules, then the equation to quantify the reaction rate will often follow the stoichiometry such that:

$$-r_B = k_B C_B^b C_C^c \cdots \tag{4.23}$$

$$-r_C = k_C C_B^b C_C^c \cdots \tag{4.24}$$

$$r_S = k_S C^b_B C^c_C \cdots \tag{4.25}$$

$$r_T = k_T C_B^b C_C^c \cdots \tag{4.26}$$

- where r_i = reaction rate for Component *i* (kmol·m⁻³·s⁻¹) k_i = reaction rate constant for Component *i* ([kmol·m⁻³]^{NC-b-c-···s⁻¹})
 - NC = is the number of components in the rate expression
 - C_i = molar concentration of Component *i* (kmol·m⁻³)

The exponent for the concentration (b, c, ...) is known as the *order of reaction*. The reaction rate constant is a function of temperature, as will be discussed in the next chapter.

Thus, from Equations 4.22 to 4.26:

$$\frac{k_B}{b} = \frac{k_C}{c} = \dots = \frac{k_S}{s} = \frac{k_T}{t} = \dots$$
(4.27)

Reactions for which the rate equations follow the stoichiometry as given in Equations 4.23 to 4.26 are known as *elementary reactions*. If there is no direct correspondence between the reaction stoichiometry and the reaction rate, these are known as *nonelementary reactions* and are often of the form:

$$-r_B = k_B C_B^{\beta} C_C^{\delta} \cdots C_S^{\epsilon} C_T^{\xi} \cdots$$
(4.28)

$$-r_C = k_C C_B^{\beta} C_C^{\delta} \cdots C_S^{\epsilon} C_T^{\xi} \cdots$$
(4.29)

$$r_S = k_S C_B^{\beta} C_C^{\delta} \cdots C_S^{\epsilon} C_T^{\xi} \cdots$$
(4.30)

$$r_T = k_T C_B^{\beta} C_C^{\delta} \cdots C_S^{\epsilon} C_T^{\xi} \cdots$$
(4.31)

where β , δ , ε , ξ = order of reaction

The reaction rate constant and the order of reaction must be determined experimentally. If the reaction mechanism involves multiple steps involving chemical intermediates, then the form of the reaction rate equations can be of a more complex form than Equations 4.28 to 4.31.

If the reaction is reversible, such that:

$$bB + cC + \cdots \equiv sS + tT + \cdots$$
 (4.32)

then the rate of reaction is the net rate of the forward and reverse reactions. If the forward and reverse reactions are both elementary, then:

$$-r_B = k_B C_B^b C_C^c \cdots - k_B^\prime C_S^s C_T^t \cdots$$

$$(4.33)$$

$$-r_C = k_C C_B^b C_C^c \cdots - k_C^{\prime} C_C^s C_T^t \cdots$$

$$(4.34)$$

$$r_{S} = k_{S} C_{B}^{b} C_{C}^{c} \cdots - k_{S}^{\prime} C_{S}^{s} C_{T}^{t} \cdots$$

$$(4.35)$$

$$r_T = k_T C_B^b C_C^c \cdots - k_T^{\prime} C_S^s C_T^t \cdots$$
(4.36)

where k_i = reaction rate constant for Component *i* for the

forward reaction k'_i = reaction rate constant for Component *i* for the reverse reaction

If the forward and reverse reactions are nonelementary, perhaps involving the formation of chemical intermediates in multiple steps, then the form of the reaction rate equations can be more complex than Equations 4.33 to 4.36.

4.5 Idealized Reactor Models

Three idealized models are used for the design of reactors (Rase, 1977; Denbigh and Turner, 1984; Levenspiel, 1999). In the first (Figure 4.1a), the *ideal-batch* model, the reactants are charged at the beginning of the operation. The contents are subjected to perfect mixing for a certain period, after which the products are discharged. Concentration changes with time, but the perfect mixing ensures that at any instant the composition and temperature throughout the reactor are both uniform.

In the second model (Figure 4.1b), the *mixed-flow* or *continuous well-mixed* or *continuous-stirred-tank reactor* (*CSTR*), feed and product takeoff are both continuous and the reactor contents are assumed to be perfectly mixed. This leads to uniform composition and temperature throughout the reactor. Because of the perfect mixing, a fluid element can leave the instant it enters the reactor or stay for an extended period. The residence time of individual fluid elements in the reactor varies.

In the third model (Figure 4.1c), the *plug-flow* model, a steady uniform movement of the reactants is assumed, with no attempt to



The idealized models used for reactor design. (Reproduced from Smith R and Petela EA (1992) Waste Minimization in the Process Industries Part 2 Reactors, Chem Eng, Dec (509–510): 17, by permission of the Institution of Chemical Engineers.)

induce mixing along the direction of flow. Like the ideal-batch reactor, the residence time in a plug-flow reactor is the same for all fluid elements. Plug-flow operation can be approached by using a number of mixed-flow reactors in series (Figure 4.1d). The greater the number of mixed-flow reactors in series, the closer is the approach to plug-flow operation.

1) *Ideal-batch reactor*. Consider a batch reactor in which the feed is charged at the beginning of the batch and no product is withdrawn until the batch is complete. It is given that:

$$\begin{bmatrix} \text{Moles of reactant} \\ \text{converted} \end{bmatrix} = -r_i = -\frac{1}{V} \frac{dN_i}{dt}$$
(4.37)

Integration of Equation 4.37 gives:

$$t = \int_{N_{i0}}^{N_{it}} \frac{dN_i}{r_i V}$$
(4.38)

where t = batch time

 N_{i0} = initial moles of Component *i* N_{it} = final moles of Component *i* after time *t*

Alternatively, Equation 4.37 can be written in terms of reactor conversion X_i :

$$\frac{dN_i}{dt} = \frac{d[N_{i0}(1-X_i)]}{dt} = -N_{i0}\frac{dX_i}{dt} = r_i V$$
(4.39)

Integration of Equation 4.39 gives:

$$t = N_{i0} \int_{0}^{X_i} \frac{dX_i}{-r_i V}$$
(4.40)

Also, from the definition of reactor conversion, for the special case of a constant density reaction mixture:

$$X_i = \frac{N_{i0} - N_{it}}{N_{i0}} = \frac{C_{i0} - C_{it}}{C_{i0}}$$
(4.41)

where C_i = molar concentration of Component *i*

 C_{i0} = initial molar concentration of Component *i* C_{it} = final molar concentration of Component *i* at time *t*

Substituting Equation 4.41 into Equation 4.39 and noting that $N_{i0}/V = C_{i0}$ gives:

$$-\frac{dC_i}{dt} = -r_i \tag{4.42}$$

Integration of Equation 4.42 gives:

$$t = -\int_{C_{i0}}^{C_{i1}} \frac{dC_i}{-r_i}$$
(4.43)

2) Mixed-flow reactor. Consider now the mixed-flow reactor in Figure 4.2a in which a feed of Component *i* is reacting. A material balance for Component *i* per unit time gives:

$$\begin{bmatrix} \text{Moles of reactant in} \\ \text{feed per unit time} \end{bmatrix} - \begin{bmatrix} \text{Moles of reactant} \\ \text{converted per unit time} \end{bmatrix}$$
$$= \begin{bmatrix} \text{Moles of reactant} \\ \text{in product per unit time} \end{bmatrix}$$
(4.44)

Equation 4.44 can be written per unit time as:

$$N_{i,in} - (-r_i V) = N_{i,out}$$
(4.45)

where $N_{i,in}$ = inlet moles of Component *i* per unit time $N_{i,out}$ = outlet moles of Component *i* per unit time

Rearranging Equation 4.45 gives:

$$N_{i,out} = N_{i,in} + r_i V \tag{4.46}$$



