…is admirable for its encyclopedic nature and its
…ent bias in the choice of emphasis or coverage. This
…art from the other books on economic dynamics
…ied and scholarly coverage of a remarkably broad
…on economic dynamics. Not only is the book valu-
…lopedic reference book for researchers, but is an
…or a textbook on economic dynamics. Since Gan-
…ed to provide background in even some of the
…areas of dynamic economic research in a readable
…y unnecessarily advanced notation.
…on A. Barnett
…rnal of International and Comparative Economics

ECONOMIC DYNAMICS

…mathematical methods used in the eco-
…nics, this book shows how they are utilised
…analyze dynamical models. Accordingly,
…n the methods, and every new mathemati-
…e introduced is followed by its application
…omic models. The mathematical methods
…e from elementary linear difference and
…quations to simultaneous systems to the
…alysis of non-linear dynamical systems.
…siderations are stressed throughout, includ-
…vanced topics. Bifurcation and chaos theo-
…ealt with. The reader is guided through a
…analysis of each topic, be it a mathematical
…n economic model. The Study Edition also
…reader with solutions to the numerous

…0-62760-X

Gandolfo

ECONOMIC DYNAMICS

Giancarlo Gandolfo

ECONOMIC DYNAMICS

STUDY EDITION

Giancarlo Gandolfo

Economic Dynamics

Study Edition

With 65 Figures
and 6 Tables

Springer

Professor Dr. Giancarlo Gandolfo
University of Rome "La Sapienza"
Faculty of Economics
Via del Castro Laurenziano 9
I-00161 Roma, Italy

1st ed. 1997. Corr. 2nd printing

ISBN 3-540-62760-X Springer-Verlag Berlin Heidelberg New York

The Library of Congress has cataloged the 3rd completely rev. and enl. ed. as follows:
Gandolfo, Giancarlo. Economic dynamics / Giancarlo Gandolfo. – 3rd, Completely rev. and enl. ed. p. cm. Rev. ed. of: Economic dynamics, methods and models. 2nd, rev. ed. 1980, c1971. Includes bibliographical references and index. ISBN 3-540-60988-1 (hardcover: alk. paper) 1. Economics, Mathematical. 2. Economics – Mathematical models. 3. Statics and dynamics (Social sciences) I. Gandolfo, Giancarlo. Economic dynamics, methods and models. II. Title. HB135.G36 1996 330'.01'51–dc20

Cover design: Erich Kirchner, Heidelberg

SPIN 11504658 42/3153 - 5 4 3 2 1 – Printed on acid-free paper

To my wife
Luciana
(finally!)

Preface

This is a completely rewritten and much expanded version of the book published in 1980 (this, in turn, was an updated edition of my *Mathematical Methods and Models in Economic Dynamics*, originally published in 1971) and reprinted several times since then. In the last fifteen years so many new mathematical methods have been introduced in the study of dynamic economic models, that a revision was not sufficient, but a completely new edition was called for.

This book gives a comprehensive, but simple, treatment of the mathematical methods commonly used in dynamical economics. It also shows how they are applied to build and analyse economic models. Accordingly, the focus is on methods, and every mathematical technique introduced is followed by its application to selected economic models that serve as examples. The unifying principle in the exposition of the different economic models is then seen to be the common mathematical technique. This process will enable the reader not only to understand the basic literature, but also to build and analyse her or his own models.

'Comprehensive' means that this book contains an unusually broad set of mathematical methods. The standard constant-coefficient linear difference and differential equations and simultaneous systems are of course thoroughly explained from scratch. Mixed differential-difference equations are introduced by a discussion of discrete and continuous time in economic modelling. An extensive treatment of non-linear dynamic equations and qualitative methods (from phase diagrams to non-linear oscillations and bifurcation theory) is given. The saddle-path properties of solutions to dynamic optimization problems and to rational expectations models with jump variables are explained rigorously. Stability considerations are stressed throughout, including advanced topics such as Liapunov's second method, structural stability, conditional stability, and so on. The theory of chaotic dynamics is examined at some length and appraised in relation to stochastic erraticity. Synergetics and catastrophe theory are also (briefly) treated.

'Simple' means that the lowest possible level of mathematical prerequisites is presupposed in the reader. He or she is assumed to have no previous familiarity with the topics treated. Accordingly, every subject is worked out in great detail, and no essential step in the argument is omitted. The reader is guided through a step-by-step analysis of each topic, be it a mathematical method or an economic model. This 'user-friendly' feature, which is also present in the exercises, will undoubtedly be appreciated by students.

The required background for Part I consists of elementary algebra, for Part II of the rudiments of calculus. Advanced matrix algebra is used when necessary, but the main propositions can be understood without it. More mathematical background is needed for Part III; however, this does not go much beyond the knowledge acquired in any basic course of mathematics for economists. It also should be pointed out that the organisation of Part III is different from that of Parts I and II, where mathematical methods and illustrative economic models are treated in separate chapters. This separation is didactically convenient when the mathematical chapter deals with a specific and limited topic, e.g. difference equations of a given order, linear and with constant coefficients. Since each chapter of Part III usually has a much broader mathematical subject matter—bifurcation theory, for example—the economic illustrations are more conveniently presented with the mathematical methods.

This book has been written for economists, but applied mathematicians interested in getting a bird's eye view of the economic applications of the mathematical methods under consideration may find the economic parts of the book useful. In fact, although the main selection criterion has been the models' suitability to illustrate the mathematical point, both old classics and new research results have been included, as well as both microeconomic and macroeconomic models. Also, students in time series analysis who need a grounding in difference equations will find Part I useful.

* * *

I am grateful to the students from all over the world who have written me over the years to indicate unclear points and misprints, and to Flavio Casprini, Nicola Cetorelli, Giuseppe De Arcangelis, Vivek H. Dehejia, Daniela Federici, Maria Maddalena Giannetti, Michael D. Intriligator, Giovanna Paladino, Maria Luisa Petit, Francesca Sanna Randaccio, for their advice and comments. Claude Hillinger and Karlhans Sauernheimer discussed with me the project of this book when I was visiting CES (Center for Economic Studies of the University of Munich) in 1992 and 1994, and gave me many useful suggestions. Finally, I would like to thank the University of Rome 'La Sapienza' for generous support in terms of research funds and a year of sabbatical leave. None of the persons and institutions mentioned has any responsibility for possible deficiencies that might remain.

Giancarlo Gandolfo, University of Rome 'La Sapienza', December 1995

PREFACE TO THE SOFTCOVER EDITION

This student edition differs from the previous hardcover edition in one important respect: the addition of a new section containing the answers to *all* exercises. The help of Daniela Federici and Maria Maddalena Giannetti is gratefully acknowledged. Also, misprints and errors have been corrected thanks to Serena Sordi.
In the 2005 reprint further misprints and errors have been corrected thanks to Lisbeth Fajstrup and Federico Trionfetti.

University of Rome 'La Sapienza', December 1996 and April 2005

Contents

PREFACE . . . VII

1 Introduction **1**
1.1 Definition . . . 1
1.2 Functional equations . . . 2
1.3 References . . . 3

I LINEAR DIFFERENCE EQUATIONS **5**

2 Difference Equations: General Principles **7**
2.1 Definitions . . . 7
2.2 Linear difference equations with constant coefficients . . . 9
2.2.1 The homogeneous equation . . . 10
2.2.2 The non-homogeneous equation . . . 12
2.3 Determination of the arbitrary constants . . . 13
2.4 References . . . 14

3 First-order Difference Equations **17**
3.1 Solution of the homogeneous equation . . . 17
3.2 Particular solution of the non-homogeneous equation . . . 21
3.2.1 $g(t)$ is a constant . . . 21
3.2.2 $g(t)$ is an exponential function . . . 22
3.2.3 $g(t)$ is a polynomial function of degree m . . . 23
3.2.4 $g(t)$ is a trigonometric function of the sine-cosine type . 23
3.2.5 $g(t)$ is a combination of the previous functions . . . 24
3.2.6 The case when $g(t)$ is a generic function of time. Backward and forward solutions . . . 24
3.3 General solution of the non-homogeneous equation . . . 27
3.4 A digression on distributed lags and partial adjustment equations . . . 28
3.5 Exercises . . . 31
3.5.1 Example . . . 31
3.5.2 Other exercises . . . 32
3.6 References . . . 33

4 **First-order Difference Equations in Economic Models** 35
4.1 The cobweb theorem . . . 35
4.1.1 The cobweb model and expectations . . . 38
4.1.1.1 The normal price . . . 39
4.1.1.2 Adaptive expectations . . . 40
4.2 The dynamics of multipliers . . . 43
4.2.1 The basic case . . . 43
4.2.2 Other multipliers . . . 45
4.2.2.1 A foreign trade multiplier . . . 46
4.2.2.2 Taxation . . . 47
4.3 Exercises . . . 48
4.4 References . . . 51

5 **Second-order Difference Equations** 53
5.1 Solution of the homogeneous equation . . . 53
5.1.1 Positive discriminant ($\Delta > 0$) . . . 54
5.1.2 Null discriminant ($\Delta = 0$) . . . 55
5.1.3 Negative Discriminant ($\Delta < 0$) . . . 56
5.1.4 Stability conditions . . . 58
5.2 Solution of the non-homogeneous equation. . . . 59
5.2.1 The operational method . . . 61
5.3 Determination of the arbitrary constants . . . 63
5.4 Exercises . . . 65
5.4.1 Example . . . 65
5.4.2 Other exercises . . . 68
5.5 References . . . 69

6 **Second-order Difference Equations in Economic Models** 71
6.1 Multiplier-accelerator interaction: the prototype model . . . 71
6.1.1 Graphical location of the roots . . . 73
6.2 Market adjustments and rational expectations . . . 75
6.3 Exercises . . . 76
6.4 References . . . 80

7 **Higher-order Difference Equations** 83
7.1 Solution of the homogeneous equation . . . 83
7.2 Particular solution of the non-homogeneous equation . . . 84
7.2.1 The operational method . . . 85
7.3 Determination of the arbitrary constants . . . 87
7.4 Stability conditions . . . 87
7.4.1 Necessary and sufficient stability conditions (Samuelson's form) . . . 88
7.4.2 Necessary and sufficient stability conditions (Cohn-Schur form) . . . 89

7.5 Exercises . . . 91
7.5.1 Example . . . 91
7.5.2 Other exercises . . . 92
7.6 References . . . 92

8 **Higher-order Difference Equations in Economic Models** 95
8.1 Inventory cycles . . . 95
8.2 Distributed lags and interaction between the multiplier and the accelerator . . . 98
8.3 Exercises . . . 100
8.4 References . . . 102

9 **Simultaneous Systems of Difference Equations** 103
9.1 First-order 2×2 systems in normal form . . . 103
9.1.1 General solution of the homogeneous system: first method . . . 103
9.1.2 General solution of the homogeneous system: second (or direct) method . . . 106
9.1.2.1 Unequal real roots . . . 106
9.1.2.2 Equal real roots . . . 108
9.1.2.3 Complex roots . . . 109
9.1.3 Particular solution. Determination of the arbitrary constants . . . 110
9.2 First order n×n systems in normal form . . . 111
9.2.1 Direct matrix solution . . . 114
9.2.2 Stability conditions . . . 115
9.2.2.1 A digression on not-wholly-unstable systems . 118
9.2.2.2 Proof of the stability conditions . . . 120
9.2.3 Particular solution . . . 121
9.2.3.1 The operational method . . . 122
9.2.4 Determination of the arbitrary constants . . . 125
9.3 General systems . . . 126
9.3.1 First-order systems not in normal form . . . 126
9.3.2 Higher-order systems . . . 127
9.3.2.1 An example . . . 127
9.3.2.2 The general case . . . 129
9.3.2.3 Transformation of a higher-order system into a first-order system in normal form . . . 130
9.3.2.4 Stability conditions for higher-order systems . 132
9.4 Exercises . . . 132
9.4.1 Example . . . 132
9.4.2 Other exercises . . . 133
9.5 References . . . 134

10 Simultaneous Difference Systems in Economic Models **137**
10.1 Cournot oligopoly . . . 137
10.2 Multiplier effects in an open economy . . . 140
10.3 Exercises . . . 143
10.4 References . . . 143

II LINEAR DIFFERENTIAL EQUATIONS **145**

11 Differential Equations: General Principles **147**
11.1 Definitions . . . 147
11.2 Linear differential equations with constant coefficients . . . 148
11.2.1 The homogeneous equation . . . 149
11.2.2 The non-homogeneous equation . . . 150
11.3 Determination of the arbitrary constants . . . 152
11.4 References . . . 154

12 First-order Differential Equations **155**
12.1 Solution of the homogeneous equation . . . 155
12.2 Particular solution of the non-homogeneous equation . . . 158
12.2.1 $g(t)$ is a constant . . . 158
12.2.2 $g(t)$ is an exponential function . . . 159
12.2.3 $g(t)$ is a polynomial function of degree m . . . 159
12.2.4 $g(t)$ is a trigonometric function of the sine-cosine type . 160
12.2.5 $g(t)$ is a combination of the previous functions . . . 160
12.2.6 $g(t)$ is a generic function of time. The method of variation of parameters . . . 161
12.3 General solution of the non-homogeneous equation . . . 162
12.4 Continuously distributed lags and partial adjustment equations 163
12.5 Exercises . . . 165
12.5.1 Example . . . 165
12.5.2 Other exercises . . . 167
12.6 References . . . 167

13 First-order Differential Equations in Economic Models **169**
13.1 Stability of supply and demand equilibrium . . . 169
13.2 The neoclassical growth model . . . 175
13.2.1 Existence of a growth equilibrium . . . 176
13.2.2 Stability of growth equilibrium . . . 178
13.2.3 Refinements . . . 181
13.2.3.1 Depreciation and technical progress . . . 181
13.2.3.2 Golden rule . . . 183
13.2.4 Further developments . . . 184
13.2.4.1 Adjustment time or, how long is the long run? 184

13.2.4.2 β-convergence, σ-convergence, and all that . . 187
13.2.4.3 Endogenous growth . . . 189
13.3 Exercises . . . 189
13.4 References . . . 191

14 Second-order Differential Equations **193**
14.1 Solution of the homogeneous equation . . . 193
14.1.1 Positive discriminant ($\Delta > 0$) . . . 194
14.1.2 Null discriminant ($\Delta = 0$) . . . 195
14.1.3 Negative discriminant ($\Delta < 0$) . . . 196
14.1.4 Stability conditions . . . 198
14.2 Particular solution of the non-homogeneous equation . . . 199
14.2.1 Variation of parameters . . . 200
14.3 General solution of the non-homogeneous equation . . . 202
14.4 Determination of the arbitrary constants . . . 203
14.5 Exercises . . . 203
14.5.1 Examples . . . 203
14.5.2 Other exercises . . . 205
14.6 References . . . 205

15 Second-order Differential Equations in Economic Models **207**
15.1 The second-order accelerator . . . 207
15.2 Exercises . . . 210
15.3 References . . . 212

16 Higher-order Differential Equations **213**
16.1 Solution of the homogeneous equation . . . 213
16.2 Solution of the non-homogeneous equation . . . 215
16.2.1 Variation of parameters . . . 215
16.3 Determination of the arbitrary constants . . . 218
16.4 Stability conditions . . . 219
16.4.1 Necessary and sufficient stability conditions (Routh-Hurwitz) . . . 221
16.4.2 Necessary and sufficient stability conditions (Liénard-Chipart) . . . 223
16.5 Exercises . . . 223
16.5.1 Example . . . 223
16.5.2 Other exercises . . . 225
16.6 References . . . 225

17 Higher-order Differential Equations in Economic Models **227**
17.1 Feedback control and stabilisation policies . . . 227
17.1.1 Introduction . . . 227
17.1.2 Three types of stabilisation policy . . . 228

17.1.2.1 Proportional stabilisation policy 231
17.1.2.2 Mixed proportional-derivative stabilisation policy 232
17.1.2.3 Integral stabilisation policy 233
17.2 Exercises 234
17.3 References 235

18 Simultaneous Systems of Differential Equations 237
18.1 First-order 2×2 systems in normal form 237
18.1.1 General solution of the homogeneous system: first method 238
18.1.2 General solution of the homogeneous system: second (or direct) method 240
18.1.2.1 Unequal real roots 241
18.1.2.2 Equal real roots 243
18.1.2.3 Complex roots 244
18.1.3 Particular solution. Determination of the arbitrary constants 245
18.2 First order n×n systems in normal form 245
18.2.1 Solution of the homogeneous system 247
18.2.1.1 The matrix exponential 249
18.2.2 Stability conditions 251
18.2.2.1 D-stability, and stabilisation of matrices . . . 254
18.2.2.2 Sensitivity analysis 255
18.2.2.3 A digression on not-wholly-unstable systems . 259
18.2.2.4 Proof of the stability conditions 262
18.2.3 Particular solution 264
18.2.3.1 Variation of parameters 264
18.2.4 Determination of the arbitrary constants 265
18.3 General systems 266
18.3.1 First-order systems not in normal form 267
18.3.2 Higher-order systems 268
18.3.2.1 An example 268
18.3.2.2 The general case 270
18.3.2.3 Transformation of a higher-order system into a first-order system in normal form 271
18.3.2.4 Stability conditions for higher-order systems . 273
18.4 Exercises 273
18.4.1 Example 273
18.4.2 Other exercises 275
18.5 References 277

19 Differential Equation Systems in Economic Models 279
19.1 Stability of Walrasian general equilibrium of exchange 279
19.1.1 Static stability 280
19.1.2 Dynamic stability 283
19.2 Human capital in a growth model 286
19.3 A digression on 'arrow diagrams' 291
19.4 Balanced growth in a multi-sector economy 293
19.5 Exercises 298
19.6 References 301

III ADVANCED TOPICS 303

20 Comparative Statics and the Correspondence Principle 305
20.1 Introduction 305
20.2 The method of comparative statics 306
20.2.1 Purely qualitatively comparative statics 310
20.2.2 The inverse comparative statics problem 310
20.3 Comparative statics and optimizing behaviour 311
20.4 Comparative statics and dynamic stability of equilibrium: the 'correspondence principle' 314
20.4.1 Criticism and qualifications 316
20.5 Extrema and dynamic stability 318
20.5.1 An application to the theory of the firm 323
20.6 Elements of comparative dynamics 324
20.7 An illustrative application of the correspondence principle: the IS-LM model 325
20.8 Exercises 328
20.9 References 329

21 Stability of Equilibrium: A General Treatment 331
21.1 Introduction 331
21.2 Basic concepts and definitions 332
21.2.1 Stability 333
21.2.2 Further definitions 337
21.2.3 Structural stability 338
21.3 Qualitative methods: phase diagrams 341
21.3.1 Single equations 342
21.3.2 Two-equation simultaneous systems 346
21.3.2.1 Introduction: phase plane and phase path . . 346
21.3.2.2 Singular points 347
21.3.2.3 Graphical construction of the trajectories . . 350
21.3.2.4 Linear systems 356
21.4 Quantitative methods 360

21.4.1 Linearisation . . . 360
21.5 Elements of the qualitative theory of difference equations . . . 363
21.5.1 Single difference equations . . . 363
21.5.2 Two simultaneous difference equations . . . 368
21.6 Economic applications . . . 368
21.7 Exercises . . . 369
21.8 References . . . 370

22 Saddle Points and Economic Dynamics 373
22.1 Saddle points in optimal control problems . . . 374
22.1.1 Introduction . . . 374
22.1.2 The maximum principle . . . 375
22.2 Optimal economic growth . . . 377
22.2.1 Optimal growth: traditional . . . 378
22.2.1.1 The setting of the problem . . . 378
22.2.1.2 The optimality conditions in the basic neoclassical model . . . 381
22.2.1.3 Saddle-point transitional dynamics in the basic neoclassical model . . . 384
22.2.2 Optimal growth: endogenous . . . 386
22.2.2.1 A model of optimal endogenous growth . . . 386
22.2.2.2 The conditions for optimal endogenous growth 388
22.2.2.3 Optimal endogenous growth: saddle-point transitional dynamics . . . 390
22.3 Rational expectations and saddle points . . . 394
22.3.1 Introduction . . . 394
22.3.2 Rational expectations, saddle points, and overshooting 397
22.3.3 Rational expectations and saddle points: the general case . . . 401
22.4 Exercises . . . 403
22.5 References . . . 405

23 Liapunov's Second Method 407
23.1 General concepts . . . 407
23.2 The fundamental theorems . . . 408
23.3 Some economic applications . . . 413
23.3.1 Global stability of Walrasian general equilibrium . . . 413
23.3.2 Rules of thumb in business management . . . 420
23.3.3 Price adjustment and oligopoly under product differentiation . . . 421
23.4 Exercises . . . 425
23.5 References . . . 426

24 Introduction to Nonlinear Dynamics 429
24.1 Preliminary remarks . . . 429
24.2 Some integrable differential equations . . . 431
24.2.1 First-order and first-degree exact equations . . . 431
24.2.2 Linear equations of the first order with variable coefficients . . . 434
24.2.3 The Bernoulli equation . . . 436
24.3 Limit cycles and relaxation oscillations . . . 437
24.3.1 Limit cycles: the general theory . . . 437
24.3.2 Limit cycles: relaxation oscillations . . . 439
24.3.3 Kaldor's non-linear cyclical model . . . 441
24.3.3.1 The model . . . 441
24.3.3.2 Kaldor via relaxation oscillations . . . 445
24.3.3.3 Kaldor via Poincaré's limit cycle . . . 447
24.4 The Lotka-Volterra equations . . . 449
24.4.1 Construction of the integral curves . . . 454
24.4.2 Conservative and dissipative systems, and irreversibility . . . 456
24.4.3 Goodwin's growth cycle . . . 458
24.4.3.1 The model . . . 458
24.4.3.2 The phase diagram of the model . . . 461
24.5 Exercises . . . 464
24.6 References . . . 466

25 Bifurcation Theory 469
25.1 Introduction . . . 469
25.2 Bifurcations in continuous time systems . . . 469
25.2.1 Codimension-one bifurcations . . . 471
25.2.2 The Hopf bifurcation . . . 475
25.2.3 Sensitivity analysis and bifurcations: a reminder . . . 479
25.2.4 Kaldor's non-linear cyclical model again . . . 480
25.2.5 Oscillations in optimal growth models . . . 481
25.2.5.1 The model . . . 481
25.2.5.2 The optimality conditions . . . 483
25.2.5.3 Emergence of a Hopf bifurcation . . . 484
25.2.6 Cycles in an IS-LM model with pure money financing . 486
25.3 Bifurcations in discrete time systems . . . 488
25.3.1 Codimension-one bifurcations . . . 489
25.3.2 The Hopf bifurcation in discrete time . . . 491
25.3.3 Kaldor's cyclical model in discrete time . . . 492
25.3.4 Liquidity costs in the firm . . . 494
25.3.4.1 The model . . . 494
25.3.4.2 The dynamics . . . 497
25.4 Exercises . . . 499

25.5 References 501

26 Complex Dynamics **503**
26.1 Introduction 503
26.2 Discrete time systems and chaos 505
26.2.1 The logistic map 505
26.2.2 Intermittency 512
26.2.3 The basic theorems 512
26.2.4 Discrete time chaos in economics 515
26.2.4.1 Chaos in growth theory 515
26.2.4.2 Exchange rate dynamics and chaos 516
26.3 Continuous time systems and chaos 518
26.3.1 The Lorenz equations, strange attractors, and chaos . . 518
26.3.2 Other routes to continuous time chaos 521
26.3.2.1 The Rössler attractor 521
26.3.2.2 The Shil'nikov scenario 521
26.3.2.3 The forced oscillator 522
26.3.2.4 The coupled oscillator 522
26.3.3 International trade as the source of chaos. 526
26.3.4 A chaotic growth cycle 527
26.4 Significance and detection of chaos: Stochastic dynamics or chaos? 528
26.5 Other approaches 532
26.5.1 Introduction 532
26.5.2 Fast and slow, and synergetics 533
26.5.3 Catastrophe theory 536
26.6 Exercises 538
26.7 References 540

27 Mixed Differential-Difference Equations **545**
27.1 General concepts 545
27.2 Continuous vs discrete time in economic models 545
27.3 Linear mixed equations 550
27.4 The method of solution 550
27.5 Stability conditions 555
27.6 Delay differential equations and chaos 556
27.7 Some economic applications 556
27.7.1 Kalecki's business cycle model 557
27.7.1.1 The model 557
27.7.1.2 The dynamics 559
27.7.2 A formalization of the classical price-specie-flow mechanism of balance of payments adjustment 563
27.7.2.1 The model 563
27.7.2.2 Stability 565

27.8 Exercises 567
27.9 References 568

Bibliography **571**

Name Index **593**

Subject Index **599**

Answers to Exercises **611**

List of Figures

3.1 (a) Monotonic and divergent; (b) Oscillatory and convergent; (c) Monotonic and divergent; (d) Oscillatory and divergent; (e) Constant; (f) Oscillatory with constant amplitude. 19
3.2 Joining the points of the solution to a difference equation . . 20

4.1 The cobweb theorem 36
4.2 The cobweb theorem: abnormal cases 38
4.3 The simple dynamic multiplier 45

5.1 Second-order difference equations: real roots 60
5.2 An apparently divergent solution 64
5.3 A constant-amplitude oscillation 68

6.1 Samuelson's multiplier-accelerator diagram 74
6.2 Roots in Hicks' second-order model 79

8.1 Metzler's inventory cycle: stability regions 97

12.1 Solutions to a first-order differential equation 157

13.1 a) excess demand, b) excess demand price 171
13.2 The neoclassical aggregate growth model 178

17.1 Proportional stabilisation policy in the Phillips model 232

19.1 Stability of growth equilibrium with human capital 288
19.2 Construction of an arrow diagram 292

21.1 Stability 334
21.2 Asymptotic stability 335
21.3 First-order non-linear differential equation: phase diagram . . 343
21.4 First-order non-linear differential equation: spiral-like phase diagram 344
21.5 Singular points of 2×2 systems: the node 348
21.6 Singular points of 2×2 systems: the saddle point 349
21.7 Singular points of 2×2 systems: the focus 350
21.8 Singular points of 2×2 systems: the centre 351

21.9 Multiple singular points of 2×2 systems 352
21.10 Nature of the points not lying on $\varphi_1 = 0$ 353
21.11 Graphical determination of the nature of a singular point . . 354
21.12 Arrows and centres 356
21.13 Limit cycles 357
21.14 Phase diagram for a first-order difference equation: convergent monotonic movement 364
21.15 Phase diagram for a first-order difference equation: divergent monotonic movement 365
21.16 Phase diagram for a first-order difference equation: various oscillatory movements 367

22.1 Equilibrium point of the neoclassical optimal growth model . 383
22.2 The neoclassical optimal growth model: transitional dynamics 384
22.3 An optimal endogenous growth model: transitional dynamics 392
22.4 Rational expectations and saddle path dynamics: the overshooting phenomenon 396
22.5 Rational expectations and exchange-rate overshooting 400

23.1 Perfect complementarity 419

24.1 Kaldor's business cycle model: the investment and saving functions 442
24.2 The dynamics of Kaldor's cycle 444
24.3 Kaldor's model via Poincaré-Bendixson 448
24.4 Lotka-Volterra equations, diagram 1 452
24.5 Lotka-Volterra equations, diagram 2 453
24.6 Lotka-Volterra equations, construction of the orbits 455
24.7 The real wage rate and employment 460
24.8 Goodwin's growth cycle 462
24.9 Path of u and v in Goodwin's growth cycle 463
24.10 Goodwin's elementary non-linear cyclical model 465

25.1 Saddle-node (or fold) bifurcation, $y' = \alpha - y^2$ 470
25.2 Transcritical bifurcation, $y' = \alpha y - y^2$ 472
25.3 Pitchfork bifurcation, $y' = \alpha y - y^3$ 474
25.4 The Hopf bifurcation 478
25.5 The fold bifurcation in discrete time, $y_{t+1} = \alpha - y_t^2$ 489
25.6 The supercritical flip bifurcation, $y_{t+1} = \alpha y_t - \alpha y_t^2$ 490

26.1 The logistic curve $y_{t+1} = \alpha y_t(1 - y_t)$ 506
26.2 Flip bifurcation in the logistic equation 507
26.3 Stylized period doubling in the logistic equation 508
26.4 Windows in the logistic equation 509
26.5 Chaos in the logistic difference equation 510

26.6 An example of intermittency 511
26.7 The Lorenz equations (Guckenheimer and Holmes) 519
26.8 A two-dimensional torus in three-dimensional space 524
26.9 A catastrophe in inflation (Woodcock and Davis) 537

27.1 Roots of Kalecki's model 562

List of Tables

3.1 Numerical computation of the solution to a first-order difference equation . . . 32

4.1 Numerical example of the multiplier . . . 50

5.1 Constant-amplitude trigonometric oscillation . . . 67

6.1 Roots of Hicks' model . . . 78
6.2 Stability condition in Goodwin's model . . . 80

21.1 Singular points of the linear system $y_1' = a_{11}y_1 + a_{12}y_2, \quad y_2' = a_{21}y_1 + a_{22}y_2$ . . . 358

Chapter 1

Introduction

1.1 Definition

'A system is dynamical if its behavior over time is determined by functional equations in which variables at different points of time are involved in an essential way'.

This definition, due to Frisch (1936) and Samuelson (1947), is based on the *formal* characteristics common to all problems studied by economic dynamics, also called dynamical economics (these two expressions will be used interchangeably). Other definitions, based on the economic substance of those problems, are possible: e.g., economic dynamics is concerned with growth, or business cycles, or stability, or economic change. Such definitions have in fact been suggested: see, for example Baumol (1970), Day (1994), and the instructive survey by Machlup (1959). But they are inevitably partial, and a complete definition of this type would reduce to a cumbersome list of problems, with the danger of omitting some of them and of not including newly arising problems. The Frisch-Samuelson formal definition, on the contrary, is precise and general, and we shall adopt it.

Before commenting on it, let us recall that, according to another formal definition (Hicks, 1939), economic dynamics is identified with those parts of economic theory where every quantity must be dated, whereas in economic statics we need not trouble about dating. But in this way we would include in economic dynamics many non-dynamic phenomena. For example, think of a case in which all quantities have the same date. This may mean that a certain phenomenon has taken place at a certain point of time (and this may be important, but is not dynamics), or that a variable at time t depends on another variable at the same time t (and this too may be important – e.g. consumption is assumed to depend on current income and not on lagged income– but, again, it is not dynamics). The definition based on 'dating', then, is too vague and cannot be accepted.

1.2 Functional equations

Let us now turn back to the initial definition and intuitively explain what a *functional equation* is. The general theory of functional equations is outside the scope of the present book, and we shall give only some basic notions, which are sufficient for our purposes.

The basic concept is the following: *a functional equation is an equation where the unknown is a function.* Everybody knows that to solve an equation means to find that value (or those values) of the unknown which satisfy the equation. Now, to solve a functional equation means to find an unknown function which satisfies the functional equation identically.

Two points must be stressed in this definition. The first is that by *function* we mean the *form* of the function, apart from arbitrary constants (for example, $y = Ae^x$, where A is an arbitrary constant). As we shall see when explaining the various functional equations appearing in this book, the solution of a functional equation determines the form of the unknown function, and the determination of the arbitrary constant(s) requires additional conditions.

The second point is that by '*to satisfy identically*' we mean that the function we have to find must satisfy the functional equation for *any* admissible value of the independent variable appearing in the function. The following simple example may clarify this point.

Let us consider the functional equation $y'(x) - y(x) = 0$. We must find a specific function (in one independent variable) which identically satisfies the stated equation, i.e. a function such that, for any value of its argument (x), the value of the function and the value of its first derivative are equal. It is easy to check that this function is $y(x) = Ae^x$, since, from elementary calculus, $y'(x) = Ae^x = y(x)$ for any x. Now consider the function $y = ax+b$, which gives $y' = a$; if we let $x = (a-b)/a$, we have also $y = a$, i.e. $y' = y$. However, for any other value of x the value of the function will be different from a: therefore, the function $y = ax + b$ does *not* satisfy our functional equation identically. As a matter of terminology, from now on we shall usually omit 'identically', it being understood that to 'satisfy' a functional equation means to satisfy it identically.

In general, the symbol x can stand for any variable, not only time. This obvious remark is useful to avoid the mistake of believing that, in economics, functional equations are used only in dynamical problems (an example of a case outside economic dynamics is the classic problem of obtaining a utility function knowing the marginal rate of substitution). Since this is a book on economic dynamics only, from now on we shall assume that x stands for time, and use t instead of x. We are now ready to understand the second part of the definition of dynamical economics.

In fact, $y'(t) = y(t)$ can be considered as a relation that involves the value of y at any point of time and the value it has at an arbitrarily close point, determined by y'. The 'different points of time' clause is necessary to exclude the case, already mentioned above, of quantities dated at the same point of time. Time must enter in an 'essential' way: for example, if it enters *only* as a unit of measurement (i.e. because we are dealing with quantities that are flows per unit of time), the system is not dynamical.

The simplest types of functional equations, widely used in dynamical economics, are linear, constant-coefficient difference and differential equations (the meaning of these words will be clarified in the following treatment) and the relative simultaneous systems. These functional equations will be dealt with in Part I and Part II respectively. In Part III other types will be treated, together with more advanced material concerning the qualitative theory of differential and difference equations.

1.3 References

Throughout this book, the end-of-chapter references will be indicated only by name(s), date, title. Complete information as to publisher, place of publication, etc., is contained in the Bibliography at the end of the volume.

Baumol, W.J., 1970, *Economic Dynamics,* Chap. 1 and Appendix to Chap. 8.

Day, R., 1994, *Complex Economic Dynamics, Vol. I,* Chaps. 1, 2.

Frisch, R., 1936, On the Notion of Equilibrium and Disequilibrium.

Hicks, J.R., 1939, 1946, *Value and Capital,* Chap. IX, Sect. 1.

Hicks, J.R., 1965, *Capital and Growth,* Chap. 1.

Machlup, F., 1959, Statics and Dynamics: Kaleidoscopic Words.

Samuelson, P.A., 1947, *Foundations of Economic Analysis,* Chap. XI, pp. 311-17.

Samuelson, P.A., 1949, Dynamic Process Analysis, Sects. I, II.

Volterra, V., 1959 (1929), *Theory of Functionals and of Integral and Integro-Differential Equations,* Chap. 1, pp. 1-7.

Part I

LINEAR DIFFERENCE EQUATIONS

Chapter 2

Difference Equations: General Principles

2.1 Definitions

Given a function $y = f(t)$, its first difference is defined as the difference between the value of the function when the argument assumes the value $t+h$, $(h > 0)$, and the value of the function corresponding to the value t of the argument. In symbols, $\Delta y = f(t+h) - f(t)$. It should be noted that it is unimportant whether the values run forwards or backwards, namely we could as well define the first difference of the function as $\Delta y = f(t) - f(t-h)$.

Without loss of generality we can assume unit increments of the dependent variable, i.e. $\Delta y = f(t+1) - f(t)$, or $\Delta y = f(t) - f(t-1)$. From now on we shall conventionally use forward-running values of t.

If we consider successive equally-spaced values of the independent variable $(t+1, t+2, t+3$, etc.), we can obtain successive first differences:

$$\begin{aligned}
\Delta y_t &= f(t+1) - f(t) &&= y_{t+1} - y_t,\\
\Delta y_{t+1} &= f(t+2) - f(t+1) &&= y_{t+2} - y_{t+1},\\
\Delta y_{t+2} &= f(t+3) - f(t+2) &&= y_{t+3} - y_{t+2},
\end{aligned}$$

and so on. We can then compute the *second differences*, i.e. the sequence of differences between two successive first differences:

$$\begin{aligned}
\Delta^2 y_t &= \Delta y_{t+1} - \Delta y_t = (y_{t+2} - y_{t+1}) - (y_{t+1} - y_t) = y_{t+2} - 2y_{t+1} + y_t,\\
\Delta^2 y_{t+1} &= \Delta y_{t+2} - \Delta y_{t+1} = y_{t+3} - 2y_{t+2} + y_{t+1},\\
\Delta^2 y_{t+2} &= \Delta y_{t+3} - \Delta y_{t+2} = y_{t+4} - 2y_{t+3} + y_{t+2},
\end{aligned}$$

and so on. Note that the superscript 2 means that the operation of computing the difference has been repeated twice, i.e. that the difference operator Δ has been applied twice.

Proceeding similarly we can compute the differences between two successive second differences and obtain the *third differences* of the function:

$$\begin{aligned}\Delta^3 y_t &= \Delta^2 y_{t+1} - \Delta^2 y_t = (\Delta y_{t+2} - \Delta y_{t+1}) - (\Delta y_{t+1} - \Delta y_t)\\ &= \Delta y_{t+2} - 2\Delta y_{t+1} + \Delta y_t\\ &= (y_{t+3} - y_{t+2}) - 2(y_{t+2} - y_{t+1}) + (y_{t+1} - y_t)\\ &= y_{t+3} - 3y_{t+2} + 3y_{t+1} - y_t,\end{aligned}$$

$$\Delta^3 y_{t+1} = \Delta^2 y_{t+2} - \Delta^2 y_{t+1} = y_{t+4} - 3y_{t+3} + 3y_{t+2} - y_{t+1},$$

and so on. Higher-order differences can be computed by the reader as an exercise.

We can now define an *ordinary difference equation* as a functional equation involving one or more of the differences Δy, $\Delta^2 y$, etc., of an unknown function of time. Since the argument t varies in a discontinuous way, taking on equally-spaced values, it follows that our unknown function will be defined only corresponding to these values of t (i.e. the graph of the function will be a succession of separate points, as we shall see in detail in Chap. 3).

We have called this equation *ordinary* because the unknown function is a function of only one argument. When the partial differences of a function having more than one argument are involved, the equation becomes a partial difference equation, a type of difference equation that will not be treated in this book.

The *order* of a difference equation is that of the highest difference appearing in the equation. If, for example, the highest difference contained is the third difference, the equation is of the third order; note that the equation is of the third order independently of the fact that the lower-order differences are or are not contained in the equation.

Since the differences of any order can be expressed, as we have seen above, in terms of values of the function at different points of time, a difference equation may also be defined as a functional equation involving two or more of the values y_t, y_{t+1}, etc., of an unknown function of time. For example, the difference equation $a\Delta y_t + by_t = 0$ transforms, if we substitute $\Delta y_t = y_{t+1} - y_t$, into $ay_{t+1} + (b-a)y_t = 0$. In this form, the order of the equation is given by the highest difference between time subscripts: if the equation, for example, contains y_{t+3}, y_{t+1} and y_t, it is of the third order. We shall consider the difference equation expressed in this second form as it is the form they commonly take in economic models.

Let us note again that it makes no difference whether the equally spaced values of t are computed forwards or backwards, so long as the structure of the time lags remains unaltered. The equation $ay_{t+1} + (b-a)y_t = 0$, for example, is identical with the equation $ay_t + (b-a)y_{t-1} = 0$. The reason is that to solve the difference equation means, as we know from the Introduction, to find a function (or functions) which satisfies (satisfy) the equation for any admissible value of t. This allows us to shift all the time subscript as we

like, provided that they are all shifted by the same amount (neglecting this proviso would alter the structure of the equation).

Consider now the equation $\Delta y_t = a$, i.e. $y_{t+1} - y_t = a$. In words, the problem is: find a function such that its first difference equals the given constant a for any value of t. It can be checked that the linear function $y = at + b$ satisfies the equation, since

$$y_{t+1} - y_t = [a(t+1) + b] - (at + b) = a.$$ [1]

Note that in the solution function an arbitrary constant (b) appears. This is not surprising, since the constancy of first differences is not affected by a parallel shift of the straight line. More generally, in the operation of differencing, the presence of an arbitrary constant, that is eliminated in the course of the operation, does not alter the result. Therefore, an arbitrary constant always appears in the solution of a first-order difference equation, and no more than one can appear.

Proceeding further, consider the equation $\Delta^2 y_t = 0$ (find a function such that its second difference equals zero for any value t). The solution is always the linear function $y = at + b$, but now both a and b are arbitrary constants; in fact, *any* straight line has a zero second difference. In general, the computation of second difference eliminates in succession *two* (and only two) arbitrary constants.

We shall see later on how the arbitrary constant(s) can be determined through additional conditions; what interests us here is to note that we can induce, from the reasoning above, the following important theorem:

Theorem 2.1 *The general solution of a difference equation of order* n *is a function of* t *involving exactly* n *arbitrary constants.*

2.2 Linear difference equations with constant coefficients

We can now summarize the scope of our treatment. In Part I we shall be concerned with linear, constant-coefficient difference equations. The general n-th order form of such equations is

$$c_n y_{t+n} + c_{n-1} y_{t+n-1} + \ldots + c_1 y_{t+1} + c_0 y_t = g(t), \tag{2.1}$$

where the c's are given constant and $g(t)$ is a known function. Some c's may be zero, but of course *both* c_n and c_0 must be different from zero if the

[1] Actually, this function is also the only one that satisfies the equation. This is shown by the 'existence and uniqueness' theorem, which we shall not treat. All types of equations considered in this book are 'well-behaved', i.e. their solution exists and is unique.

equation is of order n. Eq. (2.1) is called the forward form; the equivalent backward form is

$$c_n y_t + c_{n-1}y_{t-1} + ... + c_1 y_{t-n+1} + c_0 y_{t-n} = g(t). \tag{2.2}$$

In order to avoid cumbersome sentences, from now on we shall use the expression 'difference equations' (or even, where there is no danger of misunderstanding, simply 'equations') in the sense of 'ordinary difference equations, linear and with constant coefficients'.

We must now distinguish between homogeneous and non-homogeneous equations. Eq. (2.1) is non-homogeneous; the corresponding n-th order homogeneous equation is

$$c_n y_{t+n} + c_{n-1}y_{t+n-1} + ... + c_1 y_{t+1} + c_0 y_t = 0. \tag{2.3}$$

The reason for dealing with these two forms separately is that the solution of Eq. (2.1) can be obtained in a relatively simple manner when the solution of Eq. (2.3) is known.

2.2.1 The homogeneous equation

The following theorems are fundamental in the theory of homogeneous difference equations:

Theorem 2.2 *If $y_1(t)$ is a solution of the homogeneous equation, then $Ay_1(t)$, where A is an arbitrary constant, is also a solution.*

The proof is simple. Assume that $y_1(t)$ satisfies Eq. (2.3). Substitute $Ay_1(t)$ in the same equation, obtaining

$$c_n Ay_1(t+n) + c_{n-1}Ay_1(t+n-1) + ... + c_1 Ay_1(t+1) + c_0 Ay_1(t) = 0;$$

therefore

$$A[c_n y_1(t+n) + c_{n-1}y_1(t+n-1) + ... + c_1 y_1(t+1) + c_0 y_1(t)] = 0.$$

If $Ay_1(t)$ has to be a solution the last relation must be satisfied. Since $y_1(t)$ is a solution of Eq. (2.3), the expression in square brackets vanishes, and so the relationship

$$A[c_n y_1(t+n) + c_{n-1}y_1(t+n-1) + ... + c_1 y_1(t+1) + c_0 y_1(t)] = 0$$

is satisfied. This proves the theorem.

Before going on to the next theorem, it is as well to recall the notion of *linearly independent* functions.

Given n functions $y_1(t), y_2(t), ..., y_n(t)$, they are said to be *linearly dependent* if n constants $A_1, A_2, ..., A_n$ exist, which do not all vanish, and such that the equation

$$A_1 y_1(t) + A_2 y_2(t) + ... + A_m y_m(t) = 0 \tag{2.4}$$

is identically satisfied for all admissible values of t. Otherwise the functions are *linearly independent.*

Theorem 2.3 *If $y_1(t), y_2(t)$ are two distinct (i.e., linearly independent) solutions of the homogeneous equation ($n>1$), then $A_1y_1(t) + A_2y_2(t)$ is also a solution for any two constants A_1, A_2.*

The proof is similar to that of Theorem 2.2 and is left as an exercise.

Theorem 2.3—called the *superposition theorem*—can easily be extended to any number $k \le n$ of distinct solutions of Eq. (2.3), and gives us the procedure to obtain the general solution of Eq. (2.3). This procedure consists in finding n distinct solutions $y_1(t), y_2(t), ..., y_n(t)$ and combining them linearly, as stated in Theorem 2.4:

Theorem 2.4 *The general solution of Eq. (2.3) is given by*

$$f(t; A_1, A_2, ..., A_n) = A_1y_1(t) + A_2y_2(t) + ... + A_ny_n(t), \tag{2.5}$$

where $y_1(t), y_2(t), ..., y_n(t)$ are n *linearly independent solutions of Eq. (2.3), and $A_1, A_2, ..., A_n$ are arbitrary constants.*

The proof is straightforward: by Theorem 2.3, the function (2.5) is a solution of the difference equation (2.3). Since this function contains exactly n arbitrary constants, we can conclude—from Theorem 2.1—that it is the general solution of Eq. (2.3). The practical problem of how to find the n functions $y_1(t)$, $y_2(t)$, ..., $y_n(t)$ will be tackled in the following chapters; for the moment we observe that, given a homogeneous equation of order n, a set of n linearly independent solutions is called a *fundamental set.* The condition for a set of n solutions to form a fundamental set is contained in Theorem 2.5.

Theorem 2.5 *Let $y_1(t), y_2(t), ..., y_n(t)$ be* n *solutions of Eq. (2.3). They are linearly independent (i.e. form a fundamental set) if, and only if, the following determinant* (called the Casorati determinant)

$$D(t) = \begin{vmatrix} y_1(t) & y_2(t) & ... & y_n(t) \\ y_1(t+1) & y_2(t+1) & ... & y_n(t+1) \\ ... & ... & ... & ... \\ y_1(t+n-1) & y_2(t+n-1) & ... & y_n(t+n-1) \end{vmatrix} \tag{2.6}$$

is different from zero for all admissible values of t.

To prove this theorem, consider Eq. (2.4): since the functions we are considering are solutions to Eq. (2.3), they hold for any t by definition of solution, and hence also for $t+1, ..., t+n-1$. Hence we can write the following linear system

$$\begin{array}{lllll}
A_1y_1(t) & +A_2y_2(t) & +...+ & A_ny_n(t) & =0, \\
A_1y_1(t+1) & +A_2y_2(t+1) & +...+ & A_ny_n(t+1) & =0, \\
... & ... & ... & ... & ... \\
A_1y_1(t+n-1) & +A_2y_2(t+n-1) & +...+ & A_ny_n(t+n-1) & =0,
\end{array} \tag{2.7}$$

which is a system of homogeneous linear equations. According to a well-known theorem in elementary algebra, when $D(t) \neq 0$, system (2.7) admits only the null solution, i.e. holds true if, and only if, $A_1 = A_2 = ... = A_n = 0$, which is the definition of linearly independent functions (see above). On the contrary, $D(t) = 0$ is the necessary and sufficient condition for system (2.7) to possess non-trivial solutions (i.e., solutions with at least one non-zero A_i, $i = 1, 2, ..., n$), which is the definition of linearly dependent functions.

2.2.2 The non-homogeneous equation

We have so far dealt with the homogeneous difference equation. We now prove the basic theorem concerning the solution of the non-homogeneous difference equation.

Theorem 2.6 *If $\bar{y}(t)$ is any particular solution of the non-homogeneous equation [i.e., $\bar{y}(t)$ is any function that satisfies (2.1)], the general solution of the same equation is obtained adding $\bar{y}(t)$ to the general solution of the corresponding homogeneous equation, namely*

$$y(t) = \bar{y}(t) + f(t; A_1, A_2, ..., A_n) \tag{2.8}$$

is the general solution of the non-homogeneous equation.

The proof of the theorem can be given substituting (2.8) into (2.1) and checking that the latter is satisfied. Since the function (2.8) contains exactly n arbitrary constants, it is the general solution of Eq. (2.1).

The general solution of the homogeneous equation is thus only a part of the general solution of the non-homogeneous equation, and so it is not 'general' with respect to the latter. This means that the expression 'general solution' must always be qualified. As a matter of terminology, note the following: (1) some authors use the world 'integral' (particular or general) instead of 'solution' but with the same meaning; (2) the expression 'particular solution' is also used (a) in the sense of a solution obtained from the general solution by giving specific values to the arbitrary constants, and (b) in the sense of any single non-general solution of the homogeneous equation (i.e., to indicate any one of $y_1(t), y_2(t)$, etc.); (3) the expression 'complementary function' is used to indicate the general solution of the homogeneous equation when considered as part of the general solution of the non-homogeneous equation, and the expression 'reduced equation' is used to indicate the homogeneous part of a non-homogeneous equation, i.e. the corresponding homogeneous equation obtained putting $g(t) \equiv 0$ in the course of the procedure to solve a non-homogeneous equation. To avoid confusion, we shall not adopt these uses.

Theorem 2.6 contains the method to follow for solving the non-homogeneous equation:

(a) find a particular solution $\bar{y}(t)$ of the non-homogeneous equation;

(b) put $g(t) \equiv 0$ and solve the resulting homogeneous equation (often called the 'reduced' equation);

(c) add the two results.

Steps (a) and (b) can be taken in any order; step (c) gives the general solution of the non-homogeneous equation.

The particular solution of the non-homogeneous equation will depend, *ceteris paribus*, on the form of the known function $g(t)$. This suggest the following general approach: *to find a particular solution of the non-homogeneous equation, try a function having the same form of $g(t)$ but with undetermined constant(s) (e.g., if $g(t)$ is a constant, try an undetermined constant; if it is an exponential function, try the same exponential function with an undetermined multiplicative constant, and so on). Substitute this function in the non-homogeneous equation and determine the coefficient(s) so that the equation is satisfied.*

This method—called method of *undetermined coefficients*—will be expounded in more detail in the following chapter, where we shall also examine the cases in which it cannot be applied.

It is interesting to note, from the economic point of view, that in the general solution of the non-homogeneous equation the particular solution $\bar{y}(t)$ may usually be interpreted as the *equilibrium state* of the variable y (a stationary equilibrium or a moving equilibrium according to whether $\bar{y}(t)$ is a constant or a function of t). The component $f(t; A_1, A_2, ..., A_n)$ in Eq. (2.8) may then be interpreted as giving the *deviations* from the equilibrium. Of course, from the mathematical point of view it is always true that $y(t) - \bar{y}(t) = f(t; A_1, A_2, ..., A_n)$, independently of the possibility of giving an economic interpretation to the particular solution $\bar{y}(t)$.

2.3 Determination of the arbitrary constants

The problem remains of how to determine the arbitrary constants A_i. To do this we need an adequate number of additional conditions. This need derives from the fact that the solution—namely Eq. (2.5) or Eq. (2.8) as the case may be—of the difference equation under consideration gives only the *form* of the function $y(t)$ but not its position in the Cartesian plane (t, y). However, as soon as the function is constrained to pass through n given points, its position—which depends on n arbitrary constants—is determined, and the arbitrariness of the constants disappears.

More formally, to determine the n arbitrary constants, n additional conditions are needed, which usually take the form

$$\begin{array}{llll} y(t) & = y_0 & \text{for } t & = 0, \\ y(t) & = y_1 & \text{for } t & = 1, \\ \ldots & \ldots & \ldots \quad \ldots & \ldots \\ y(t) & = y_{n-1} & \text{for } t & = n-1, \end{array}$$

where $y_0, y_1, \ldots\ldots. y_{n-1}$ are known values (whence the name of *initial conditions*). Substituting such values in the general solution, we obtain a system of n linear equations in the n unknowns $A_1, A_2, \ldots\ldots., A_n$. Consider for example the general solution of Eq. (2.1)

$$y(t) = A_1 y_1(t) + A_2 y_2(t) + \ldots\ldots. + A_n y_n(t) + \bar{y}(t),$$

where $y_1(t), y_2(t), y_n(t)$ are n distinct solutions of the corresponding homogeneous equation. Substituting the given values y_0 etc. in the place of $y(0)$ etc. we obtain, after rearranging terms,

$$\begin{array}{ll} A_1 y_1(0) + A_2 y_2(0) + \ldots + A_n y_n(0) & = y_0 - \bar{y}(0), \\ A_1 y_1(1) + A_2 y_2(1) + \ldots + A_n y_n(1) & = y_1 - \bar{y}(1), \\ \ldots & \ldots \\ A_1 y_1(n-1) + A_2 y_2(n-1) + \ldots + A_n y_n(n-1) & = y_{n-1} - \bar{y}(n-1), \end{array} \tag{2.9}$$

where of course $\bar{y}(0), \bar{y}(1), \ldots\ldots\ \bar{y}(n-1)$ are absent if we consider the solution of the homogeneous equation (2.3). System (2.9) is a linear system whose determinant is

$$D(0) = \begin{vmatrix} y_1(0) & y_2(0) & \ldots & y_n(0) \\ y_1(1) & y_2(1) & \ldots & y_n(1) \\ \ldots & \ldots & \ldots & \ldots \\ y_1(n-1) & y_2(n-1) & \ldots & y_n(n-1) \end{vmatrix}.$$

It is easy to see that $D(0)$ coincides with the determinant $D(t)$—as defined in Eq. (2.6)—for $t = 0$. Since the functions $y_1(t), y_2(t), \ldots, y_n(t)$ form a fundamental set, $D(t)$ is different from zero for any t, and so also for $t = 0$. It follows that $D(0) \neq 0$. Thus system (2.9) can always be solved.

We now have enough general principles to pass on to a detailed treatment of the difference equations of the various orders.

2.4 References

Boole, G., 1960 (1872), *A Treatise on the Calculus of Finite Differences*, Chap. IX, pp. 157-161.

Goldberg, S., 1958, *Introduction to Difference Equations*, Chaps. 1, 2, 3.
Kenkel, J.L., 1974, *Dynamic Linear Economic Models,* Chaps. 3, 5.
Milne-Thomson, L.M., 1960, *The Calculus of Finite Differences,* Chap. XII.

Chapter 3

First-order Difference Equations

The general form of these equations is

$$c_1 y_t + c_0 y_{t-1} = g(t), \tag{3.1}$$

where c_0, c_1 are given constants and $g(t)$ is a known function. The constants c_0, c_1 must be both different from zero, since if even only one of them is zero the equation is no longer a difference equation.

3.1 Solution of the homogeneous equation

Let us begin with the study of the homogeneous equation, whose form is

$$c_1 y_t + c_0 y_{t-1} = 0, \tag{3.2}$$

or

$$y_t + b y_{t-1} = 0, \tag{3.3}$$

where $b \equiv c_0/c_1$. Suppose that in the initial period (i.e. for $t = 0$) the function y takes on an arbitrary value A; from Eq. (3.3) we can then compute the following sequence:

$$\begin{aligned}
y_1 &= -by_0 = -bA,\\
y_2 &= -by_1 = -b(-bA) = b^2A,\\
y_3 &= -by_2 = -b(b^2A) = -b^3A,\\
y_4 &= -by_3 = -b(-b^3A) = b^4A,\\
&\cdots\cdots\cdots\cdots\cdots\cdots
\end{aligned}$$

and so the solution appears to be

$$y_t = A(-b)^t. \tag{3.4}$$

As a check, substitute this function in Eq. (3.3) :

$$A(-b)^t + bA(-b)^{t-1} = 0. \tag{3.5}$$

If our function is a solution, Eq. (3.5) must hold identically. Now, since

$$bA(-b)^{t-1} = -(-b)A(-b)^{t-1} = -A(-b)^t,$$

it follows that Eq. (3.5) can be written as

$$A(-b)^t - A(-b)^t = 0, \tag{3.6}$$

and is indeed satisfied for any value of t.

Since the function we have found satisfies the difference equation and contains one arbitrary constant, we may conclude from general principles that it is the general solution.

The problem remains of how to determine the arbitrary constant. To do this we need an additional condition. As we know (see Sect 2.3), this need derives from the fact that relation (3.4) gives only the *form* of the function y_t but not its position in the Cartesian plane (t, y_t). As soon as the function is constrained to pass through a given point, say (t^*, y^*), its position, which depends on one arbitrary constant only, is determined and the arbitrariness of the constant disappears. More formally, the additional condition says that $y_t = y^*$ for $t = t^*$, where t^* and y^* are known values. Substituting these values in (3.4) we get $y^* = A(-b)^{t^*}$ and so

$$A = y^*/(-b)^{t^*}. \tag{3.7}$$

In economic problems the value of y in the initial period is usually assumed as known, at least in principle, i.e. $y_t = y_0$ for $t = 0$, which gives $A = y_0$. In this case, we speak of the *initial condition.* In problems involving a given time horizon (say, up to $t = T$) it is also possible to take as given the value of y in the final period, i.e. $y_t = y_T$ for $t = T$; in such a case we speak of a *terminal condition.*

The behaviour over time of the function $y_t = A(-b)^t$ depends on both the sign and the absolute value of the parameter b.

As for the sign, if b is negative then $-b$ is positive and the movement is monotonic. On the other hand, if b is positive then $-b$ is negative and the value of the function will alternate in sign, since the power of a negative number is positive (negative) if the exponent is an even (odd) integer. This case is usually described as an 'oscillatory' movement. However, to distinguish terminologically this kind of movement from the trigonometric (sine and cosine) oscillations (which, we shall see, can arise only in second- or higher-order equations), we suggest the expression 'improper oscillations' or 'alternations'. 'Proper oscillations' or simply 'oscillations' would then specifically indicate trigonometric oscillations.

As for the absolute value, if b is in absolute value less (greater) than unity, the movement will be convergent (divergent). This conclusion is a consequence of the properties of powers: the absolute value of a power, as

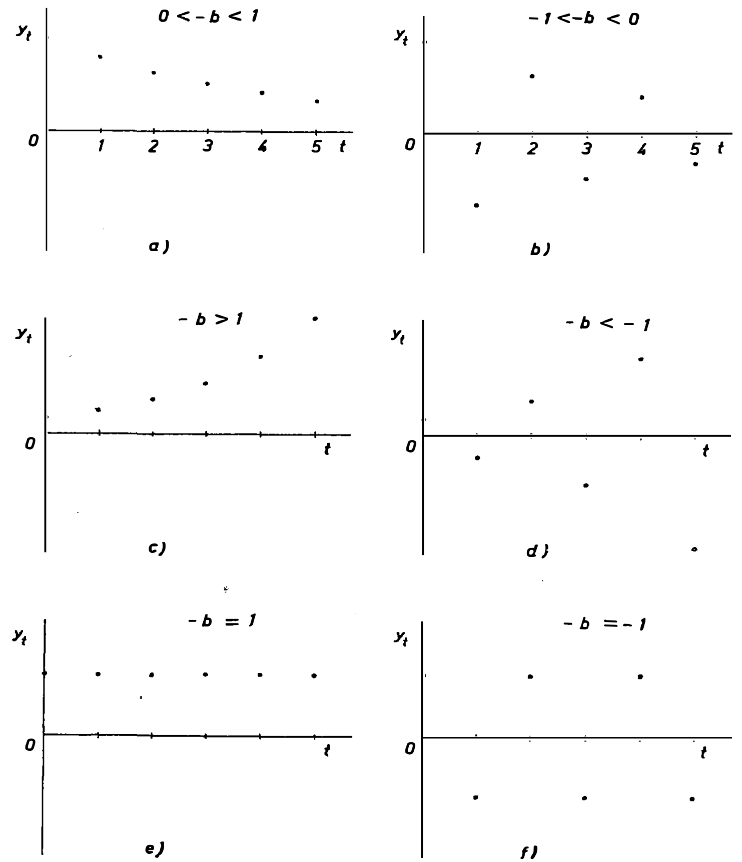


Figure 3.1: (a) Monotonic and convergent; (b) Oscillatory and convergent; (c) Monotonic and divergent; (d) Oscillatory and divergent; (e) Constant; (f) Oscillatory with constant amplitude.

the exponent increases, tends to zero (to infinity) if the absolute value of the base is less (greater) than one. In the particular case of $|b| = 1$, the function shows improper oscillations of constant amplitude (when $b = 1$) or takes the constant value A (when $b = -1$).

In Fig. 3.1 all kinds of movements are shown (A is assumed to be positive; if it were negative, the qualitative behaviour of the solution would not change). Note that the diagrams show only a succession of points. This is because, as we know, t varies over a set of equally spaced values $(0, 1, 2, 3,$ etc.) and so the solution function is defined only corresponding to equally spaced values of t. The graphical counterpart to this is a succession of points.

Of course, in reality time is a continuous variable. When we formalize an economic problem in difference equations terms (this is also called 'period analysis'), we (implicitly or explicitly) assume that, for all relevant purposes, only what happens at the end of each time interval does matter, so that the variables we are analysing may be thought of as varying by discrete 'jumps'. What happens during the period is not considered, in the sense that all relevant economic activity of each period is assumed to be concentrated in a single point of time (the end of the period, which is the same as the beginning of the following period). These assumptions may or may not be justified according to the nature of the problem we are examining; for some further comments on these points, as well as on the related point of the use of discrete or continuous time tools in economics, see Chap 27, Sect. 27.2.

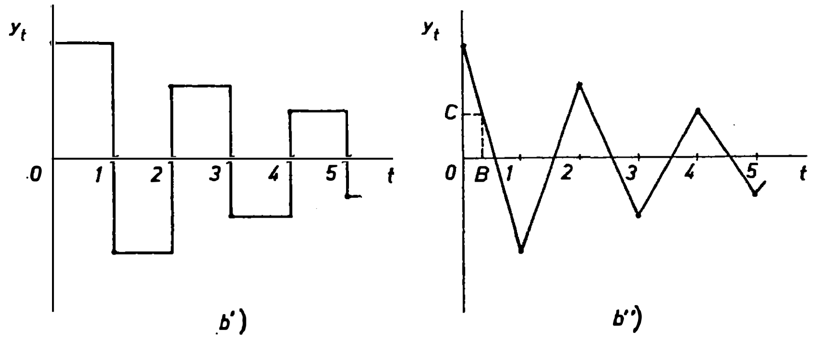


Figure 3.2: Joining the points of the solution to a difference equation

Going back to the diagrams, the points are usually joined by segments. Fig. 3.2 shows two alternative ways of doing this (diagram (b) of Fig. 3.1 is exemplified); the right-hand panel is usually called a 'sawtooth' diagram. It must be emphasized that *the joining of the successive points is performed only to help the eye to follow the movement of the solution over time.* It would be a gross mistake to interpret the segments as describing the movement of y_t in each instant of the period: it is *not* possible to say, for example, that for $t = \overline{OB}$ the value of y is $\overline{OC}$. Such an inference would be wrong, since y_t is defined only for $t = 0, 1, 2, 3, \ldots$, as represented in Fig. 3.1. If that is understood, graphical representations of the kind depicted in Fig. 3.2 may

be safely adopted as a visual aid.

3.2 Particular solution of the non–homogeneous equation

To complete the study of first-order equations we must now explain how to find a particular solution of the non-homogeneous equation. The application of the method of *undetermined coefficients* set forth in Chap. 2, Sect. 2.2.2, will be explained in relation to the most common functions.

In Sect. 3.2.6 we shall introduce a different and more general method, the *operational method,* to deal with the cases in which the method of undetermined coefficients cannot be applied.

3.2.1 $g(t)$ is a constant

In this case, equation (3.1) becomes

$$c_1 y_t + c_0 y_{t-1} = a, \tag{3.8}$$

where a is a given constant. As a particular solution, try an undetermined constant, and call it μ. Substitution into (3.8) yields

$$(c_1 + c_0)\mu = a,$$

from which

$$\mu = a/(c_1 + c_0) \tag{3.9}$$

and so

$$\bar{y}_t = a/(c_1 + c_0) \tag{3.10}$$

is a particular solution.

The method obviously breaks down if $c_1 + c_0 = 0$. In this case, equation (3.8) may be written as

$$y_t - y_{t-1} = a/c, \tag{3.11}$$

where $c = c_1 = -c_0$.

As a particular solution, try now μt. Substituting in (3.11) we have

$$\mu t - \mu(t-1) = a/c,$$

whence

$$\mu = a/c. \tag{3.12}$$

A particular solution is then

$$\bar{y}_t = (a/c)t. \tag{3.13}$$

It is important to note that the above treatment illustrates the following general prescription (which is a necessary complement of the general principle explained in Chap. 2): *if the function you try as a particular solution does not work, try the same function multiplied by t.*[1]

Since the general solution of the homogeneous equation $y_t - y_{t-1} = 0$ is $y_t = A$, where A is an arbitrary constant, the general solution of Eq. (3.11) is $y_t = A + (a/c)t$. The same result could be obtained directly, seeing that the function whose first difference is constant is a linear function (see Chap. 2, for a similar case), but it is important to understand how the same result is obtained by applying general principles.

3.2.2 $g(t)$ is an exponential function

When $g(t) = Bd^t$, where B and d are given constants,[2] as a particular solution try Cd^t, C being an undetermined constant. Substituting in (3.1) we have

$$c_1Cd^t + c_0Cd^{t-1} = Bd^t. \tag{3.14}$$

Therefore

$$d^{t-1}(c_1Cd + c_0C - Bd) = 0. \tag{3.15}$$

The last equation is satisfied for any t if, and only if,

$$c_1Cd + c_0C - Bd = 0, \tag{3.16}$$

whence

$$C = \frac{Bd}{c_1d + c_0}. \tag{3.17}$$

A particular solution is then

$$\bar{y} = \frac{Bd}{c_1d + c_0}d^t. \tag{3.18}$$

The method fails if $c_1d + c_0 = 0$; note that this implies that the general solution of the corresponding homogeneous equation is Ad^t. As a particular solution, try then tCd^t. Substitute in (3.1) and obtain

$$c_1tCd^t + c_0(t-1)Cd^{t-1} = Bd^t; \tag{3.19}$$

thus

[1]In second- or higher-order equations it may be necessary, as we shall see, to multiply by t^2, t^3, etcetera. More generally, the prescription is to try the same function multiplied by a suitable polynomial in t.

[2]If $g(t) = B\alpha^{\lambda t}$, where B, α, λ are given constants, put $\alpha^\lambda = d$ and proceed as before.

$$d^{t-1}[(c_1d + c_0)tC - c_0C - Bd] = 0. \tag{3.20}$$

Since $c_1d + c_0 = 0$ by assumption, it follows that

$$d^{t-1}(-c_0C - Bd) = 0, \tag{3.21}$$

which is satisfied for any t if, and only if,

$$-c_0C - Bd = 0, \tag{3.22}$$

so that

$$C = -Bd/c_0. \tag{3.23}$$

A particular solution is then

$$\bar{y}_t = \frac{-Bd}{c_0}td^t. \tag{3.24}$$

3.2.3 $g(t)$ is a polynomial function of degree m

Consider for example $g(t) = a_0 + a_1t$, where a_0 and a_1 are given constants. Try $\bar{y}_t = \alpha + \beta t$ as a particular solution, α and β being undetermined constants. Substituting in (3.1) we have

$$c_1(\alpha + \beta t) + c_0[\alpha + \beta(t-1)] = a_0 + a_1t. \tag{3.25}$$

Separating the terms containing t from the others, we get

$$(c_1\beta + c_0\beta - a_1)t = a_0 - (c_0 + c_1)\alpha + c_0\beta. \tag{3.26}$$

The only way to satisfy Eq. (3.26) for any t is to let both $c_1\beta + c_0\beta - a_1$ and $a_0 - (c_0 + c_1)\alpha + c_0\beta$ be equal to zero, and so we have the following linear system:

$$\begin{aligned} (c_0 + c_1)\beta &= a_1, \\ (c_0 + c_1)\alpha - c_0\beta &= a_0, \end{aligned} \tag{3.27}$$

whose solution determines the values of α and β.

The student should be able, by now, to examine by himself the case where the method fails (i.e. when $c_1 + c_0 = 0$), as well as the case where the degree of the polynomial is higher than one.

3.2.4 $g(t)$ is a trigonometric function of the sine-cosine type

In this case, $g(t) = B_1 \cos \omega t + B_2 \sin \omega t$, where both B_1, B_2, ω are known constants. As a particular solution, try the function

$$\alpha \cos \omega t + \beta \sin \omega t, \tag{3.28}$$

where α and β are undetermined constants. Note that the trial function must be $\alpha \cos \omega t + \beta \sin \omega t$ also when $g(t) = B_1 \cos \omega t$ or $g(t) = B_2 \sin \omega t$. Substitution in (3.1) gives

$$\begin{aligned} & c_1 \alpha \cos \omega t + c_1 \beta \sin \omega t + c_0 \alpha \cos(\omega t - \omega) + c_0 \beta \sin(\omega t - \omega) \\ & = B_1 \cos \omega t + B_2 \sin \omega t. \end{aligned} \tag{3.29}$$

We now recall from elementary trigonometry the addition formulae for the sine and cosine, namely

$$\begin{aligned} \cos(\omega t \pm \omega) &= \cos \omega t \cos \omega \mp \sin \omega t \sin \omega, \\ \sin(\omega t \pm \omega) &= \sin \omega t \cos \omega \pm \sin \omega \cos \omega t. \end{aligned}$$

By using these formulae Eqs. (3.29) become, after simple manipulations,

$$\begin{aligned} & [(c_1 + c_0 \cos \omega)\alpha - c_0 \beta \sin \omega - B_1] \cos \omega t \\ & + [\alpha c_0 \sin \omega + (c_1 + c_0 \cos \omega)\beta - B_2] \sin \omega t = 0. \end{aligned} \tag{3.30}$$

This equation is satisfied for any t if, and only if,

$$\begin{aligned} (c_1 + c_0 \cos \omega)\alpha - (c_0 \sin \omega)\beta - B_1 &= 0, \\ (c_0 \sin \omega)\alpha + (c_1 + c_0 \cos \omega)\beta - B_2 &= 0, \end{aligned} \tag{3.31}$$

which is a linear system whose solution determines the values of α and β. The student may examine, as an exercise, the case in which the method fails.

3.2.5 $g(t)$ is a combination of the previous functions

When $g(t)$ is a combination of two or more of the functions treated above in the various cases, as a particular solution we can try the same combination of functions with undetermined coefficients. This is then substituted in Eq. (3.1), and the resulting expression—after collecting the terms containing t separately from the others if the case—usually gives the equations (in terms of the undetermined coefficients) that must hold for Eq. (3.1) to be identically satisfied. The solution of these equations gives the sought-for values of the undetermined coefficients.

Although the cases exemplified above cover the usual situations, there may be the case in which $g(t)$ does not belong to any of the categories examined above. Worse still, we might find ourselves in the case in which $g(t)$ is generic function of time, whose functional form we do not actually know.

3.2.6 The case when $g(t)$ is a generic function of time. Backward and forward solutions

It may happen that we do not know the functional form of $g(t)$, although we know the actual succession of values that such a function takes through

time. In other words, $g(t)$ is simply a sequence of known real values. A typical case is, for example, when (in the econometrician's terminology), y_t is an endogenous (or output) variable while $g(t)$ represents an exogenous (or input) variable. Or we might have a difference equation that associates the values of an unknown and unobservable variable at different points of time (for example, changes in price expectations) with the observed values of a related variable (the actual prices: for an example see below Sect. 4.1.1).

In such cases the method of undetermined coefficients cannot be applied, as it requires the knowledge of the functional form of $g(t)$. It is, however, possible to find a particular solution by applying operational methods, which were introduced in the study of the solution of difference equations by Boole (1872; see also Milne-Thomson, 1960, Chap. XIII). We have already met the difference operator Δ in Sect. 2.1; we now introduce the lag operator L, i.e. the operator such that

$$Ly_t = y_{t-1}, \quad L^n y_t = y_{t-n}, \quad L^{-1} y_t = y_{t+1}. \tag{3.32}$$

For most practical purposes, operators can be treated just as algebraic quantities and manipulated with the rules valid for these quantities (for an elementary introduction see Allen, 1960, Appendix A). For example, using a well-known infinite-series expansion, we have

$$(1 - \alpha L)^{-1} = 1 + \alpha L + \alpha^2 L^2 + \ldots = \sum_{i=0}^{\infty} \alpha^i L^i. \tag{3.33}$$

A heuristic justification of (3.33) consists of multiplying through by $(1 - \alpha L)$, whence

$$\begin{aligned} (1 - \alpha L)^{-1}(1 - \alpha L) &= 1 \\ &= (1 + \alpha L + \alpha^2 L^2 + \ldots)(1 - \alpha L) \\ &= (1 + \alpha L + \alpha^2 L^2 + \ldots) - (\alpha L + \alpha^2 L^2 + \ldots). \end{aligned} \tag{3.34}$$

The last expression on the r.h.s. of Eq. (4.34) is indeed equal to 1, since all the other terms cancel out.

Let us now consider the difference equation (3.1) and let us suppose that $g(t) = x_t$, where x_t is a sequence of known real values. Dividing through by c_1 and letting $X_t \equiv (1/c_1)x_t$, we have

$$y_t + by_{t-1} = X_t, \tag{3.35}$$

that is

$$\begin{aligned} (1 + bL)y_t &= X_t, \\ y_t &= (1 + bL)^{-1} X_t. \end{aligned} \tag{3.36}$$

If we let $\alpha = -b$ and apply the expansion (3.33) we find that

$$\overline{y}_t = \sum_{i=0}^{\infty}(-b)^i L^i X_t = \sum_{i=0}^{\infty}(-b)^i X_{t-i} \tag{3.37}$$

satisfies our non-homogeneous difference equation, hence it is a particular solution. The reader may wish to check as an exercise that if one plugs Eq. (3.37) into Eq. (3.35) the latter is indeed identically satisfied.

Note that $\overline{y}_t$ as given in Eq. (3.37) is a bounded sequence if $|b| < 1$, but will be divergent if $|b| > 1$. In this latter case consider an alternative expansion, which derives from the fact that, formally,

$$1 - \alpha L = -\alpha L\left(1 - \frac{1}{\alpha L}\right)$$

and

$$\left(1 - \frac{1}{\alpha L}\right)^{-1} = 1 + (\alpha L)^{-1} + (\alpha L)^{-2} + \ldots = 1 + \frac{1}{\alpha}L^{-1} + \left(\frac{1}{\alpha}\right)^2 L^{-2} + \ldots$$

Therefore,

$$\begin{aligned}(1-\alpha L)^{-1} &= -\frac{1}{\alpha L}\left(1 - \frac{1}{\alpha L}\right)^{-1} = \frac{-1}{\alpha}L^{-1}\left(1 + \frac{1}{\alpha}L^{-1} + \left(\frac{1}{\alpha}\right)^2 L^{-2} + \ldots\right) \\ &= \frac{-1}{\alpha}L^{-1} - \left(\frac{1}{\alpha}\right)^2 L^{-2} - \left(\frac{1}{\alpha}\right)^3 L^{-3} - \ldots\end{aligned}$$

and, finally,

$$(1-\alpha L)^{-1} = -\sum_{i=1}^{\infty}\left(\frac{1}{\alpha}\right)^i L^{-i} \tag{3.38}$$

is the alternative expansion we need. Hence, if we apply (3.38) instead of Eq. (3.33), we obtain the particular solution

$$\overline{y}_t = -\sum_{i=1}^{\infty}\left(-\frac{1}{b}\right)^i L^{-i} X_t = -\sum_{i=1}^{\infty}\left(-\frac{1}{b}\right)^i X_{t+i}, \tag{3.39}$$

which is a bounded sequence if $|1/b| < 1$.

Equation (3.37) shows the particular solution to be a geometrically declining weighted sum of all the *past* values of X_t, and for this reason is called the *backward* solution, while Eq.(3.39) shows such solution to be a geometrically declining weighted sum of all the *future* values of X_t, and for this reason is called the *forward* solution.

Since $|b| < 1$ ($|b| > 1$) corresponds to the case in which the solution of the homogeneous part of the equation is stable (unstable), it follows that the particular solution of the non-homogeneous equation will be of the backward or forward type according as the solution of the homogeneous equation is stable (unstable).

Let us conclude by observing that the operational method is quite general and could be applied to the functional forms of $g(t)$ described in the previous paragraphs, although at the cost of some complications. Consider, for example, the case in which $g(t) = a$, a constant. In the case $|b| < 1$ we can apply the form (3.37), with $X_t = a/c_1$. Thus, recalling that the infinite sum $\sum_{i=0}^{\infty}(-b)^i$ equals $(1+b)^{-1}$ when $|b| < 1$, we have

$$\overline{y}_t = \sum_{i=0}^{\infty}(-b)^i \frac{a}{c_1} = \frac{a}{c_1}\sum_{i=0}^{\infty}(-b)^i = \frac{a}{c_1}\frac{1}{1+b} = \frac{a}{c_1 + c_0},$$

which of course coincides with Eq. (3.10). When $|b| > 1$, we use Eq. (3.39) to obtain the same result.

3.3 General solution of the non-homogeneous equation

After finding a particular solution of the non-homogeneous equation, we can add it to the general solution of the corresponding homogeneous equation, thus obtaining the general solution of the non-homogeneous equation:

$$y_t = A(-b)^t + \overline{y}_t. \tag{3.40}$$

We must now determine the arbitrary constant A. It is important to keep in mind that this constant must be determined, given an additional condition, in relation to the general solution of the equation concerned. This means that, if the equation is non-homogeneous, we cannot use formula (3.7) but must find a new one. The method, however, is the same: given $y_t = y^*$ for $t = t^*$, substitute in the general solution (3.40). This yields

$$A = \frac{y^* - \overline{y}_{t^*}}{(-b)^{t^*}}. \tag{3.41}$$

In economics, as we have already said, the initial value of y is usually assumed known, at least in principle, so that $A = y_0 - \overline{y}_0$.

Let us finally recall (see Sect. 2.2.2), for use in the economic examples to be examined in the next chapter, that in the general solution of the non-homogeneous equation the particular solution $\overline{y}_t$ may usually be interpreted as the *equilibrium state* of the variable y (a stationary equilibrium or a moving equilibrium according to whether $\overline{y}_t$ is a constant or a function of t). The component $A(-b)^t$ in (3.40) may then be interpreted as giving the *deviations* from equilibrium, since $y_t - \overline{y}_t = A(-b)^t$.

3.4 A digression on distributed lags and partial adjustment equations

In economic models it is often assumed that the value that a variable y takes at time t depends on the present and past values of some other variable x (that for the moment we leave unspecified), namely

$$y_t = b_0 x_t + b_1 x_{t-1} + ... + b_n x_{t-n} + ... \tag{3.42}$$

where $b_0, b_1, ..., b_n, ...$ are known constants, and are usually assumed to be non negative. Equation (3.42) is called a *distributed lag* equation, and may contain a finite or infinite number of terms, but in any case the sum of the b's is assumed to be a finite number, which is taken to be positive. Thus we have

$$\sum_i b_i = b, \tag{3.43}$$

which is assumed to hold both when the number of terms is finite and when it is infinite.

There are several reasons for the presence of distributed lags in economic models. In many situations there is an inherent time lag between a decision made by an economic agent and the completion of the corresponding action, due for example to technical reasons. If a firm decides to carry out an investment project involving the construction of a new plant, some time will elapse before it is completed. Another reason may be habit persistence on the side of consumers. We also know that macroeconomic variables react more or less slowly to changes in policy instruments (money supply, government expenditure, etc.).

It is also frequently assumed that optimizing agents, after calculating the optimal value of their decision variable(s), cannot immediately adjust the actual to the desired (optimal) value of the variable due to frictions and imperfections of various types. A wealth holder, for example, after calculating the optimal wealth composition might be unable immediately to bring the actual to the desired composition because of the time required to sell and/or buy indivisible physical assets. In such cases we are in the presence of a *partial adjustment equation*, according to which, given a discrepancy between the desired and actual value of a variable, the latter is adjusted toward the former only gradually, according to a coefficient of reaction or speed of adjustment, namely

$$y_t - y_{t-1} = \alpha(x_t - y_{t-1}), \quad 0 < \alpha < 1, \tag{3.44}$$

where y_t denotes the actual value of the variable and x_t is its desired value, whose determinants need not concern us here. The adjustment speed α has been assumed smaller than one because $\alpha = 1$ means that the discrepancy between the actual and desired value of the variable is completely eliminated

within the period, i.e., we have a total rather than a partial adjustment. We shall presently see that distributed lag equations and partial adjustment equations are closely related.

If we go back to distributed lags and define a new set of coefficients

$$w_i = \frac{b_i}{b},$$

then Eq. (3.42) can be rewritten as

$$y_t = b[w_0 x_t + w_1 x_{t-1} + ... + w_n x_{t-n} + ...], \qquad \sum_{i=0}^{\infty} w_i = 1. \tag{3.45}$$

From Eq. (3.45) it can be seen that if each value of x had been increased by one unit, the value of y_t would increase by b units. This is the long-term effect of a sustained unit increase in x (i.e., all values of x increase by one unit and maintain the new value through time). Thus b is called the *long-term distributed lag multiplier.* The coefficient b_0 is called *impact multiplier,* and the coefficients b_i are called *delay-i multipliers.*

We can define a *mean time-lag* (or average lag) as the weighted arithmetic mean of the time lags, where the weights are the coefficients w_i. Since the lags are amounts of time, their mean is an amount of time as well. Formally we have

$$0 \cdot w_0 + 1 \cdot w_1 + ... + n \cdot w_n + ... = \sum_{i=0}^{\infty} i w_i \tag{3.46}$$

or, equivalently,

$$\sum_{i=0}^{\infty} i b_i / \sum_{i=0}^{\infty} b_i = \sum_{i=0}^{\infty} i b_i / b. \tag{3.47}$$

In practical application it is often assumed that the coefficients follow some kind of known distribution (for a survey of the various types of distribution see any econometrics textbook). A frequently used lag distribution is the Koyck distribution, whose coefficients decline geometrically, namely

$$b_i = k b_{i-1}, \quad i = 1, 2, ... \quad 0 < k < 1. \tag{3.48}$$

Equation (3.48) can be considered as a first-order difference equation, whose solution gives

$$b_i = b_0 k^i.$$

With these lag coefficients, the sum of the infinite geometric series of coefficients (3.43) becomes

$$b = \sum_{i=0}^{\infty} b_i = b_0 \frac{1}{1-k},$$

whence

$$b_0 = b(1-k), \tag{3.49}$$

and so the first coefficient of the series is fixed at $b(1-k)$ in order that the sum of the infinite geometric series of coefficients is b. Therefore the distributed lag equation (3.42) becomes

$$y_t = b(1-k)(x_t + kx_{t-1} + k^2 x_{t-2} + \dots). \tag{3.50}$$

The useful feature of the Koyck scheme is that it can be reduced to a relation involving y_t, y_{t-1}, and x_t only. In fact, if we shift all the time subscripts in Eq. (3.50) backwards by one unit and multiply throughout by k we have

$$ky_{t-1} = b(1-k)(kx_{t-1} + k^2 x_{t-2} + \dots). \tag{3.51}$$

Subtraction of Eq. (3.51) from Eq. (3.50) and elimination of all common terms yields

$$y_t - ky_{t-1} = b(1-k)x_t,$$

from which

$$y_t = b(1-k)x_t + ky_{t-1}. \tag{3.52}$$

Equation (3.52) can be put into an alternative form by subtracting y_{t-1} from both members and assuming that the coefficients of the distribution were normalized from the beginning, i.e. $b = \sum_i b_i = 1$. Thus we get

$$y_t - y_{t-1} = (1-k)(x_t - y_{t-1}), \tag{3.53}$$

which is the typical form of the partial adjustment equation (3.44), with $\alpha = 1-k$. The adjustment speed α is inversely related to the mean time-lag, which is $(1 - 1/\alpha)$ (see Sect. 3.5.2, exercise 3). When $\alpha \to 1$ (which means $k \to 0$) the mean time-lag $\to 0$, since the adjustment tends to be complete within the period.

Thus we have shown that a Koyck distributed lag equation is equivalent to a partial adjustment equation. The converse is also true, for it can be demonstrated that a partial adjustment equation gives rise to a Koyck equation. In fact, we can consider Eq. (3.44) as a first-order non-homogeneous difference equation, where x_t is an unspecified function. Hence we can apply the method explained in Sect. 3.2.6 to find a particular solution, and obtain

$$\overline{y}_t = \alpha \sum_{i=0}^{\infty} (1-\alpha)^i x_{t-i}, \tag{3.54}$$

which coincides with Eq. (3.50) for $\alpha = 1-k$ and $b = \sum_i b_i = 1$, as already assumed above. To be precise, the partial adjustment equation is slightly more general than the Koyck equation, since the general solution of Eq. (3.44) also contains the general solution of the corresponding homogeneous equation, so that

$$y_t = A(1-\alpha)^t + \alpha \sum_{i=0}^{\infty} (1-\alpha)^i x_{t-i},$$

where, however, the term $(1-\alpha)^t$ is negligible for t sufficiently great.

3.5 Exercises

3.5.1 Example

Let us solve the following non-homogeneous equation

$$y_t + 1.5y_{t-1} = 60, \qquad y_0 = 25. \tag{3.55}$$

The reader can immediately recognize that it belongs to the type described in subsection 3.2.1. Therefore the particular solution is obtained setting $y_t = y_{t-1} = \overline{y}$ (an undetermined constant) in (3.55), whence

$$\overline{y} + 1.5\overline{y} = 60,$$

therefore

$$\overline{y} = 24. \tag{3.56}$$

The corresponding homogeneous equation is

$$y_t + 1.5y_{t-1} = 0,$$

whose general solution is

$$y_t = A(-1.5)^t, \tag{3.57}$$

where A is an arbitrary constant. Adding (3.56) and (3.57) we obtain the general solution of (3.55):

$$y_t = A(-1.5)^t + 24. \tag{3.58}$$

Since -1.5 is negative and in absolute value greater than unity (improper oscillations of ever increasing amplitude), the stationary equilibrium is unstable.

Given the initial condition $y_0 = 25$, we can determine the value of the arbitrary constant A. From (3.58) we have, letting $t = 0$,

$$y_0 = A(-1.5)^0 + 24,$$

Table 3.1: Numerical computation of the solution to a first-order difference equation

t	$y_t = -1.5y_{t-1} + 60$	$y_t = -(1.5)^t + 24$
0	y_0=25	y_0=25
1	y_1=-1.5×22.5+60=22.5	y_1=(-1.5)^1+24=-1.5+24=22.5
2	y_2=-1.5×22.5+60=26.25	y_2=(-1.5)^2+24=2.25+24=26.25
3	y_3=-1.5×26.25+60=20.625	y_3=(-1.5)^3+24=-3.375+24=20.625
4	y_4=-1.5×20.625+60=29.0625	y_4=(-1.5)^4+24=5.0625+24=29.0625

and so

$$25 = A + 24,$$
$$A = 1.$$

We have then

$$y_t = (-1.5)^t + 24. \tag{3.59}$$

To compute the numerical values of y_t manually, we can either use Eq. (3.55) recursively or utilise Eq. (3.59). Table 3.1 shows the computations in the two cases.

From the computational point of view the first method of finding the successive values of y_t is more efficient, since in general it involves a smaller number of operations. The second method, however, is preferable when we have to calculate the value of y_t for a single given value of y_t, since we do not have to compute all the preceding values.

3.5.2 Other exercises

1. (a) Find the general solution of the following difference equations:

(*i*)	$y_t - 0.5y_{t-1} = 10,$	(t^*, y^*)	$= (1, 10);$
(*ii*)	$4y_t + 2y_{t-1} = 90,$	(t^*, y^*)	$= (0, 18);$
(*iii*)	$y_t + y_{t-1} = 90,$	(t^*, y^*)	$= (0, 50);$
(*iv*)	$2y_{t+1} + 3y_t + 2 = 0,$	(t^*, y^*)	$= (0, -1).$

(b) Find the equilibrium values for the equations above and check their stability.

2. Find the general solution of the following difference equations:

(*i*)	$y_{t+1} - \frac{1}{2}y_t = t + 3,$	$(t^*, y^*) =$	$(0, 3);$
(*ii*)	$y_t \top 3y_{t-1} = 4^t,$	$(t^*, y^*) =$	$(0, 0);$
(*iii*)	$y_{t+1} + \frac{\sqrt{2}}{2}y_t = \cos\frac{\pi t}{4},$	$(t^*, y^*) =$	$(0, 0).$

3. Consider the Koyck distribution and show that the mean time-lag is $k/(1-k) = (\alpha - 1)/\alpha$. [Hint: assume $b = 1$. Then the definition of mean time-lag becomes $\sum_{i=0}^{\infty} i(1-k)k^i = (1-k)(0 + k + 2k^2 + 3k^3 + \ldots)$].

3.6 References

Allen, R.G.D., 1959, *Mathematical Economics*, Sect. 6.3 and Appendix A.
Baumol, W.J., 1970, *Economic Dynamics,* Chap. 9.
Boole, G., 1960 (1872), *A Treatise on the Calculus of Finite Differences,* Chaps. II, IX, XI.
Goldberg, S., 1958, *Introduction to Difference Equations,* Chap. 2, Sect. 2.4.
Milne-Thomson, L.M., 1960, *The Calculus of Finite Differences,* Chaps. XI, XIII.

Chapter 4

First-order Difference Equations in Economic Models

4.1 The cobweb theorem

This is a dynamic model derived from the static supply and demand model. Assume that supply reacts to price with a lag of one period, while demand depends on current price, and that both functions are linear. In symbols:

$$D_t = a + bp_t,$$
$$S_t = a_1 + b_1 p_{t-1}.$$

Why should supply behave in that way? First of all, the model relates to goods whose production is not instantaneous nor continuous, but requires a fixed period of time (which we take as the unit of measurement of time). At the end of each period the output 'started' at the beginning of the period materializes, and the market determines its price (e.g., agricultural production). Producers believe—and this is the crucial assumption—that this price will hold also in the next period and so start the new production according to the current price. When output materializes (one period later), the price by which it has been determined is obviously the price of the previous period.

A last assumption is the market-clearing assumption: in each period, the market determines the price in such a way that demand absorbs exactly the quantity supplied, i.e. no producer is left with unsold output and no consumer with unsatisfied demand. This means that

$$D_t = S_t,$$

whence

$$bp_t - b_1 p_{t-1} = a_1 - a. \tag{4.1}$$

The general solution of the corresponding homogeneous equation is $A(b_1/b)^t$ and the particular solution of (4.1) is $p_e = (a_1 - a)/(b - b_1)$. On the assumption that p_0 is known, the arbitrary constant is determined as $A = p_0 - p_e$.

The general solution of Eq. (4.1) is then

$$p_t = (p_0 - p_e)\left(\frac{b_1}{b}\right)^t + p_e. \tag{4.2}$$

Can we give an economic interpretation to the particular solution p_e?

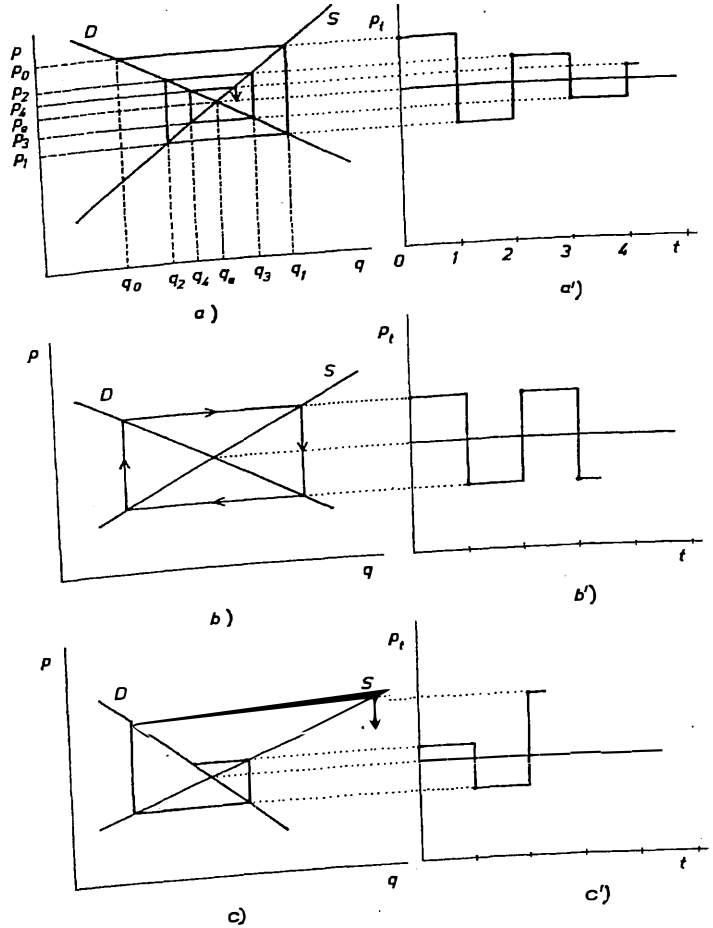


Figure 4.1: The cobweb theorem

The answer is yes: it is the static equilibrium value of price (the subscript e stands for 'equilibrium'). We can see from (4.2) that if the initial price

happens to be p_e, then $p_t = p_e$, i.e. price stays fixed at p_e (of course if no exogenous disturbances occur) and that is what we mean by (static) equilibrium. Another way to arrive to the same interpretation is to check that the value of p_e, given above, is the same as that obtained by the solution of the static supply and demand model:

$$\begin{aligned} D &= a + bp, \\ S &= a_1 + b_1 p, \\ D &= S. \end{aligned}$$

Let us now analyse the behaviour over time of price, as given by Eq. (4.2).

Usually demand has a negative slope ($b < 0$) and supply a positive one ($b_1 > 0$). Then $b_1/b < 0$, and so the price will have an oscillatory movement around its equilibrium value. These improper oscillations will be explosive, of constant amplitude, or damped according as $|b_1| \gtreqless |b|$, i.e. according as supply has a slope greater than, equal to, or smaller than, the absolute value of the slope of demand.

In Fig. 4.1 three cases are shown. The cobweblike aspect of the diagrams on the left justifies the name given by Kaldor (1934, but the model had been introduced earlier by other authors, see Waugh, 1964) to the model. Take, for example, diagram (a). Let us suppose that in the initial period the system is not in equilibrium because of an exogenous disturbance (e.g., a drought), and let q_0 be the initial quantity; the corresponding initial price p_0. The price p_0 induces entrepreneurs to produce the quantity q_1, which, as we know, materializes in period 1. This quantity will be exactly absorbed by demand at price p_1, which in turns induces output q_2 and so on.

The succession of prices over time is shown in diagram (a'). The movement converges toward equilibrium, i.e. we are in a case where the slope of supply is smaller in absolute value than the slope of demand (in checking this, it must be remembered that demand and supply functions have been written with price as the independent variable; slopes must be computed with reference to the p axis).

We have so far considered only 'normal' cases, i.e. those in which demand has negative slope and supply has a positive slope. Two among the conceivable 'abnormal 'cases are shown in Fig. 4.2. Here the movement is monotonic. The conditions on slope hold here too.

As we can see from the general solution (4.2), the stability condition, i.e. the condition that price converges—no matter how—to its equilibrium value, is in any case $|b_1/b| < 1$, i.e. $|b_1| < |b|$.

This completes the examination of the traditional cobweb theorem.

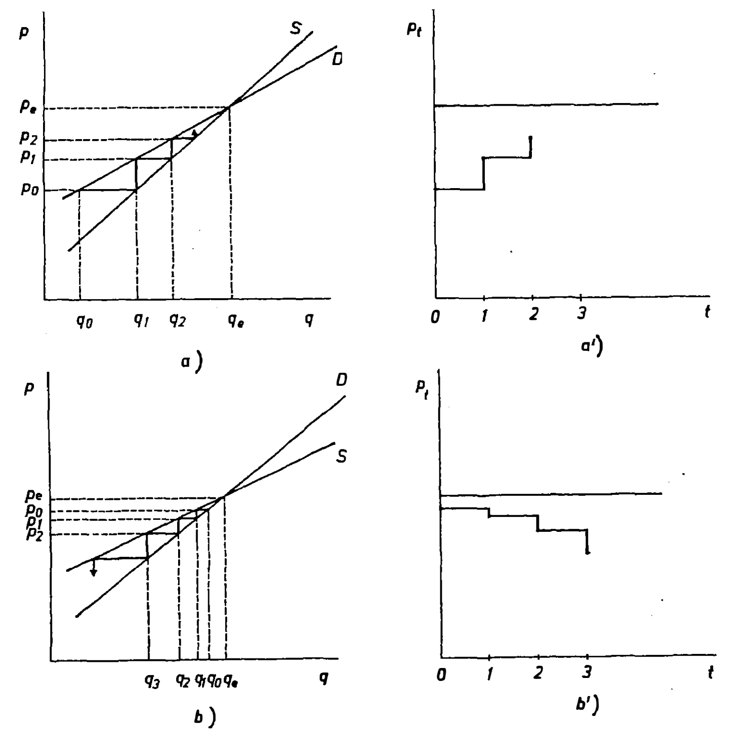


Figure 4.2: The cobweb theorem: abnormal cases

4.1.1 The cobweb model and expectations

The traditional cobweb model can be considered as a particular case of the more general model

$$\begin{aligned} D_t &= a + bp_t, \\ S_t &= a_1 + b_1 p_t^e, \\ D_t &= S_t, \end{aligned} \tag{4.3}$$

where p_t^e indicates the price expected by producers, namely the price that producers, at the moment of 'starting' the production, think will hold when output materializes. In the original cobweb theorem the (implicit or explicit) assumption is, as we have seen above, that $p_t^e = p_{t-1}$.

Now, it seems implausible to assume that producers continue to expect that the price will remain constant to its previous level when, on the contrary, it continues varying period after period. Everybody learns from experience. The formation of expectations inherent in the traditional cobweb theorem

cannot be accepted as a plausible description of reality. There are several ways to improve on this situation.

4.1.1.1 The normal price

One possibility is to introduce the concept of 'normal price' as that price which producers think will sooner or later obtain in the market. Then if the current price is different from the normal price, they think the former will modify, moving toward the latter. The simplest way to formalize this is to put

$$p_t^e = p_{t-1} + c(p_N - p_{t-1}), \qquad 0 < c < 1. \tag{4.4}$$

If, for example, the current price is lower than the normal price, an increase in price is expected for the next period. The fact that the positive constant c is assumed to be less than unity is equivalent to the assumption that producers do not think that the normal price will immediately become the normal price (this would be the case only if $c = 1$, whence $p_t^e = p_N$), but think that the process will require a certain amount of time (measured by the reciprocal of c). Let us note, incidentally, that when $c = 0$ we are back to the original cobweb theorem.

What we have said above would be little more than a tautology if we did not specify how to determine the normal price. A simple way of doing this is to assume that p_N equals p_e, the static equilibrium price. This would be the case if, for example, producers have perfect information (i.e., they know the model governing the determination of the market price) but know that, due to frictions and other elements, the price cannot immediately go back to the value p_e when it is displaced from equilibrium.

Now let us turn to the mathematics. Substituting (4.4) , where $p_N = p_e$, in (4.3) we obtain

$$b_1(1-c)p_{t-1} - bp_t = a - a_1 - b_1 c p_e, \tag{4.5}$$

whose solution is

$$p_t = A\left[\frac{b_1(1-c)}{b}\right]^t + p_e. \tag{4.6}$$

The stability condition is

$$|b_1(1-c)| < |b|. \tag{4.7}$$

The quantity $1 - c$ is positive and less than unity, since $0 < c < 1$. Then the absolute value of $b_1(1-c)$ is smaller than the absolute value of b_1. If we compare our new model with the original cobweb theorem we obtain the following results (the first term in the comparisons refers to the original cobweb theorem):

(1) A convergent movement remains convergent and the convergence is faster. In fact, since

$$\left|\frac{b_1(1-c)}{b}\right| < \left|\frac{b_1}{b}\right| < 1,$$

the absolute value of $[b_1(1-c)/b]^t$ tends to zero more rapidly than the absolute value of $(b_1/b)^t$.

(2) An improper oscillation of constant amplitude becomes damped. This is so because, if $|b_1| = |b|$, then $|b_1(1-c)| < |b|$.

(3) A divergent movement may become convergent (or of a constant amplitude) if the parameter c is sufficiently close to unity and when it remains divergent, the divergence is slower. This is so because the greater is c the smaller is $(1-c)$—remember that in any case $0 < c < 1$—and so the more likely it is that $|b_1(1-c)| \leq |b|$ even if $|b_1| > |b|$. In any case

$$\left|\frac{b_1(1-c)}{b}\right| < \left|\frac{b_1}{b}\right|,$$

and so, if the movement remains divergent, the absolute value of $[b_1(1-c)/b]^t$ increases at a slower rate than the absolute value of $(b_1/b)^t$. Note that a greater value of c means that the producers expect a faster approach to the current price towards its equilibrium value.

Our conclusions show that in any case the introduction of expectations based on the normal price (assumed to be equal to the equilibrium price) makes the model more stable, and this is a sensible result.

4.1.1.2 Adaptive expectations

Another way of 'improving' the original cobweb theorem by means of more realistic expectations is to use *adaptive* expectations. Actually, this form of expectations, used in other fields of economics as well, was introduced by Nerlove in his study on cobweb phenomena (Nerlove, 1958). According to this formulation, expectations are revised or 'adapted' in each period on the basis of the discrepancy between the observed value and the previously expected value, that is

$$p_t^e - p_{t-1}^e = \beta(p_{t-1} - p_{t-1}^e), \tag{4.8}$$

where β is a positive coefficient smaller than one. This equation tells us that, if the observed value of the previous period has been greater (smaller) than the expected value for the same period, i.e. if $p_{t-1} - p_{t-1}^e \gtreqless 0$, then the new expected value is revised upwards, left constant, or revised downwards. Note that Eq. (4.8) by itself does note gives us the expected value, but only the rule of its variation. It is however possible to obtain p_t^e by considering Eq.(4.8) as a difference equation in the unknown function p_t^e. First of all, write it in the form

$$p_t^e - (1-\beta)p_{t-1}^e = \beta p_{t-1}, \tag{4.9}$$

and note that the general solution for the corresponding homogeneous equation is

$$p_t^e = A(1-\beta)^t. \tag{4.10}$$

As regards a particular solution, we are in the case described in Sect. 3.2.6. Since $|1-\beta| < 1$, we can apply (3.37) and obtain

$$\overline{p_t^e} = \beta\sum_{i=0}^{\infty}(1-\beta)^i p_{t-1-i}. \tag{4.11}$$

Thus the general solution of Eq. (4.9) is

$$p_t^e = A(1-\beta)^t + \beta\sum_{i=0}^{\infty}(1-\beta)^i p_{t-1-i}, \tag{4.12}$$

where the term $A(1-\beta)^t$ can be neglected for t sufficiently great. Hence adaptive expectations mean that the expected price is a weighted average (with geometrically declining weights) of all past observed prices[1].

The use of (4.12) would bring us beyond first-order equations; however , it is possible to solve the model without using (4.12). As $S_t = a_1 + b_1 p_t^e$, we have $S_{t-1} = a_1 + b_1 p_{t-1}^e$, whence

$$p_t^e = \frac{S_t - a_1}{b_1}, \qquad p_{t-1}^e = \frac{S_{t-1} - a_1}{b_1},$$

and substituting these values in (4.9) we obtain

$$\frac{S_t - a_1}{b_1} = (1-\beta)\frac{S_{t-1} - a_1}{b_1} + \beta p_{t-1},$$

whence

$$S_t = (1-\beta)S_{t-1} + a_1\beta + b_1\beta p_{t-1}. \tag{4.13}$$

[1] The sum of weights, owing to the fact that $0 < \beta < 1$ (the case $\beta = 1$ can be excluded because we would fall back into the original cobweb theorem), is equal to one, as it should. In fact,

$$\sum_{i=0}^{\infty}\beta(1-\beta)^i = \beta\sum_{i=0}^{\infty}(1-\beta)^i = \beta\frac{1}{1-(1-\beta)} = 1.$$

Let us note, incidentally, that (4.12) is a distributed lag equation of the Koyck type: see note 1 Chap. 8. In general, partial adjustment equations (of which adaptive expectations are a case) are equivalent to a Koyck scheme.

Now $D_t = S_t$ for all t by assumption, and since $D_t = a + bp_t$, we can substitute $a + bp_t$ and $a + bp_{t-1}$ respectively in the place of S_t and S_{t-1} in (4.13), obtaining

$$a + bp_t = (1-\beta)a + (1-\beta)bp_{t-1} + a_1\beta + b_1\beta p_{t-1},$$

which gives

$$p_t - \left[\left(\frac{b_1}{b} - 1\right)\beta + 1\right]p_{t-1} = \frac{(a_1 - a)\beta}{b}. \tag{4.14}$$

A particular solution of (4.14) is p_e, that is the equilibrium price. The general solution of the corresponding homogeneous equation is

$$p_t = A\left[\left(\frac{b_1}{b} - 1\right)\beta + 1\right]^t,$$

and so the general solution of (4.14) is

$$p_t = A\left[\left(\frac{b_1}{b} - 1\right)\beta + 1\right]^t + p_e, \tag{4.15}$$

where $A = p_0 - p_e$ is the initial deviation. The stability condition is

$$\left|\left(\frac{b_1}{b} - 1\right)\beta + 1\right| < 1, \tag{4.16}$$

that is

$$-1 < \left(\frac{b_1}{b} - 1\right)\beta + 1 < 1.$$

Adding -1 to all members we get

$$-2 < \left(\frac{b_1}{b} - 1\right)\beta < 0.$$

Dividing through by β and adding $+1$ to all members we obtain

$$1 - \frac{2}{\beta} < \frac{b_1}{b} < 1. \tag{4.17}$$

Let us now compare Eq.(4.17) with the stability condition holding in the original cobweb theorem, $|b_1/b| < 1$. The latter can also be written as

$$-1 < \frac{b_1}{b} < 1. \tag{4.18}$$

Since $0 < \beta < 1$, we have $2/\beta > 2$, and so

$$1 - \frac{2}{\beta} < -1. \tag{4.19}$$

It follows that (4.17) is *less* stringent than (4.18) as far as the left side is concerned. The latter is the side which matters in the case of 'normal' functions: we conclude also here that the introduction of adaptive expectations makes the model more stable.

We shall examine later (Chap. 6, Sect. 6.3, exercise 5) the consequence of introducing other types of expectations in the cobweb theorem.

4.2 The dynamics of multipliers

4.2.1 The basic case

We know from elementary Keynesian macroeconomics that in a model without government or foreign sector, an increase ΔI in autonomous investment (or more generally in autonomous expenditure) brings about an increase in national income according to the multiplier equation

$$\Delta Y = \frac{1}{1-b}\Delta I,$$

where b is the marginal propensity to consume. This is a result of the comparative statics kind[2], i.e., given an equilibrium position (where income is, say, Y_0) and given an autonomous shift of investment from I to $I + \Delta I$, the *new* equilibrium value of income is

$$Y_0 + \frac{1}{1-b}\Delta I.$$

This result, however, does not say anything about the movement from the old to the new equilibrium—we do not know if income will move towards (or away from) its new equilibrium value. Only a dynamical model can elucidate these points. The usual dynamical assumption is that consumption depends on income with a one-period lag, i.e.

$$C_t = a + bY_{t-1}, \qquad a \geq 0, \qquad 0 < b < 1. \tag{4.20}$$

Investment, for the moment, is assumed to be wholly autonomous, and in the initial period shifts from I_0 to $I_0 + \Delta I$ (remaining at this level in all subsequent periods):

$$I_t = I_0 + \Delta I. \tag{4.21}$$

The equation

$$Y_t = C_t + I_t \tag{4.22}$$

[2]The reader wanting to know more about comparative statics is referred to Part III, Chap. 20.

closes the model [3].

Substituting from (4.20) and (4.21) in (4.22) we obtain

$$Y_t - bY_{t-1} = a + I_0 + \Delta I, \tag{4.23}$$

whose solution is

$$Y_t = A(b)^t + \frac{a + I_0 + \Delta I}{1 - b}. \tag{4.24}$$

The new equilibrium value $(a+I_0+\Delta I)/(1-b)$ and the initial [4] equilibrium value $(a + I_0)/(1 - b)$ of income differ by the quantity $(\Delta I)/(1 - b)$,which is the result stated at the beginning. But Eq. (4.24) says something more. Since $0 < b < 1$, the term $A(b)^t$ tends to zero as time goes on and this means that income actually moves (with a monotonic movement) towards its new equilibrium value. In Fig 4.3 a graphical representation of this approach is given. The static part of the diagram is the well-known Keynesian 'cross'. The initial level of investment is $\overline{OR}$ (this means that we assume here $a = 0$) and $\overline{OR}'$ is its new level (the shift has been exaggerated for graphical convenience). The corresponding equilibria are Y_0 and Y_E respectively. Let us now examine the dynamic process. In period 1, consumption -which depends on income in period zero- is $\overline{A_1P_1}$. Adding it to the (new) investment $\overline{OR}' = \overline{A_1Y_o}$ we obtain income in period 1, which is then $\overline{Y_0P_1}$. By means of the 45° line we transfer this segment to the axis, obtaining point Y_1. In period 2, consumption is $\overline{A_2P_2}$ and income is $\overline{Y_1P_2} = \overline{A_2P_2} + \overline{A_2Y_1}$; by means of the 45° line we obtain point Y_2 on the abscissae, and so on. As we see, the system tends monotonically towards point P_E, i.e. towards Y_E.

We have so far examined the case in which investment is entirely autonomous. Let us now assume that it is partly autonomous and partly depending on income (with a one-period lag[5]) according to the marginal propensity to invest h, $0 < h < 1$. Eq. (4.21) above becomes

$$I_t = hY_{t-1} + I_0 + \Delta I,$$

and Eq. (4.23) becomes

$$Y_t - (b + h)Y_{t-1} = a + I_0 + \Delta I,$$

whose solution is

$$Y_t = A(b + h)^t + \frac{a + I_0 + \Delta I}{1 - b - h}.$$

[3] Eq. (4.22) involves the well known problem of '*ex-ante*' and '*ex-post*' (or equilibrium conditions and identities in national income analysis). See any macroeconomics textbook.

[4] If we assume that in the initial situation income was at its equilibrium value, then $Y_0 = a + bY_0 + I_0$, whence $Y_0 = (a + I_0)/(1 - b)$.

[5] The student may check as an exercise that, if we put $I_t = hY_t + I_0 + \Delta I_0$, the multiplier and the stability condition below are not altered.

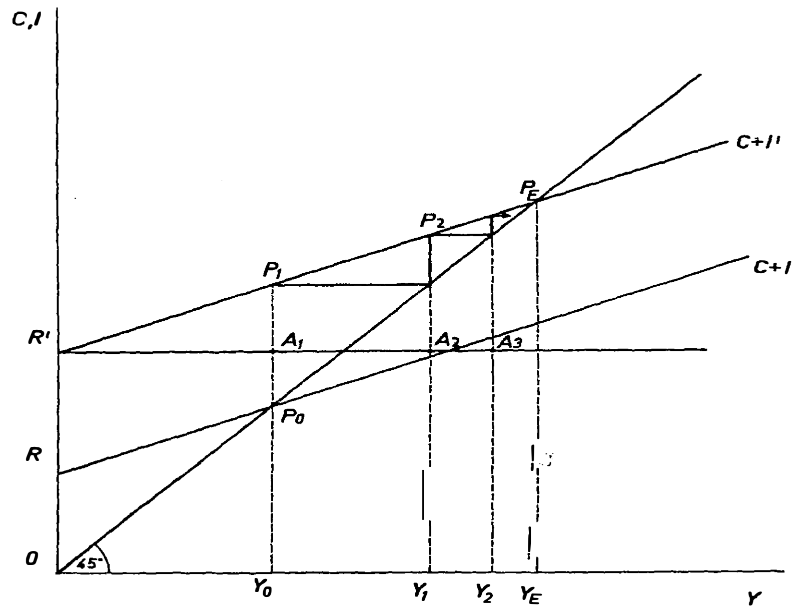


Figure 4.3: The simple dynamic multiplier

Since b and h are both positive magnitude, the movement is monotonic. Stability requires that $b + h < 1$, i.e.

$$h < 1 - b. \tag{4.25}$$

Since $1 - b$ is the marginal propensity to save, condition (4.25) says that the marginal propensity to invest must be smaller than the marginal propensity to save if equilibrium is to be stable. Let us note, incidentally, that from (4.25) it follows that $0 < 1 - b - h$, and this ensures that the multiplier $1/(1 - b - h)$ is positive[6].

4.2.2 Other multipliers

Multiplier analysis offers a wealth of examples. We shall examine two other well-known cases.

[6] This is an application of the 'correspondence principle'. See Part III, Chap. 20.

4.2.2.1 A foreign trade multiplier

As a further example of multiplier analysis, we shall now examine a foreign trade multiplier. Imports are a function of income and exports are assumed to be wholly exogenous (for a more general model see Sect. 10.2). In the open economy, aggregate supply is the sum of national product and imports; aggregate demand is national consumption plus national investment plus exports. The equilibrium condition is then no longer $Y = C+I$, but $Y+M = C+I+X$ or $Y = C+I+X-M$. The formal model is then the following:

$$\begin{aligned}
C_t &= a + bY_{t-1}, \\
I_t &= hY_{t-1} + I_0 + \Delta I, \\
X_t &= X_0 + \Delta X, \\
M_t &= mY_{t-1} + M_0, \qquad 0 < m < 1, \\
Y_t &= C_t + I_t + X_t - M_t.
\end{aligned}$$

Substituting from the first four equations of the above system into the fifth, we obtain

$$Y_t - (b+h-m)Y_{t-1} = a + I_0 + X_0 - M_0 + \Delta I + \Delta X, \tag{4.26}$$

whose solution is

$$Y_t = A(b+h-m)^t + \frac{a + I_0 + X_0 - M + \Delta I + \Delta X}{1-b-h+m}. \tag{4.27}$$

Note that the multiplier is now $1/(1-b-h+m)$. The 'leakage' due to the imports is represented by m, the marginal propensity to import.

The sum $b+h$ is surely greater than m. This apparently unwarranted statement can be justified on the following economic grounds. Take a unit increment in income, which causes an increment in both consumption and investment; this increment is measured by $b+h$. It follows that $b+h-m$ which measures the marginal propensity to spend on domestic goods, is a positive quantity. Hence the movement given by (4.27) is monotonic.

The stability condition is that $b+h-m < 1$. This inequality is equivalent to the inequality $1-b-h+m > 0$ and so the stability condition ensures that the multiplier $1/(1-b-h+m)$ is positive. The condition may also be written in the form

$$h < 1 - b + m, \tag{4.28}$$

i.e. the marginal propensity to invest must be smaller than the sum of the marginal propensity to save and the marginal propensity to import.

A related and interesting question is whether the foreign trade multiplier brings about a complete adjustment in the balance of trade. Assume that trade is initially balanced (i.e. $X = M$) and that export autonomously

increase. Income increases according to the foreign trade multiplier and also imports increase, since they are an increasing function of income. Will the (induced) increase in imports exactly offset the (exogenous) increase in exports? Formally, we have

$$\Delta Y = \frac{1}{1-b-h+m}\Delta X$$

and

$$\Delta M = m\Delta Y = \frac{m}{1-b-h+m}\Delta X.$$

It follows that $\Delta M = \Delta X$ if, and only if, $m/(1-b-h+m) = 1$, so that $h = 1-b$. It must be emphasized that the last equality cannot be ruled out by stability considerations. It is true that in a closed economy h must be smaller that $1-b$ (see condition (4.25) above), but here we are in an open economy and the stability condition (4.28) is compatible with h being equal to $1-b$. In general $\Delta M \gtreqless \Delta X$ according as $h \gtreqless 1-b$, any of the three cases being in principle possible.

4.2.2.2 Taxation

As a last example, let us now examine the case in which taxation is present. To simplify the problem we consider the closed economy model and assume that taxation is a simple linear function of income,

$$T = T_a + \tau Y, \qquad 0 < \tau < 1.$$

Consumption is now a function of disposable income Y_d, that in our simplified model can be calculated as $Y - d - T + r$, where d is depreciation and r net transfers, both assumed exogenous. Thus we have

$$\begin{aligned}
C_t &= a + bY_{d_{t-1}}, \\
I_t &= hY_{t-1} + I_0 + \Delta I_0, \\
T_t &= T_a + \tau Y_t, \\
Y_{d_t} &= Y_t - d - T_t + r, \\
Y_t &= C_t + I_t + G_t,
\end{aligned} \tag{4.29}$$

where G is government expenditure. Substitution from the first four equations in the fifth and rearrangement of terms yields

$$Y_t - [b(1-\tau)+h]Y_{t-1} = a - bT_a + I_0 + \Delta I_0 + G + br - bd,$$

whose solution is

$$Y_t = A[b(1-\tau)+h]^t + \frac{a + I_0 + \Delta I_0 + G + b(-T_a + r - d)}{1 - b(1-\tau) - h}$$

Since both $b(1-\tau)$ and h are positive, the movement is monotonic. The multiplier $1/[1-b(1-\tau)-h]$ is smaller than the multiplier holding when taxation is absent, $1/(1-b-h)$. The condition is $b(1-\tau)+h<1$, that is

$$h < 1 - b + b\tau, \tag{4.30}$$

which is less stringent than condition (4.28). Therefore, the introduction of taxation makes the model more stable.

4.3 Exercises

1. Consider the following model

$$\begin{aligned} S_t &= sY_{t-1}, \\ I_t &= k(Y_t - Y_{t-1}), \\ S_t &= I_t, \end{aligned} \tag{4.31}$$

where s is the propensity to save (marginal=average) and k is the acceleration coefficient. The third equation states the equilibrium condition for the determination of national income Y (*ex ante* saving = *ex ante* investment). Substitution from the two first equations into the third immediately gives

$$Y_t - \frac{k+s}{k}Y_{t-1} = 0.$$

The solution of this homogeneous difference equation is

$$Y_t = A\left(\frac{k+s}{k}\right)^t = Y_0\left(1+\frac{s}{k}\right)^t, \tag{4.32}$$

where the arbitrary constant A has been determined letting $Y_t = Y_0$ for $t=0$. Eq. (4.32) tells us that equilibrium income increases over time at the constant rate of growth s/k. This rate is called by Harrod—whose dynamic growth model is represented in an admittedly simplified form in Eqs. (4.31)—the *warranted* rate of growth. It is a rate such that, when income grows according to it, there is a continuous equality between *ex ante* saving and *ex ante* investment, i.e. a dynamic equilibrium obtains. This is clear, as we have obtained Y_t as a function of time in Eq. (4.32) imposing the condition $S_t = I_t$.

2. Modify model (4.31) assuming that saving is in proportion to current income, $S_t = sY_t$, while the investment function is the same. Discuss the solution and show that growth requires $k > s$ (since $s<1$ and $k>1$, this condition usually holds). Why is the warranted rate of growth greater than Harrod's s/k?

3. Consider the following model, due to Cagan (1956):

$$\begin{aligned} l_t &= \alpha(p^e_{t+1} - p_t), \quad \alpha < 0, \\ m_t &= \text{exogenous}, \\ l_t &= m_t - p_t, \end{aligned} \tag{4.33}$$

where l_t= demand for real money balances, p^e_{t+1}= expected price for next period, p_t= current price, m_t=nominal money supply. All variables are expressed in logarithms. Hence $p^e_{t+1} - p_t$ is the expected inflation rate, and the first equation states that current demand for real balances varies inversely with expected inflation. The third equation is the usual equilibrium condition. To solve the model we must make assumptions on p^e_{t+1}. Suppose first that economic agents expect the future rate of inflation to be equal to the current one multiplied by a constant, i.e.

$$p^e_{t+1} - p_t = \gamma(p_t - p_{t-1}), \quad \gamma > 0. \tag{4.34}$$

Then model (4.33) becomes

$$(1+\alpha\gamma)p_t - \alpha\gamma p_{t-1} = m_t, \tag{4.35}$$

which is the dynamic equation describing the behaviour of prices. The general solution of the homogeneous part of this equation is

$$p_t = A\left(\frac{\alpha\gamma}{1+\alpha\gamma}\right)^t. \tag{4.36}$$

The behaviour of prices will be stable if $|\alpha\gamma| < |1+\alpha\gamma|$, unstable in the opposite case.

To complete the exercise, the reader must find a particular solution (Hint: see Sect. 3.2.6 and apply the appropriate expansion according to the stable or unstable nature of (4.36)).

4. In model (4.33) assume perfect foresight, i.e. $p^e_{t+1} = p_{t+1}$. The basic difference equation becomes

$$p_t - \frac{\alpha-1}{\alpha}p_{t-1} = \frac{1}{\alpha}m_{t-1}. \tag{4.37}$$

Show that now the model is monotonically unstable. Hence find a particular solution by solving forwards.

5. Let us assume that in a certain period equilibrium income is at the level $Y_0 = 100$, while C_0 is 60 and I_0 is 40. The consumption function is $C_t = 0.60Y_{t-1}$ and investment is entirely autonomous. Suddenly, for whatever reason, the investment shifts from 40 to 50. Therefore it is interesting to analyse whether the new equilibrium is a stable position.

Table 4.1: Numerical example of the multiplier

Period	Investment	Consumption	Income
0	40	60	100
1	50	60	110
2	50	66	116
3	50	69.6	119.6
4	50	71.76	121.76
5	50	73.056	123.056
6	50	73.8336	123.8336
....	....	....	.....
$t \to \infty$	50	$C_t \to 75.00$	$Y_t \to 125$

The increment in income is obtained applying the static multiplier $1/(1-0.6)$ to the increment in autonomous investment; the result is 25. The new equilibrium value of income is then 125. Substituting the given functions in the equilibrium $Y_t = C_t + I_t$, we obtain

$$Y_t = 0.60Y_{t-1} + 50,$$

whose solution is

$$Y_t = A(0.60)^t + 125,$$

where

$$A = Y_0 - 125 = 100 - 125 = -25,$$

and so

$$Y_t = -25(0.60)^t + 125.$$

Since $(0.60)^t$ tends monotonically to zero as t increases, Y_t tends to 125, and so the equilibrium is stable.

As a further exercise, Table 4.1 can be computed.

Let us now relax the assumption on complete exogeneity of investments. Consider $I_t = 0.20Y_{t-1} + 20$ (the consumption function is the same); the autonomous component in the investment function shifts from 20 to 30. This assumption implies that the multiplier $1/(1 - 0.6 - 0.2)$, applied to the increment in autonomous investment, gives an increment in income of 50. The new equilibrium value of income is then 150. Substituting the given functions in equation $Y_t = C_t + I_t$, we obtain

$$Y_t = 0.60Y_{t-1} + 0.20Y_{t-1} + 30,$$

therefore

$$Y_t - 0.80Y_{t-1} = 30,$$

whose solution is

$$Y_t = A(0.80)^t + 150,$$

where

$$A = Y_0 - 150 = 100 - 150 = -50,$$

and so

$$Y_t = -50(0.80)^t + 150.$$

Since $(0.80)^t$ tends to zero monotonically as t increases, Y_t tends to 150, consequently the equilibrium is characterized as stable.

6. A way of taking into account the influence of habit formation on consumption is to introduce the consumption of the previous period as an additional explanatory variable in the consumption function. In the simplest case we have the model

$$\begin{aligned} C_t &= a + bY_t + cY_{t-1}, \qquad 0 < c < 1, \\ Y_t &= C_t + I_t, \\ I_t &\ \text{to be defined.} \end{aligned}$$

(a) Examine the stability condition in the case of a constant value of investment, $I_t = I_0$.

(b) Solve the model assuming that investment grows at the exogenous constant rate g, namely $I_t = I_0(1+g)^t$.

7. Given the static demand and supply functions: $D_t = a + bp_t$, $S_t = a_1 + b_1p_t$, consider the dynamic mechanism that expresses the standard adjustment of price to excess demand in a market:

$$\Delta p_t = \alpha(D_t - S_t),$$

where $\alpha > 0$ is the reactivity of price to excess demand. Examine the stability of the static equilibrium point.

4.4 References

Allen, R.G.D., 1959, *Mathematical Economics*, Chap. 1, Sects. 1.1, 1.2, 1.4; Chap. 2, Sects. 2.1, 2.5-2.8.
Baumol, W.J., 1970, *Economic Dynamics*, Chap. 4.
Buchanan, N.S., 1939, A Reconsideration of the Cobweb Theorem.
Cagan, P., 1956, The Monetary Dynamics of Hyperinflation.
Ezekiel, M., 1938, The Cobweb Theorem.
Harrod, R.F., 1939, An Essay in Dynamic Theory.
Kaldor, N., 1960 (1934), Determinateness of Static Equilibrium, pp. 4, 30-33.
Nerlove, M., 1958, Adaptive Expectations and Cobweb Phenomena.

Wan, H.Y., Jr., 1971, *Economic Growth*, Chap. 1, Sect. 1.2.
Waugh, F.V., 1964, Cobweb Models.

Chapter 5

Second-order Difference Equations

The general form of these equation is

$$c_2 y_t + c_1 y_{t-1} + c_0 y_{t-2} = g(t), \tag{5.1}$$

where c_0, c_1, c_2 are given constants and $g(t)$ is a known function. The coefficients c_2 and c_0 must be both different from zero, since if either one is zero the equation becomes of the first order.

5.1 Solution of the homogeneous equation

Consider now the homogeneous equation

$$c_2 y_t + c_1 y_{t-1} + c_0 y_{t-2} = 0, \tag{5.2}$$

which can be written in the form

$$y_t + a_1 y_{t-1} + a_2 y_{t-2} = 0, \tag{5.3}$$

where

$$a_1 \equiv c_1/c_2, \qquad a_2 \equiv c_0/c_2.$$

We have seen in Chap. 3 that the general solution of the first-order equation involves a function of the type λ^t, where λ is a constant determined by means of the coefficients of the equation. Analogy leads us to think that the solution-function of Eq. (5.3) might be of the same type. Let us then substitute $y_t = \lambda^t$ in Eq. (5.3), the constant λ being for the moment undetermined. We obtain

$$\lambda^t + a_1\lambda^{t-1} + a_2\lambda^{t-2} = 0,$$

whence

$$\lambda^{t-2}(\lambda^2 + a_1\lambda + a_2) = 0. \tag{5.4}$$

If λ^t has to be a solution, Eq. (5.4) must be satisfied for any value of t, and this is true—apart from the trivial case $\lambda = 0$—if, and only if, the expression in parentheses, which does not involve t, is zero, i.e.

$$\lambda^2 + a_1\lambda + a_2 = 0. \tag{5.5}$$

Eq. (5.5) is called the *characteristic* (or *auxiliary*) *equation* of the difference equation (5.3). It is remarkable that in this way we have reduced the solution of a functional equation to the solution of an algebraic equation. The two roots of Eq. (5.5) are given by the well-known formula

$$\lambda_1, \lambda_2 = \frac{-a_1 \pm (a_1^2 - 4a_2)^{1/2}}{2}.$$

We must now examine in some detail the nature of these roots—and hence of the solution of Eq. (5.3)—according to the sign of the discriminant $\Delta \equiv a_1^2 - 4a_2$. Three cases are possible.

5.1.1 Positive discriminant ($\Delta > 0$)

The roots λ_1, λ_2 are real and distinct. Then both λ_1^t and λ_2^t satisfy Eq. (5.3) and—according to general principles (see Chap. 2, Theorem 2.4)—we can combine them linearly and obtain the general solution

$$y_t = A_1\lambda_1^t + A_2\lambda_2^t, \tag{5.6}$$

where A_1, A_2 are two arbitrary constants.

The kind of movement (monotonic or oscillatory or combination of the two) depends on the *sign* of λ_1, λ_2. We know that the t-th power of a positive number is always positive (monotonic movement), whereas the power of a negative number is positive or negative according as to whether the exponent is even or odd (improper oscillations, as we have defined them in Chap. 3). A great variety of movements is possible since, in principle, each root may have any sign (no root, however, can be zero, since $a_2 \neq 0$). The sign of the roots can be ascertained by **Descartes' theorem**: *In any algebraic equation, complete or incomplete, the number of positive roots cannot exceed the number of changes of signs of the coefficients, and in any complete equation the number of negative roots cannot exceed the number of continuations in the signs of the coefficients.* In our case, given of course that $\Delta > 0$, there are the following possibilities:

- $+ \ + \ +$ two negative roots,
- $+ \ + \ -$ one negative and one positive root (the negative root being greater in absolute value),
- $+ \ - \ +$ two positive roots,
- $+ \ - \ -$ one negative and one positive root (the positive root being greater in absolute value),
- $+ \ 0 \ -$ one positive and one negative root (with the same absolute value).

In any case the movement will be *convergent* if, and only if, *both* roots are in absolute value less than unity, since in that case both terms on the right-hand side of Eq. (5.6) tend to zero in absolute value.

Finally, we note that, as t increases, the movement of y_t will be dominated by the root numerically greater, which is called the *dominant* root.

5.1.2 Null discriminant ($\Delta = 0$)

Equation (5.5) has two real and equal roots: $\lambda_1 = \lambda_2 = -\frac{1}{2}a_1$, say λ^*, which is also called a multiple root with multiplicity two. We must find another solution of Eq. (5.3), since we have only one solution, λ^{*t}. Let us try $t\,\lambda^{*t}$. Substituting in Eq. (5.3) we have

$$t\lambda^{*t} + a_1(t-1)\lambda^{*t-1} + a_2(t-2)\lambda^{*t-2} = 0; \tag{5.7}$$

therefore

$$\lambda^{*t-2}[t\lambda^{*2} + a_1(t-1)\lambda^* + a_2(t-2)] = 0,$$

from which, using $\lambda^* = -\frac{1}{2}a_1$,

$$\left(-\frac{1}{2}a_1\right)^{t-2}\left(\frac{1}{4}a_1^2t - \frac{1}{2}a_1^2t + \frac{1}{2}a_1^2 + a_2t - 2a_2\right) = 0,$$

and so

$$\left(-\frac{1}{2}a_1\right)^{t-2}\left(-\frac{1}{4}a_1^2t + a_2t + \frac{1}{2}a_1^2 - 2a_2\right) = 0. \tag{5.8}$$

If $t\lambda^{*t}$ has to be a solution , Eq. (5.8) must be satisfied for any t. Since $a_1^2 - 4a_2 = 0$ by assumption, then $a_2 = \frac{1}{4}a_1^2$, and if we substitute this relation into the expression in the second set of parentheses in Eq. (5.8), this expression will vanish, thus proving that

$$t\lambda^{*t} \tag{5.9}$$

is a solution of Eq. (5.3) in the case under consideration. Thus λ^{*t} and $t\lambda^{*t}$ are two distinct solutions of (5.3). The general solution is

$$y_t = A_1\lambda^{*t} + A_2t\lambda^{*t} = (A_1 + A_2t)\lambda^{*t}. \tag{5.10}$$

Let us note that, when $|\lambda^*| < 1$, the solution will be damped, since, in the term $t\lambda^{*t}$, the damping due to λ^{*t} dominates the explosive tendency of the multiplicative t. More formally, $\lim_{t\to\infty} t\lambda^{*t} = 0$ if $|\lambda^*| < 1$.

5.1.3 Negative Discriminant ($\Delta < 0$)

In this case the roots are two complex conjugate numbers, i.e. numbers having the form $\alpha \pm i\theta$, where $i = +\sqrt{-1}$ is the imaginary unit and α, θ are real numbers. α is called the real part of the complex number and $i\theta$ is the imaginary part. Here $\alpha = \frac{1}{2}a_1$, $\theta = \frac{1}{2}(4a_2 - a_1^2)^{1/2}$. The solution is then

$$y_t = A'(\alpha + i\theta)^t + A''(\alpha - i\theta)^t,$$

where A' and A'' are two arbitrary constants which we may take as arbitrary complex conjugate numbers. To write the solution in a more suitable form we need a few results in elementary complex number theory, which we now recall.

(1) **Polar form of complex numbers**. Any complex number in Cartesian form $\alpha \pm i\theta$ can be written in the equivalent trigonometric form $r(\cos\omega \pm i \sin\omega)$ by a transformation from Cartesian to polar coordinates, i.e.

$$r\cos\omega = \alpha, \quad r\sin\omega = \theta, \quad r^2 = \alpha^2 + \theta^2.$$

The positive number $r = +(\alpha^2 + \theta^2)^{1/2}$ is called the *modulus* or *absolute value* of the complex number.

(2) **De Moivre's Theorem**. The relation

$$(\cos\omega \pm i\sin\omega)^n = \cos n\omega \pm i \sin n\omega$$

holds for any positive integer n.

The proof is by induction. Consider $n = 1$: the relation obviously holds. Next we show that if it holds for $n-1$, then it holds also for n; this complete the proof.

Assume that $(\cos\omega \pm i \sin\omega)^{n-1} = \cos(n-1)\omega \pm i \sin(n-1)\omega$ is true. We must prove that $(\cos\omega \pm i \sin\omega)^n = \cos n\omega \pm i \sin n\omega$ is also true. Now

$$\begin{aligned} & (\cos\omega \pm i\sin\omega)^n \\ = \; & (\cos\omega \pm i\sin\omega)(\cos\omega \pm i\sin\omega)^{n-1} \\ = \; & (\cos\omega \pm i\sin\omega)[\cos(n-1)\omega \pm i\sin(n-1)\omega]. \end{aligned}$$

Performing the multiplications (the order of the signs is of corse plus with plus and minus with minus) and rearranging terms we have

$$[\cos\omega\cos(n-1)\omega - \sin\omega\sin(n-1)\omega] \pm i[\cos\omega\sin(n-1)\omega + \cos(n-1)\omega\sin\omega].$$

Now, from elementary trigonometry, the first expression in square brackets is $\cos n\omega$ (write $\cos n\omega$ as $\cos[\omega + (n-1)\omega]$ and apply the addition formulae) and that in the second brackets is $\sin n\omega$, so that we have obtained the expression $\cos n\omega \pm i \sin n\omega$. Q.E.D.

We can now go back to our solution and write it in a more suitable form. First, change $\alpha \pm i\theta$ into $r(\cos\omega \pm i\sin\omega)$ and obtain

$$\begin{aligned} y_t &= A'[r(\cos\omega + i\sin\omega)]^t + A''[r(\cos\omega - i\sin\omega)]^t \\ &= A'r^t(\cos\omega + i\sin\omega)^t + A''r^t(\cos\omega - i\sin\omega)^t \\ &= r^t[A'(\cos\omega + i\sin\omega)^t + A''(\cos\omega - i\sin\omega)^t]. \end{aligned}$$

Then apply De Moivre's theorem and obtain

$$\begin{aligned} y_t &= r^t[A'(\cos\omega + i\sin\omega)^t + A''(\cos\omega - i\sin\omega)^t] \\ &= r^t[(A' + A'')\cos\omega t + (A' - A'')i\sin\omega t]. \end{aligned}$$

Now, A' and A'' are arbitrary complex conjugate numbers, say $a \pm ib$, where a and b are arbitrary real numbers. Then $A' + A'' = 2a$, which is a real number (call it A_1), and $(A' - A'')i = (2ib)i = 2i^2b = -2b$, which is a real number too (call it A_2). The final formula is then

$$y_t = r^t(A_1\cos\omega t + A_2\sin\omega t), \tag{5.11}$$

where r and ω are determined by the relations

$$r\cos\omega = -\frac{1}{2}a_1, \qquad r\sin\omega = \frac{1}{2}(4a_2 - a_1^2)^{1/2}. \tag{5.12}$$

An alternative form of the solution is

$$y_t = Ar^t\cos(\omega t - \epsilon), \tag{5.13}$$

where r and ω are the same as before and the arbitrary constants A and ϵ are connected to the arbitrary constants A_1 and A_2 by the transformation

$$\begin{aligned} A_1 &= A\cos\epsilon, \\ A_2 &= A\sin\epsilon. \end{aligned} \tag{5.14}$$

In fact, if we substitute (5.14) into (5.11), we have

$$y_t = r^t(A\cos\epsilon\cos\omega t + A\sin\epsilon\sin\omega t),$$

i.e.

$$y_t = Ar^t(\cos\epsilon\cos\omega t + \sin\epsilon\sin\omega t). \tag{5.15}$$

Now, from elementary trigonometry,

$$\cos\epsilon\cos\omega t + \sin\epsilon\sin\omega t = \cos(\omega t - \epsilon), \tag{5.16}$$

and so, substituting into (5.15), we obtain (5.13). This second form is perhaps easier to interpret, since it involves only one trigonometric function instead of two. The first form, however, is more suitable for the determination of the value of the arbitrary constants. In any case the resulting movement is the same: a trigonometric oscillation whose period[1] is $2\pi/\omega$ and whose amplitude will be increasing, constant or decreasing if, respectively, r is greater than equal or smaller than unity.

Of course, we must remember that t is a discontinuous variable which may take on only the values $0, 1, 2, 3$, etc., and so y_t can be represented graphically by a succession of discrete points. These point may be connected, *as a mere visual aid* (see the remarks in Chap. 3, page 20) with a continuous line, which in our case is a sinusoidal function; for some further remarks on this point, see below, exercise 5.4.1).

5.1.4 Stability conditions

We have seen above that the amplitude of the oscillation is governed by the magnitude of r. A simple formula connecting r to the coefficients of the characteristic equation can be obtained. Return to Eq. (5.12), square both members of each equation and add the corresponding members, obtaining $r^2(\cos^2\omega + \sin^2\omega) = a_2$. Since $\cos^2\omega + \sin^2\omega = 1$ for any ω, we have $r^2 = a_2$, and so

$$r = \sqrt{a_2}, \tag{5.17}$$

where the square root is to be taken with the positive sign (this is because r, the modulus or absolute value of the complex number $\alpha \pm i\theta$, is positive by definition). Now, since $\sqrt{a_2} \gtreqless 1$ according as $a_2 \gtreqless 1$, it follows that the oscillation will have an increasing, constant or decreasing amplitude according to whether $a_2 \gtreqless 1$. The *stability condition*, i.e. the necessary and sufficient condition for the oscillation to be damped, is then

$$a_2 < 1. \tag{5.18}$$

At this point we may wonder whether also in the case of real roots conditions exist by means of which we can check *on the coefficients of the characteristic equation* whether the roots are in absolute value less than unity. Such condition do exist, and it can be proved that the following inequalities:

$$\begin{aligned} 1 + a_1 + a_2 &> 0, \\ 1 - a_2 &> 0, \\ 1 - a_1 + a_2 &> 0, \end{aligned} \tag{5.19}$$

[1]The *period* of an oscillation is the interval time in which a complete oscillation takes place. This notion is defined only for *periodic* oscillations, in which distinct situations are repeated at fixed intervals of time, as is always the case with the sine and cosine functions. Aperiodic oscillations also exist, and will be treated in Chap. 26. When the oscillation is periodic, we can also define the *frequency* as the number of oscillations per unit of time, which is the reciprocal of the period.

constitute a set of necessary and sufficient conditions for the roots *-be they real or complex-* of the characteristic equation (5.5) to be less than unity in absolute value. A heuristic proof is the following.

First note that the second inequality in (5.19) coincides with inequality (5.18). The complex-roots case is then included in (5.19)[2]. For the real roots consider the graph of the parabola $f(\lambda) = \lambda^2 + a_1\lambda + a_2$. The real roots of $\lambda^2 + a_1\lambda + a_2 = 0$ are the points of intersection of $f(\lambda)$ with the λ axis. Now, we have

$$f(1) = 1 + a_1 + a_2, \qquad f(-1) = 1 - a_1 + a_2.$$

The first and third inequalities in (5.19) then mean that both $f(1)$ and $f(-1)$ are positive. This is necessary and sufficient to exclude the following cases:

(1) $+1$ and/or -1 is a root of the equation (since in this case it would be $f(1) = 0$ and/or $f(-1) = 0$);

(2) one root is less than -1, the other greater than $+1$ (in this case it would be $f(-1) < 0, f(1) < 0$, see e.g. Fig. 5.1;

(3) one root is less than -1, the other lying between -1 and $+1$ (in this case it would be $f(-1) < 0, f(1) > 0$, as the student may check graphically);

(4) one root lying between -1 and $+1$ (check graphically that in this case it would be $f(-1) > 0$, $f(1) < 0$). Only three cases then remain, all compatible with $f(-1)$ and $f(1)$ both positive (check graphically):

(a) both roots negative and less than -1;

(b) both roots positive and greater than 1;

(c) both roots lying between -1 and $+1$.

Now, in cases (a) and (b) the product of the roots is positive and greater than unity. Since the product of the roots is a_2 (as the reader may check by straightforward multiplication of the explicit expression for λ_1 and λ_2), the second inequality in (5.19) is necessary and sufficient to exclude cases (a) and (b). Only case (c) is then possible, and this completes the proof.

We should like to stress the importance of the stability conditions (5.19), since they are of great help in the analysis of the qualitative behaviour of second-order economic models (see the next chapter).

5.2 Solution of the non-homogeneous equation

All that we need in this case is the particular solution of the non-homogeneous equation. We shall exemplify the general method of undetermined coefficients

[2]Note that, when the roots are complex, the first and third inequalities in (5.19) are automatically satisfied, so that they do not impose any additional restraint. In fact, when the roots are complex, the parabola $f(\lambda) = \lambda^2 + a_1\lambda + a_2$ lies wholly in the positive half plane above the λ axis. This means that $f(-1)$, which is $1 - a_1 + a_2$, and $f(1)$,which is $1 + a_1 + a_2$, are both positive quantities.

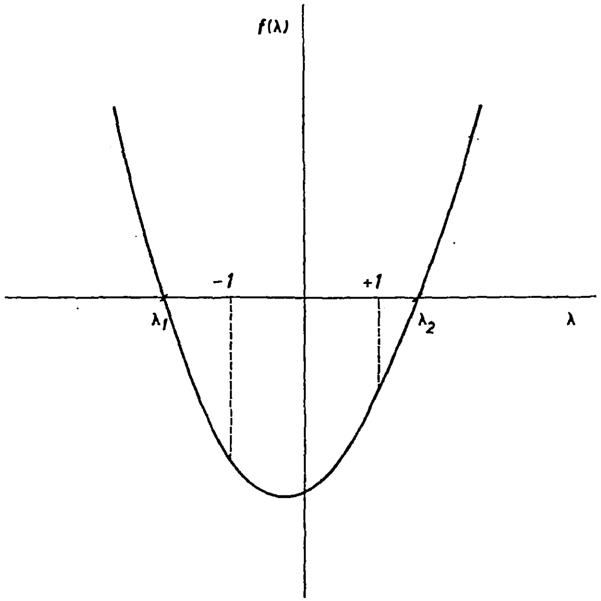


Figure 5.1: Second-order difference equations: real roots

(see 2.2.2.) in the case in which $g(t) = G$, a constant. In other cases the student may proceed along the same lines as in the examples we worked out dealing with first-order equations (see Chap. 3, Sect. 3.2).

Assume then that

$$c_2 y_t + c_1 y_{t-1} + c_0 y_{t-2} = G. \tag{5.20}$$

As a particular solution try $\bar{y}_t = B$, where B is an undetermined constant. Direct substitution in (5.20) gives $c_2 B + c_1 B + c_0 B = G$, from which

$$B = \frac{G}{c_0 + c_1 + c_2}.$$

If $c_0 + c_1 + c_2 = 0$, try $\bar{y}_t = Bt$. Substitution in (5.20) gives

$$c_2 Bt + c_1 B(t-1) + c_0 B(t-2) = G;$$

therefore

$$(c_2 + c_1 + c_0)Bt - B(c_1 + 2c_0) = G,$$

which gives

$$B = \frac{-G}{c_1 + 2c_0}.$$

If also $c_1 + 2c_0 = 0$, try $\bar{y}_t = Bt^2$. Substitution in (5.20) yields, after some manipulation

$$B = \frac{G}{2c_0}.$$

Since c_0 must be different from zero if the equation is of the second order, no further complication may arise.

The particular solution of the non-homogeneous equation may usually be interpreted, as we shall see in the economic applications as the (stationary or moving) equilibrium of the variable y_t.

5.2.1 The operational method

Let us now consider the case in which the functional form of $g(t)$ is not known, and apply the operational calculus, already introduced in Sect. 3.2.6. We now have

$$c_2 y_t + c_1 y_{t-1} + c_0 y_{t-2} = x_t, \tag{5.21}$$

from which, dividing through by c_2 and letting $X_t \equiv (1/c_2)x_t$,

$$y_t + a_1 y_{t-1} + a_2 y_{t-2} = X_t,$$

namely, in operational notation,

$$(1 + a_1 L + a_2 L^2) y_t = X_t. \tag{5.22}$$

Thus a particular solution to Eq. (5.22) is given by

$$\bar{y}_t = \frac{1}{1 + a_1 L + a_2 L^2} X_t. \tag{5.23}$$

As we have already recalled in Sect. 3.2.6, operators can be treated for our purposes as ordinary algebraic quantities. Thus we can write the polynomial $1 + a_1 L + a_2 L^2$ as $L^2(F^2 + a_1 F + a_2)$ where $F \equiv L^{-1}$ can be seen as the forward operator (inverse lag operator). We now observe that $(F^2 + a_1 F + a_2)$ can be factorised into $(F - F_1)(F - F_2)$, where F_1, F_2 are the roots of $F^2 + a_1 F + a_2 = 0$. Note that this equation coincides with the characteristic equation (5.5), hence F_1, F_2 coincide with λ_1, λ_2, the roots of the characteristic equation.

Thus we have

$$\begin{aligned} 1 + a_1 L + a_2 L^2 &= L^2 (F - \lambda_1)(F - \lambda_2) \\ &= (LF - \lambda_1 L)(LF - \lambda_2 L) \\ &= (1 - \lambda_1 L)(1 - \lambda_2 L). \end{aligned} \tag{5.24}$$

We can now expand the rational polynomial in L on the r.h.s. of Eq. (5.23) in partial fractions (for the partial fraction expansion see any calculus textbook or Turnbull, 1957, Sect. 19), that is, assuming distinct roots,

$$\frac{1}{1+a_1L+a_2L^2} = \frac{\theta_1}{1-\lambda_1 L} + \frac{\theta_2}{1-\lambda_2 L}, \tag{5.25}$$

where it turns out that

$$\begin{aligned} \theta_1 &= \frac{\lambda_1}{\lambda_1-\lambda_2}, \\ \theta_2 &= \frac{-\lambda_2}{\lambda_1-\lambda_2}, \end{aligned} \tag{5.26}$$

as the reader may wish to check by direct substitution. Thus the particular solution (5.23) becomes

$$\overline{y}_t = \sum_{r=1}^{2} \frac{\theta_r}{1-\lambda_r L} X_t, \tag{5.27}$$

namely, by using the expansion of $(1-\lambda_r L)^{-1}$ derived in Sect. 3.2.6, Eq. (3.33),

$$\overline{y}_t = \sum_{r=1}^{2} \theta_r \sum_{i=0}^{\infty} (\lambda_r)^i X_{t-i}. \tag{5.28}$$

Equation (5.28) is the backward solution, to be used when $|\lambda_r| < 1$. In the contrary case we shall use the forward solution, namely the alternative expansion explained in Chap. 3, Eq. (3.38). If one root is stable and the other unstable, we shall of course use the respective appropriate expansion.

Equation (5.25) holds independently of the nature of the roots (real or complex), but cannot be applied in the case of equal roots, $\lambda_1 = \lambda_2 = \lambda^*$. In this case we have $1 + a_1L + a_2L^2 = (1-\lambda^*L)(1-\lambda^*L)$ and the expansion is

$$\begin{aligned} \frac{1}{(1-\lambda^*L)}\frac{1}{(1-\lambda^*L)} &= (1+\lambda^*L+\lambda^{*^2}L^2+...)(1+\lambda^*L+\lambda^{*^2}L^2+...) \\ &= (1+\lambda^*L+\lambda^{*^2}L^2+...)+\lambda^*L(1+\lambda^*L+\lambda^{*^2}L^2+...) \\ &\quad +\lambda^{*^2}L^2(1+\lambda^*L+\lambda^{*^2}L^2+...)+... \\ &= 1+2\lambda^*L+3\lambda^{*^2}L^2+... \\ &= \sum_{i=0}^{\infty}(i+1)\lambda^{*^i}L^i, \end{aligned} \tag{5.29}$$

that can be used to obtain the particular solution

$$\overline{y}_t = \sum_{i=0}^{\infty}(i+1)\lambda^{*^i}X_{t-i}. \tag{5.30}$$

5.3 Determination of the arbitrary constants

In the general solution, as we have seen above, two arbitrary constants appear. In order to determine them, we need two additional conditions. These conditions will take the form

$$\begin{aligned} y_t &= y^* \quad \text{for} \quad t = t^*, \\ y_t &= y^{**} \quad \text{for} \quad t = t^{**}, \end{aligned}$$

where t^*, t^{**}, y^*, y^{**} are all known values. Substituting these values in the general solution, we obtain a set of two linear equations in the two unknown values of A_1, A_2. The solution of this system (that always exists, as we have shown in Chap. 2, Sect. 2.3) yields the values of the two arbitrary constants which satisfy the additional conditions.

Let us note that the determination of the arbitrary constants is to be made according to the general solution of the equation under consideration (i.e., if the equation is non-homogeneous, its own general solution, and not the general solution of the corresponding homogeneous equation, must be used).

In principle, the values t^* and t^{**} may be any two different values of t. In economic models, however, the additional conditions are usually given for $t = 0$ and $t = 1$ and this is why they are called 'initial' conditions.

Let us end the study of the second-order equations with a final warning. When the numerical values of the coefficients c_0, c_1, c_2, of the parameters appearing in $g(t)$, and of the initial conditions are known, the successive values of y_t can be computed recursively, without any apparent need to find the general solution of the equation. The warning is that, after computing the succession of y_t up to some t, it may be dangerous to extrapolate the behaviour we have observed, so that to avoid mistakes it is better to find the solution of the equation also in this case. For example, consider the equation

$$y_t + 1.8y_{t-1} + 0.8y_{t-2} = 0, \qquad y_0 = 0, \qquad y_1 = -2.$$

To compute the successive values of y_t recursively, write the equation in the form

$$y_t = -1.8y_{t-1} - 0.8y_{t-2}.$$

Then

$$\begin{aligned} y_2 &= -1.8y_1 - 0.8y_0 = -1.8 \times (-2) - 0.8 \times 0 = 3.6, \\ y_3 &= -1.8 \times 3.6 - 0.8 \times (-2) = -4.88, \\ y_4 &= -1.8 \times (-4.88) - 0.8 \times 3.6 = 5.904, \\ y_5 &= -1.8 \times 5.904 - 0.8 \times (-4.88) = -6.7232. \end{aligned}$$

The values alternate in sign and are increasing in absolute value. We might then be tempted to say that the movement is an improper oscillation of increasing amplitude, i.e. a divergent movement. Let us now check by solving the difference equation. Its characteristic equation is

$$\lambda^2 + 1.8\lambda + 0.8 = 0,$$

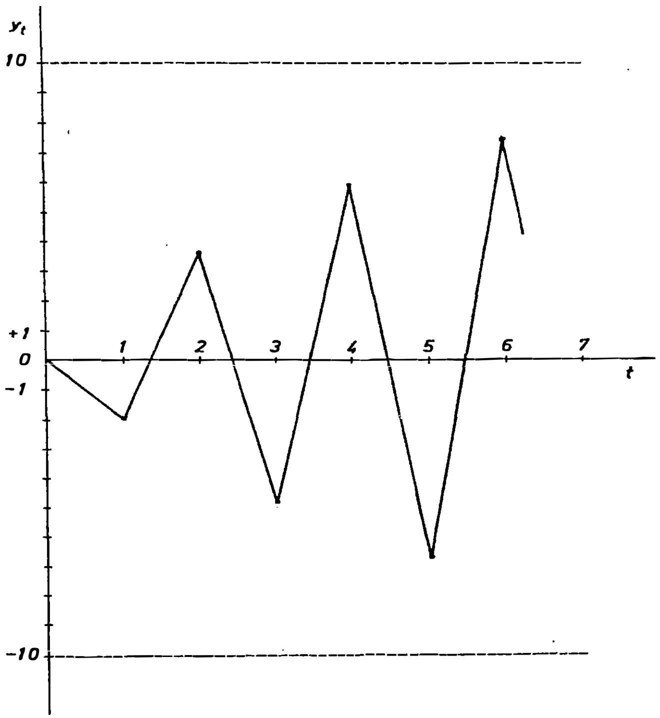


Figure 5.2: An apparently divergent solution

whose roots are -1, -0.8. The general solution of the difference equation under consideration is then

$$y_t = A_1(-1)^t + A_2(-0.8)^t.$$

The arbitrary constants are now determined according to the initial conditions. We have

$$\begin{aligned} y_0 &= A_1(-1)^0 + A_2(-0.8)^0 = A_1 + A_2 = 0, \\ y_1 &= A_1(-1)^1 + A_2(-0.8)^1 = -A_1 - 0.8A_2 = -2. \end{aligned}$$

The solution of the system

$$\begin{aligned} A_1 + A_2 \quad &= \; 0, \\ -A_1 - 0.8A_2 &= \; -2, \end{aligned}$$

yields $A_1 = 10, A_2 = -10$, and so

$$y_t = 10(-1)^t - 10(-0.8)^t.$$

From this equation we can see that, as t increases, y_t tends to an improper oscillation of constant amplitude, given by the term $10(-1)^t$, since the term $-10(-0.8)^t$ tends to zero. The overall fluctuation is actually of increasing amplitude, but it does not diverge, since it tends to a limit cycle of constant amplitude. Fig. 5.2 shows the actual fluctuation and its limits (broken lines). The inference we were tempted to make on the basis of the succession of values computed recursively is clearly wrong.

Other examples could be shown, e.g. the case of a trigonometric oscillation whose period is relatively great, so that it may happen that the value of t at which we stop our recursive computations is situated before that value of t where the first turning point in y_t occurs, so that the data leads us to think that the movement is monotonic. But we think that the warning is now sufficiently clear.

5.4 Exercises

5.4.1 Example

Let us solve the following difference equation

$$y_t - \sqrt{3}y_{t-1} + y_{t-2} = 100. \tag{5.31}$$

A particular solution of (5.31) is

$$\bar{y}_t = \frac{100}{2-\sqrt{3}}. \tag{5.32}$$

The homogeneous equation corresponding to the non-homogeneous equation (5.31) has the characteristic equation

$$\lambda^2 - \sqrt{3}\lambda + 1 = 0, \tag{5.33}$$

whose roots are

$$\lambda_1, \lambda_2 = \frac{1}{2}(\sqrt{3} \pm \sqrt{-1}) = \frac{1}{2}\sqrt{3} \pm \frac{1}{2}\mathrm{i}.$$

Since $r = \sqrt{1} = 1$, the general solution of the homogeneous equation is

$$y_t = A_1 \cos \omega t + A_2 \sin \omega t, \tag{5.34}$$

where

$$\cos\omega = \tfrac{1}{2}\sqrt{3},$$

$$\sin\omega = \tfrac{1}{2}.$$

From the trigonometric tables we find that the angle whose sine is $\frac{1}{2}$ and whose cosine is $\frac{1}{2}\sqrt{3}$, is 30°. Then Eq. (5.34) becomes

$$y_t = A_1 \cos 30°t + A_2 \sin 30°t,$$

and the general solution of Eq. (5.31) is

$$y_t = A_1 \cos 30°t + A_2 \sin 30°t + \frac{100}{2-\sqrt{3}}. \tag{5.35}$$

Given the initial conditions $y_0 = 0, y_1 = 100$ we have, substituting in Eq. (5.35),

$$\begin{aligned} 0 &= A_1 \cos 0 + A_2 \sin 0 + \frac{100}{2-\sqrt{3}}, \\ 100 &= A_1 \cos 30° + A_2 \sin 30° + \frac{100}{2-\sqrt{3}}. \end{aligned}$$

Since $\cos 0 = 1, \sin 0 = 0$, from the first equation we obtain immediately $A_1 = -100/(2-\sqrt{3})$. Substituting into the second equation this value and also the values of $\cos 30°$ and $\sin 30°$, we have

$$\frac{-100\sqrt{3}}{(2-\sqrt{3})2} + \frac{1}{2}A_2 + \frac{100}{2-\sqrt{3}} = 100,$$

which gives, after some simple manipulation, $A_2 = 100$. Eq. (5.35) then becomes

$$y_t = \frac{-100}{2-\sqrt{3}} \cos 30°t + 100 \sin 30°t + \frac{100}{2-\sqrt{3}}. \tag{5.36}$$

If we want to put the solution into the alternative form mentioned above, Eq. (5.13), we use the transformation

$$\begin{aligned} -100/(2-\sqrt{3}) &= A \cos \epsilon, \\ 100 &= A \sin \epsilon, \end{aligned}$$

so that $A = +(A_1^2 + A_2^2)^{1/2}$ and $\tan \epsilon = -(2-\sqrt{3})$. The result is $A = 200/(2-\sqrt{3})^{1/2}$ and $\epsilon = 165°$. The alternative form of the solution is then

$$y_t = \frac{200}{(2-\sqrt{3})^{1/2}} \cos(30°t - 165°) + \frac{100}{2-\sqrt{3}}. \tag{5.37}$$

As a check we may compute y_3. Using Eq.(5.36) we have (remember that $\sin 90° = 1$ and $\cos 90° = 0$)

$$y_3 = 100 + \frac{100}{2-\sqrt{3}} \simeq 100 + 373.205 = 473.205.$$

Table 5.1: Constant-amplitude trigonometric oscillation

t	0	1	2	3	4
y_t	0	100	273.205	473.205	646.410
t	5	6	7	8	9
y_t	746.410	746.410	646.410	473.205	273.205
t	10	11	12	13	14
y_t	100	0	0	100	273.205

Using Eq. (5.37) we have

$$y_3 = \frac{200}{(2-\sqrt{3})^{1/2}} \cos(-75°) + \frac{100}{2-\sqrt{3}}.$$

Now, $\cos(-75°) = \cos 75° = \frac{1}{4}\sqrt{2}(\sqrt{3}-1)$. Substituting in the last equation and using the fact that $(\sqrt{3}-1)\sqrt{2} = 2(2-\sqrt{3})^{1/2}$, we have

$$y_3 = 100 + \frac{100}{2-\sqrt{3}} \simeq 473.205,$$

as before. The succession of points arising from the solution of this difference equation is given in Table 5.1 up to $t = 14$.

This exercise offers the opportunity for a few additional considerations on the nature of trigonometric oscillations arising from complex roots of difference equations. When the roots of the characteristic equations of a second-order difference equation are complex conjugate with modulus 1, we know that the resulting movement is a constant-amplitude trigonometric oscillation. However, in these cases the recursive computation of y_t will usually show an oscillation whose amplitude is not precisely constant (take, for example, exercise $a.(v)$ below, let $y_0 = 0, y_1 = 100$, and compute the solution recursively). The reason is that the oscillation would be of constant amplitude if time were considered as a continuous variable in the sinusoidal function $A_1 \cos \omega t + A_2 \sin \omega t$. But, since t can take on only the discrete values $0, 1, 2, 3, ...$, only the corresponding points of the underlying continuous sinusoidal function can be considered. Now, there is no reason why the turning points of the actual succession should always coincide with the peaks and troughs of the sinusoidal function. Fig. 5.3 may serve as an illustration (T_1, T_2, T_3 are the turning points of the actual succession).

Thus, the statement made in Sect. (5.1.3), that when the roots of the characteristic equation are complex and $r = 1$ a constant-amplitude oscillation follows, must be qualified: what happens is that the points representing the succession of the values of y_t will lie on a sinusoidal function whose amplitude is constant. Therefore the actual oscillation as t increases can be neither explosive nor damped, although it may not show an exactly constant amplitude. As a matter of terminology, we shall continue to use the expression 'constant-amplitude oscillation' according to the meaning clarified here.

A similar problem arises, of course, when $r \neq 1$ (explosive or damped oscillations), but in such cases it is not apparent, since the fact that the turning points of the succession may not coincide with those of the sinusoidal function is of no consequence as regards the increasing or decreasing amplitude of the oscillation shown by the succession.

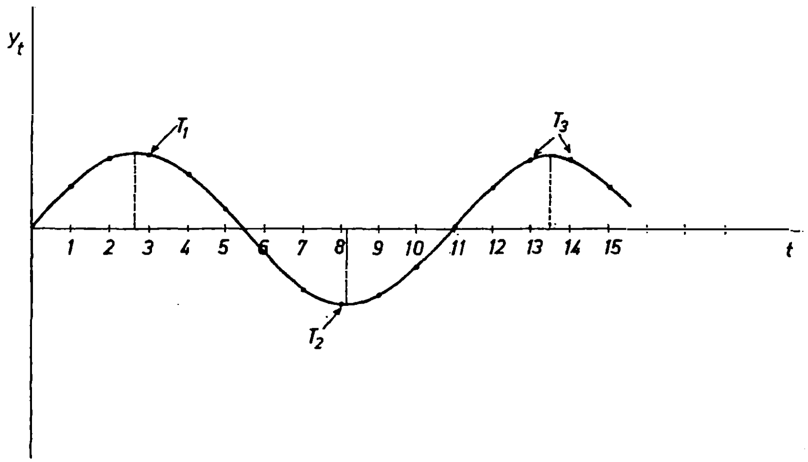


Figure 5.3: A constant-amplitude oscillation

5.4.2 Other exercises

(a) Find the solution of the following non-homogeneous difference equations:

$$\begin{array}{rl} (i) & y_t = 0.8y_{t-1} - 2.4y_{t-2} + 100 \\ (ii) & y_t = 1.1y_{t-1} - 0.6y_{t-2} + 1100 \\ (iii) & y_t = -0.25y_{t-2} + 225 \\ (iv) & y_t = 4.6y_{t-1} - 5.6y_{t-2} \\ (v) & y_t = 1.8y_{t-1} - y_{t-2} + 100 \end{array}$$

(b) Find the solution of the following homogeneous difference equations:

$$\begin{array}{rl} (i) & y_{t+2} - 3y_{t+1} + 2y_t = 0 \quad y_0 = 2, y_1 = 3 \\ (ii) & y_{t+2} - 6y_{t+1} + 9y_t = 0 \quad y_0 = 3, y_1 = 2 \end{array}$$

and determine the values of the arbitrary constants.

5.5 References

Allen, R.G.D., 1959, *Mathematical Economics*, Chap. 4, Chap. 6.

Baumol, W.J., 1958, Topology of Second Order Difference Equations with Constant Coefficients.

Baumol, W.J., 1970, *Economics Dynamics*, Chaps. 10-11.

Goldberg, S., 1958, *Introduction to Difference Equations*, Chaps. 3-4.

Kenkel, J.L., 1974, *Dynamic Linear Economic Models*, Chap. 5.

Turnbull, H.W., 1957, *Theory of Equations*, Chap. III, Sect. 19.

Chapter 6

Second-order Difference Equations in Economic Models

6.1 Multiplier-accelerator interaction: the prototype model

This model, which was built by Samuelson (1939) following a suggestion by Hansen, may justly be considered as the pioneer of all multiplier-accelerator models of income determination and the business cycle. The 'ingredients' of such models are a consumption function, an investment function (in which both induced and autonomous investment appear) and the relation which defines the equilibrium value of income. In the model under consideration we have the following equations:

$$C_t = bY_{t-1}, \qquad 0 < b < 1, \tag{6.1}$$

where consumption depends on national income with a one-period lag. The constant b is the (marginal and average) propensity to consume. As far as investment is concerned, we distinguish between induced investment $I_t^{'}$ and autonomous investment $I_t^{''}$. Then

$$I_t = I_t^{'} + I_t^{''}, \tag{6.2}$$

where I_t is total investment.

Autonomous investment (essentially public expenditure) is assumed constant:

$$I_t^{''} = G, \tag{6.3}$$

where G is a positive constant.

Induced investment depends on the variation in the demand for consumption goods, according to the acceleration principle:

$$I_t' = k(C_t - C_{t-1}), \tag{6.4}$$

where k is the acceleration coefficient (the 'relation' in Hansen's terminology).

The equilibrium condition

$$Y_t = C_t + I_t \tag{6.5}$$

closes the model.

Simple substitutions yield the following second-order equation

$$Y_t - b(1+k)Y_{t-1} + bkY_{t-2} = G. \tag{6.6}$$

The solution of this functional equation gives the behaviour over time of national income; substitution in Eqs. (6.1) and (6.4) will then give the behaviour over time of consumption and of induced investment.

A particular solution of Eq. (6.6) is easily found by trying $\bar{Y}_t =$ constant, from which

$$\bar{Y} = \frac{G}{1-b}. \tag{6.7}$$

This is the value we obtain applying the multiplier $1/(1-b)$ to autonomous expenditure G. The particular solution determines the (stationary) equilibrium value of national income.

The deviations from this value will be given by the general solution of the homogeneous equation corresponding to Eq. (6.6), i.e. of

$$Y_t - b(1+k)Y_{t-1} + bkY_{t-2} = 0.$$

The characteristic equation is

$$\lambda^2 - b(1+k)\lambda + bk = 0. \tag{6.8}$$

A qualitative analysis of Eq. (6.8) will now be made. As a first step, we apply the stability conditions (see (5.19) in the previous chapter). We obtain

$$\begin{aligned} &1 - b(1+k) + bk = 1 - b > 0,\\ &1 - bk > 0,\\ &1 + b(1+k) + bk > 0. \end{aligned} \tag{6.9}$$

The first inequality is satisfied since we have assumed that the marginal propensity to consume is less than unity (and this is an empirically plausible assumption); the third inequality is satisfied too, the left-hand side being a sum of quantities which are all positive. The crucial inequality is then the second one. Therefore we can say that the stability condition is

$$bk < 1$$

or

$$b < 1/k. \tag{6.10}$$

In order to ascertain the type of the movement (monotonic, oscillatory, etc.), let us begin to note that the succession of the signs of the coefficients of Eq. (6.8) is $+-+$. This means (Descartes' rule of signs) that no negative root may occur. We can then exclude movements involving 'improper' oscillations. The next step is to compute the discriminant of Eq. (6.8). It is

$$\Delta = b^2(1+k)^2 - 4bk, \tag{6.11}$$

and so $\Delta \gtreqless 0$ if $b^2(1+k)^2 \gtreqless 4bk$, whence

$$\Delta \gtreqless 0 \qquad as \qquad b \gtreqless \frac{4k}{(1+k)^2}. \tag{6.12}$$

6.1.1 Graphical location of the roots

Putting together (6.10) and (6.12) we have all possible cases, which can be conveniently plotted in a diagram (see Fig. 17.1), due to Samuelson.

In the diagram the function $b = 4k/(1+k)^2$ (the curve OPS) and the function $b = 1/k$ (the curve PQ) are plotted. Since $b < 1$, we are interested only in that part of the positive quadrant which is below the broken line (this is why we have not drawn the upper part of the rectangular hyperbola $b = 1/k$). Four regions are then singled out (remember that the real roots are positive).

Region A. Any point in this region lies *below* the function $b = 1/k$ and *above* the function $b = 4k/(1+k)^2$. This is the graphical counterpart of the inequalities $b < 1/k, b > 4k/(1+k)^2$. Therefore by (6.10) and (6.12) the stability condition is satisfied, and the roots are real. The system will show a monotonic movement converging towards the equilibrium value $G/(1-b)$.

Region B. Any point of this region satisfies the inequalities $b < 1/k$, $b < 4k/(1+k)^2$. The stability condition is satisfied and the roots are complex. The result is a damped oscillation around the equilibrium value.

Region C. Here we have $b > 1/k, b < 4k/(1+k)^2$. The stability condition is not satisfied and the roots are complex. The result is an explosive oscillation around the equilibrium value.

Region D. In this region $b > 1/k, b > 4k/(1+k)^2$. The stability condition is not satisfied and the roots are complex. The result is monotonically

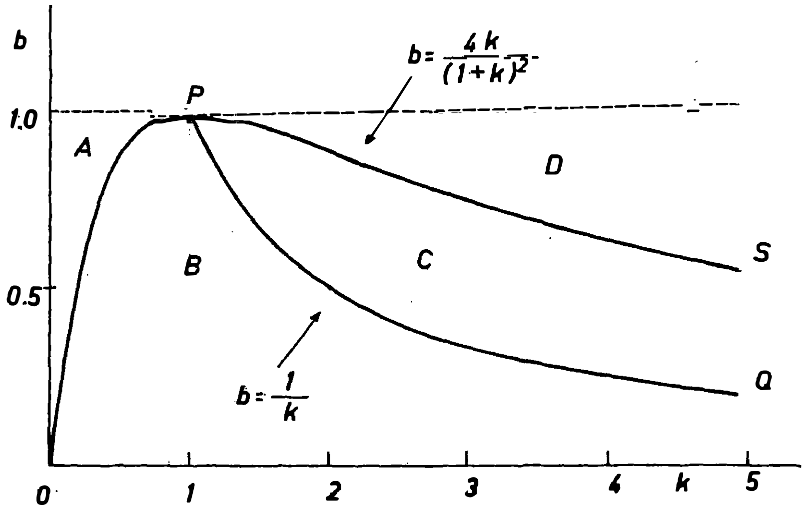


Figure 6.1: Samuelson's multiplier-accelerator diagram

explosive.

In order to complete our analysis we have to examine the points which happen to lie on the boundary lines demarcating the four regions. Since b and k are both positive in this model, and $b < 1$, we must exclude from consideration the points on the b axis, the points on the k axis, the origin, and point P. We then have the following particular cases:

(1) Points on the demarcation line between region A and region B. Here we have $b = 4k/(1+k)^2$, i.e. a real root of multiplicity two. Since the stability condition $b < 1/k$ is satisfied, this root is smaller than unity. Therefore the function $(A_1 + A_2 t)\lambda^{*t}$ will be dominated, as $t \to \infty$, by the component λ^{*t} and the movement will eventually converge monotonically toward equilibrium.

(2) Points falling on the demarcation line between regions B and C. We have $b < 4k/(1+k)^2$ and $bk = 1$, i.e. complex roots with unit modulus which imply a constant-amplitude oscillation, that is to say a movement separating stability from instability.

(3) Points falling on the line demarcating region C and region D. Here we have $b = 4k/(1+k)^2$, $b > 1/k$, i.e. a root with multiplicity two and greater than unity. The result is a movement which will be eventually dominated by the monotonically divergent term λ^{*t}.

6.2 Market adjustments and rational expectations

Consider an isolated market with an output lag. Current demand for consumption purposes C_t is assumed to depend on current price p_t, while current production P_t, due to the output lag, depends on the price p_t^e that was expected to hold in the current period. This is the setting we already met in the exposition of the cobweb model in Sect. 4.1. We now add the assumption that the commodity is non-perishable, so that inventories I_t of it can exist, and are in fact held for speculative purposes, i.e. to profit from expected changes in prices. For the sake of simplicity, storage and other costs are assumed to be negligible. Hence we have the following model, due to Muth (1961), where all the variables are measured as deviations from equilibrium:

$$\begin{aligned} C_t &= -\beta p_t, \\ P_t &= \gamma p_t^e + x_t, \\ I_t &= \alpha(p_{t+1}^e - p_t), \\ C_t + I_t &= P_t + I_{t-1}, \end{aligned} \tag{6.13}$$

where α, β, γ are positive constants and x_t represents the effect of exogenous factor (such as the weather) on supply. The fourth equation is the market-clearing condition: in any period total demand is the sum of demand for current consumption plus total inventory demand, and total supply equals current production plus the stock of inventories carried over from the previous period. Alternatively, we can put the *change* of inventories on the demand side and count just current production on the supply side, writing the market-clearing condition as $C_t + (I_t - I_{t-1}) = P_t$, that of course is formally equivalent.

Let us now assume rational expectations. These, as is well known, mean perfect foresight in a deterministic context (for a general treatment of rational expectations see Chap. 22, Sect. 22.3.1). Thus $p_t^e = p_t$, $p_{t+1}^e = p_{t+1}$, and by simple substitutions system (6.13) yields

$$\alpha p_{t+1} - (2\alpha + \beta + \gamma)p_t + \alpha p_{t-1} = x_t. \tag{6.14}$$

Let us now shift all the time subscripts backwards by one unit, divide through by α and define $X_t \equiv (1/\alpha)x_{t-1}$, thus obtaining the non-homogeneous difference equation

$$p_t - \frac{2\alpha + \beta + \gamma}{\alpha} p_{t-1} + p_{t-2} = X_t. \tag{6.15}$$

Let begin with the homogeneous part, whose characteristic equation is

$$\lambda^2 - \frac{2\alpha + \beta + \gamma}{\alpha}\lambda + 1 = 0. \tag{6.16}$$

We have

$$\Delta = \frac{(2\alpha+\beta+\gamma)^2 - 4\alpha^2}{\alpha^2},$$

which is clearly positive. Hence there will be two real roots that will be both positive since the succession of signs in the coefficients of Eq. (6.16) is $+-+$. It can also be shown that the roots will be a reciprocal pair. In fact, the product of the roots is equal to the constant term, hence $\lambda_1\lambda_2 = 1$ and $\lambda_1 = 1/\lambda_2$. Note that since the sum of the roots equals minus the coefficient of λ, they cannot be both unity. It follows that the movement will show an unstable behaviour.

Let us now come to the particular solution of the homogeneous equation. Since we do not know the form of x_t and hence of X_t, we apply the operational method explained in Sect. 5.2. Thus a particular solution is

$$\overline{p}_t = \frac{1}{1 - \frac{2\alpha+\beta+\gamma}{\alpha}L + L^2} X_t = \left(\frac{\theta_1}{1-\lambda_1 L} + \frac{\theta_2}{1-\lambda_2 L}\right) X_t, \qquad (6.17)$$

where the λ's are the roots of Eq. (6.16). Without loss of generality we can suppose that λ_1 is the root which is smaller than unity. Hence we shall use the backward expansion for $(1-\lambda_1 L)^{-1}$ and the forward expansion for $(1-\lambda_2 L)^{-1}$. Thus our particular solution becomes

$$\begin{aligned} \overline{p}_t &= \theta_1 \sum_{i=0}^{\infty} \lambda_1^i X_{t-i} - \theta_2 \sum_{i=1}^{\infty} \left(\frac{1}{\lambda_2}\right)^i X_{t+i} \qquad (6.18) \\ &= \theta_1 \sum_{i=0}^{\infty} \lambda_1^i X_{t-i} - \theta_2 \sum_{i=1}^{\infty} \lambda_1^i X_{t+i}, \end{aligned}$$

which expresses $\overline{p}_t$ as a two-sided distributed lag of $X_t \equiv (1/\alpha)x_{t-1}$, namely as a weighted sum of past, present, and future values of the exogenous shock.

Thus we have the general solution

$$p_t = A_1\lambda_1^t + A_2\lambda_2^t + \overline{p}_t = A_1\lambda_1^t + A_2\lambda_1^{-t} + \overline{p}_t,$$

where the arbitrary constants are to be determined on the basis of two side conditions on the path of p_t. If, for example, we imposed boundedness on the whole path of p_t for all bounded sequences of the exogenous shock $\{X_t\}$, namely $\lim_{t\to-\infty}|p_t| < \infty$, $\lim_{t\to+\infty}|p_t| < \infty$, then clearly $A_1 = A_2 = 0$. In fact, given $0 < \lambda_1 < 1$, λ_1^t tends to zero for $t \to +\infty$ but tends to ∞ for $t \to -\infty$. Conversely, λ_1^{-t} tends to ∞ for $t \to +\infty$ and to zero for $t \to -\infty$. Hence both A_1 and A_2 must be zero.

6.3 Exercises

1. Suppose that, in Samuelson's model, consumption is a lagged function of disposable income, namely $C_t = bY_{d_{t-1}}, Y_{d_t} = Y_t - T_t, T_t = \tau Y_t$, where $0 < \tau < 1$ is the constant tax rate.

(1.a) Why does the introduction of taxation exert a stabilizing influence?

(1.b) Suppose now that the government spends all its tax receipts plus a constant amount G, namely $I_t'' = G + T_t$. Discuss the stability condition. Consider also the case in which there is a lag between tax receipts and expenditure, whence $I_t'' = G + T_{t-1}$.

2. Consider Hicks' extension (Hicks, 1950) of Samuelson's multiplier-accelerator model:

$$\begin{aligned} Y_t &= C_t + I_t, \\ C_t &= bY_{t-1}, \\ I_t &= I_t' + I_t'', \\ I_t'' &= A_0(1+g)^t, \\ I_t' &= k(Y_{t-1} - Y_{t-2}), \end{aligned} \qquad (6.19)$$

where the following innovations should be noted. First, autonomous investment is assumed to increase over time at the constant rate of growth g. Second, accelerator-induced investment does not depend solely on the variations in consumption demand, but on the variations in total demand. Third, the variations that induce investment are lagged one period, i.e. $Y_{t-1} - Y_{t-2}$ and not $Y_t - Y_{t-1}$. By simple substitutions the model can be reduced to the following second-order non-homogeneous difference equation:

$$Y_t - (b+k)Y_{t-1} + kY_{t-2} = A_0(1+g)^t.$$

(2.a) Show that a particular solution of the non-homogeneous equation is

$$\overline{Y}_t = \frac{A_0(1+g)^2}{(1+g)^2 - (b+k)(1+g) + k}(1+g)^t,$$

and show that the denominator in this fraction is always positive (which is required for the particular solution to be economically meaningful) when the roots of the characteristic equation of the corresponding homogeneous equation are complex (hint: the denominator under consideration coincides with the characteristic equation when $1+g = \lambda$).

(2.b) Show that, when the roots of the characteristic equation are real, the denominator considered in the previous exercise is positive provided that g does *not* lie in the interval $\lambda_1 \le 1+g \le \lambda_2$, where λ_1, λ_2 are the (possible) real roots.

(2.c) Prove that the characteristic equation admits no negative real roots.

(2.d) Show that the crucial stability condition for the homogeneous equation is $k < 1$. Why is it more stringent than the stability condition ($bk < 1$) in Samuelson's model?

(2.e) Consider the discriminant of the characteristic equation, $\Delta = k^2 - (4-2b)k + b^2$, and determine the regions listed in Table 6.1 ($s \equiv 1 - b$).

Table 6.1: Roots of Hicks' model

If	Then	Roots
$k < (1-\sqrt{s})^2$	$\Delta > 0$	Real and distinct
$k = (1-\sqrt{s})^2$	$\Delta = 0$	Real and equal
$(1-\sqrt{s})^2 < k < (1+\sqrt{s})^2$	$\Delta < 0$	Complex conjugate
$k = (1+\sqrt{s})^2$	$\Delta = 0$	Real and equal
$k > (1+\sqrt{s})^2$	$\Delta > 0$	Real and distinct

(Hint: using the definition of s, it can be seen that $\Delta = 0$ has the roots $k_1, k_2 = 1 + s \pm 2\sqrt{s} = (1 \pm \sqrt{s})^2$. Then use the theory of elementary second-degree inequalities to show that $\Delta < 0$ for k lying in the interval between k_1, k_2, etc.).

(2.f) Using the results in exercises 2.d and 2.e, determine the nature of the movement around the trend (particular solution) in the four regions (and on the lines of demarcation between them) shown in Fig. 6.2.

(2.g) Since the values of the acceleration coefficient are normally much greater than one, the model is unstable. Hicks further assumed that not only the absolute deviations from the moving equilibrium solution are divergent, but also the *relative* deviations are. The relative deviations are defined as the ratio of the absolute deviations from the particular solution to the particular solution itself, i.e. $D_t \equiv Y_t / Y_0(1+g)^t$, where Y_t is given by the solution of the homogeneous equation. Consider then the homogeneous difference equation

$$\frac{1}{Y_0(1+g)^t}[Y_t - (b+k)Y_{t-1} + kY_{t-2}] = 0,$$

reduce it to a homogeneous difference equation in D_t, D_{t-1}, D_{t-2} and show that the *in*stability condition is $k/(1+g)^2 > 1$.

3. Consider the introduction of foreign trade into model (6.19). Imports are a function of income lagged one period, $M_t = mY_{t-1}$, $0 < m < 1$. Exports are determined by foreign demand, which is assumed to grow at a constant proportional rate g_x, so that $X_t = X_0(1+g_x)^t$.

(3.a) Find the particular solution and the conditions under which Y_0 is positive. Is there balance-of-payments equilibrium ($X_t = M_t$) along this solution?

(3.b) Discuss the stability conditions. Why is the crucial stability condition not affected?

4. Modify the assumptions of exercise 3 in the sense that imports are determined by unlagged consumption and investment demand separately considered, i.e. $M_t = m_1 C_t + m_2 I_t, 0 < m_1 < 1, 0 < m_2 < 1$. Exports are $X_t = X_0(1+g_x)^t$.

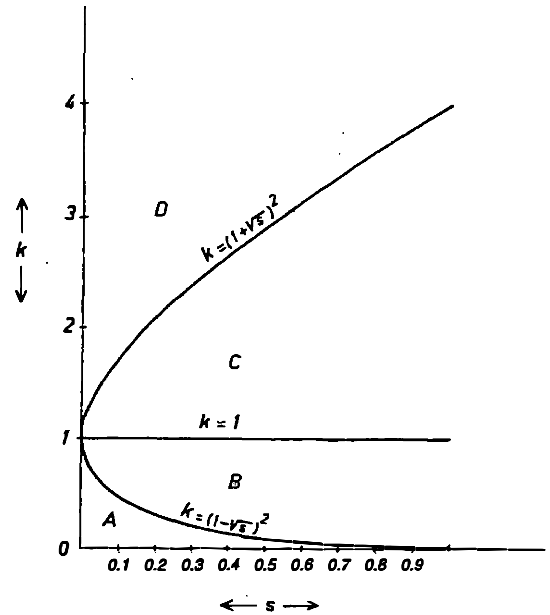


Figure 6.2: Roots in Hicks' second-order model

(4.a) Find the particular solution

(4.b) Discuss the stability conditions. Is it possible to say that they are more stringent?

5. Consider the basic cobweb model (see Sect. 4.1.1) and assume that expectations are formed according to the relation $p_t^e = p_{t-1} + \rho(p_{t-1} - p_{t-2})$, where ρ is a coefficient of expectations (see Goodwin, 1947). When $\rho > 0$ we speak of extrapolative expectations, while expectations are called regressive when $\rho < 0$. The case $\rho = 0$ corresponds to static expectations, as examined in the original cobweb model, hence we shall assume $\rho \neq 0$. After simple substitutions we obtain the second-order difference equation $bp_t - b_1(1+\rho)p_{t-1} + b_1\rho p_{t-2} = a_1 - a$, where we assume that the demand and supply functions are normal, i.e. $b < 0, b_1 > 0$.

(5.a) Show that a particular solution is the equilibrium price, $p_e = (a_1 - a)/(b - b_1)$, and that the characteristic equation of the homogeneous part of the equation is

$$\lambda^2 - \frac{b_1(1+\rho)}{b}\lambda + \frac{b_1\rho}{b} = 0.$$

Table 6.2: Stability condition in Goodwin's model

Range of values of ρ	$\rho \leq -1/3$	$\rho \geq -1/3$
Crucial stability condition	$\frac{b_1}{-b} < \frac{1}{-\rho}$	$\frac{b_1}{-b} < \frac{1}{1+2\rho}$

(5.b) Show that the roots are always real (one positive and one negative) when $\rho > 0$, while when $\rho < 0$ complex roots occur if $b_1/|b| < (-4\rho)/(1+\rho)^2$.

(5.c) Apply the stability conditions to the characteristic equation and show that, when $\rho > 0$, the stability condition turns out to be more stringent than in the basic cobweb model (Hint: the crucial stability condition is the third, whence $b_1/|b| < 1/(1+2\rho)$. Recall that the stability condition in the original cobweb model is $b_1/|b| < 1$). Hence conclude that extrapolative expectations are a destabilizing factor.

(5.d) Examine stability in the case $\rho < 0$, and show that the fact that producers expect price to reverse its movement is an element of stability, provided that the expected inversion is not 'too great' (i.e., ρ must lie in the range $-1 < \rho < 0$). In the contrary case ($\rho < -1$), in fact, such expectations would become a destabilizing factor (Hint: the second and third stability conditions are now both relevant, and which one of them is crucial depends on the absolute value of ρ, as shown in Table 6.2.
It is then easy to check that, for $-1 < \rho < 0$, the crucial stability condition is less restrictive than the stability condition in the original cobweb model; if $\rho = -1$, they are the same; if $\rho < -1$, then the crucial stability condition is more restrictive).

(5.e) For a comparison of the properties of cobweb models according to the various mechanisms of expectation-formation (classical, extrapolative, adaptive, rational, rational with speculation), see Muth (1961, Tables 5.1 and 5.2).

6.4 References

Frisch, R., 1963, Parametric Solution and Programming of the Hicksian Model.

Goodwin, R.M., 1947, Dynamical Coupling with Especial Reference to Markets Having Production Lags, Sect. IV.

Hicks, J.R., 1950, *A Contribution to the Theory of the Trade Cycle,* Chaps. I-VIII and Mathematical Appendices.

Kenkel, J.L., 1974, *Dynamic Linear Economic Models,* Chap. 10, Sects 10.1, 10.3-10.5.

Muth, J.F., 1961, Rational Expectations and the Theory of Price Movements.

Samuelson, P.A., 1939, Interactions between the Multiplier Analysis and the Principle of Acceleration.

Chapter 7

Higher-order Difference Equations

As we know the general form of an n-th order difference equation is

$$c_0 y_t + c_1 y_{t-1} + ... + c_n y_{t-n} = g(t), \qquad c_0 \neq 0, \quad c_n \neq 0. \tag{7.1}$$

Following the scheme of the previous chapters we begin by looking for the solution of the corresponding homogeneous equation.

7.1 Solution of the homogeneous equation

Let us try, as for the second-order equations, the function λ^t, where λ is a constant to be determined. Substituting in the equation we have

$$c_0 \lambda^t + c_1 \lambda^{t-1} + ... + c_n \lambda^{t-n} = 0;$$

therefore

$$\lambda^{t-n}(c_0 \lambda^n + c_1 \lambda^{n-1} + ... + c_n) = 0. \tag{7.2}$$

If λ^t is a solution, Eq.(7.2) must be satisfied for any t, and this—excluding the trivial case $\lambda = 0$—is possible if, and only if,

$$c_0 \lambda^n + c_1 \lambda^{n-1} + ... + c_n = 0, \tag{7.3}$$

which is the characteristic equation of the homogeneous difference equation. Eq.(7.3) may also be written as

$$\lambda^n + a_1 \lambda^{n-1} + ... + a_n = 0, \tag{7.4}$$

where $a_i \equiv c_i / c_0, i = 1, 2, ..., n$.

The solution of Eq.(7.4) yields exactly n roots, which may be real or complex, simple or repeated. In the case of distinct real roots, we have n

functions λ_i^t , each being the solution of the homogeneous equation. According to the general solution, we can combine them with n arbitrary constants, and obtain the general solution

$$y_t = A_1\lambda_1^t + A_2\lambda_2^t + ... + A_n\lambda_n^t. \tag{7.5}$$

If λ^* is a repeated root of multiplicity $m \leq n$, then also $t\lambda^{*t}, t^2\lambda^{*t}, ..., t^{m-1}\lambda^{*t}$ are solutions of the homogeneous equation. In general the solution of the homogeneous equation in the case of repeated real roots is

$$y_t = \sum_{j=1}^{k} P_j(t)\lambda_j^{*t}, \tag{7.6}$$

where λ_j^* are the roots of Eq.(7.4), each with its multiplicity, and $P_j(t)$ are polynomials of the type

$$P_j(t) = A_{1_j} + A_{2_j}t + ... + A_{n_j}t^{m_j-1},$$

where the A's are the arbitrary constants and m_j is the multiplicity of the j-th root.

In the case of complex roots (that always occur in conjugate pairs), each pair give raise to a trigonometric component of the kind

$$r^t(A_1\cos\omega t + A_2\sin\omega t),$$

in exactly the same way as in second-order difference equations. If some pairs of complex roots are repeated, then we shall have terms of the kind

$$\begin{array}{l} r_j^t[A_{11_j} + A_{12_j}t + ... + A_{1m_j}t^{m_j-1})\cos\omega t \\ +(A_{21_j} + A_{22_j}t + ... + A_{2m_j}t^{m_j-1})\sin\omega t], \end{array} \tag{7.7}$$

where m_j is the number of times that the j-th pair of complex roots is repeated and the A's are $2m_j$ arbitrary constants. Of course, in the same equation complex (simple or repeated) roots may occur together with real (simple or repeated) roots, and so a great variety of movements is possible.

7.2 Particular solution of the non–homogeneous equation

A particular solution of the non-homogeneous equation can usually be found by applying the general method of undetermined coefficients (See Part I, Chap. 2, and Chap. 3, Sect. 3.2). We shall exemplify the case in which $g(t) = G$, a constant. As a particular solution try $\bar{y}_t = B$, an undetermined constant. Substitution in Eq. (7.1) yields $B = G/(c_0 + c_1 + ... + c_n)$. If $c_0 + c_1 + ... + c_n = 0$, try $\bar{y}_t = Bt$. Substituting in Eq. (7.1) and collecting terms we have $(c_0 + c_1 + ... + c_n)Bt - B(c_1 + 2c_2 + ... + nc_n) = G$, and so $B = -G/(c_1 + 2c_2 + ... + nc_n)$. If also $c_1 + 2c_2 + ... + nc_n = 0$, try $\bar{y}_t = Bt^2$, and so on.

The particular solution of the non-homogeneous equation may usually be interpreted, as we shall see in the economic applications, as the (stationary or moving) equilibrium of the variable y_t.

7.2.1 The operational method

Let us now consider the case in which the functional form of $g(t)$ is not known, and apply the operational calculus (already introduced in Sects. 3.2.6 and 5.2); let us note, incidentally, that the procedure that we are going to explain can also be seen as an application of the z-transform method (see, for example, Luenberger (1979, Chap. 8, Sects. 8.1-8.3). We now have

$$c_0y_t + c_1y_{t-1} + ... + c_ny_{t-n} = x_t, \tag{7.8}$$

from which, dividing through by c_0 and letting $X_t \equiv (1/c_0)x_t$,

$$y_t + a_1y_{t-1} + ... + a_ny_{t-n} = X_t,$$

namely, in operational notation,

$$(1 + a_1L + ... + a_nL^n)y_t = X_t. \tag{7.9}$$

Thus a particular solution to Eq. (7.9) is given by

$$\bar{y}_t = \frac{1}{1 + a_1L + ... + a_nL^n}X_t. \tag{7.10}$$

As we have already recalled in Sect. 3.2.6, operators can be treated for our purposes as ordinary algebraic quantities. Hence we can write the polynomial $1 + a_1L + ... + a_nL^n$ as $L^n(F^n + a_1F^{n-1} + ... + a_n)$ where $F \equiv L^{-1}$ is the forward operator (inverse lag operator).

We then observe that $(F^n + a_1F^{n-1} + ... + a_n)$ can be factorised into $(F-F_1)(F-F_2)...(F-F_n)$, where $F_1, F_2, ..., F_n$ are the roots of $F^n + a_1F^{n-1} + ... + a_n = 0$. Note that this equation coincides with the characteristic equation (7.4), hence $F_1, F_2, ..., F_n$ coincide with $\lambda_1, \lambda_2, ..., \lambda_n$, the roots of the characteristic equation.

Thus we have

$$\begin{aligned} 1 + a_1L + ... + a_nL^n &= L^n(F - \lambda_1)(F - \lambda_2)...(F - \lambda_n) \\ &= (LF - \lambda_1L)(LF - \lambda_2L)...(LF - \lambda_nL) \\ &= (1 - \lambda_1L)(1 - \lambda_2L)...(1 - \lambda_nL). \end{aligned} \tag{7.11}$$

We can now expand the rational polynomial in L on the r.h.s. of Eq. (7.10) in partial fractions (for the partial fraction expansion see any calculus textbook or Turnbull, 1957, Sect. 19), that is, assuming distinct roots,

$$\frac{1}{1+a_1L+...+a_nL^n}=\frac{\theta_1}{1-\lambda_1L}+\frac{\theta_2}{1-\lambda_2L}+...+\frac{\theta_n}{1-\lambda_nL}, \tag{7.12}$$

where it turns out that

$$\theta_i=\frac{\lambda_i}{(\lambda_i-\lambda_1)...(\lambda_i-\lambda_{i-1})(\lambda_i-\lambda_{i+1})...(\lambda_i-\lambda_n)}=\frac{\lambda_i}{\prod\limits_{\substack{j=1\\j\neq i}}^{n}(\lambda_i-\lambda_j)}, \tag{7.13}$$

as the reader may wish to check by direct substitution. Thus the particular solution (7.10) becomes

$$\overline{y}_t=\sum_{r=1}^{n}\frac{\theta_r}{1-\lambda_rL}X_t, \tag{7.14}$$

namely, by using the expansion of $(1-\lambda_rL)^{-1}$ derived in Sect. 3.2.6, Eq. (3.33),

$$\overline{y}_t=\sum_{r=1}^{n}\theta_r\sum_{i=0}^{\infty}(\lambda_r)^iX_{t-i}. \tag{7.15}$$

Equation (7.15) is the backward solution, to be used when $|\lambda_r|<1$. In the contrary case we shall use the alternative forward expansion explained in Chap. 3, Eq. (3.38). If some roots are stable and others unstable, we shall of course use the respective appropriate expansion.

Equation (7.12) holds independently of the nature of the roots (real or complex), but cannot be applied in the case of equal roots. Although in general there can be several multiple roots (each with its own multiplicity), for the sake of simplicity we shall only examine the case of n equal roots, $\lambda_1=\lambda_2=...=\lambda_n=\lambda^*$. In this case we have $1/(1+a_1L+...+a_nL^n)=(1-\lambda^*L)^{-n}$. We can then use the binomial expansion $(1+x)^\alpha$ (see, for example, Hyslop, 1959, Sect. 19) for $\alpha=-n$ and $x=-\lambda^*L$, which gives

$$\frac{1}{(1-\lambda^*L)^n}=\sum_{i=0}^{\infty}\binom{n+i-1}{i}\lambda^{*i}L^i. \tag{7.16}$$

This can be used to obtain the particular solution

$$\overline{y}_t=\sum_{i=0}^{\infty}\binom{n+i-1}{i}\lambda^{*i}X_{t-i}. \tag{7.17}$$

Let us note, finally, that the method under consideration can also be applied to the standard cases in which we are given a specific functional form for $g(t)$, as shown in Sect. 5.2.

7.3 Determination of the arbitrary constants

In order to determine the n arbitrary constants, n additional condition are needed, which usually take the form of $y_0, y_1, ...y_{n-1}$ being known values (whence the name of *initial conditions*). Substituting these values in the general solution, we obtain a system of n linear equation in the n unknowns $A_1, A_2, ..., A_n$. Consider the general solution of Eq. (7.1)

$$y(t)=A_1y_1(t)+A_2y_2(t)+.......+A_ny_n(t)+\overline{y}(t),$$

where the $y_j(t)$ are n distinct solutions (λ^t in the case of distinct roots, and/or functions of the type $t\lambda^{*t}, t^2\lambda^{*t}$ etc. in the case of multiple roots; if there are complex roots the known transformations will be used). Substituting the given values y_0 etc. in the place of $y(0)$ etc. we obtain, rearranging terms,

$$\begin{aligned}
A_1y_1(0)+A_2y_2(0)+...+A_ny_n(0) &= y_0-\overline{y}(0),\\
A_1y_1(1)+A_2y_2(1)+...+A_ny_n(1) &= y_1-\overline{y}(1),\\
\cdots\cdots\cdots\cdots\cdots\cdots &\cdots\cdots\cdots\\
A_1y_1(n-1)+A_2y_2(n-1)+...+A_ny_n(n-1) &= y_{n-1}-\overline{y}(n-1),
\end{aligned} \tag{7.18}$$

where of course $\overline{y}(0), \overline{y}(1), ...\ \overline{y}(n-1)$ are absent if the equation is homogeneous. System (7.18) is a linear system whose determinant is

$$\begin{vmatrix} y_1(0) & y_2(0) & \cdots & y_n(0)\\ y_1(1) & y_2(1) & \cdots & y_n(1)\\ \cdots & \cdots & \cdots & \cdots\\ y_1(n-1) & y_2(n-1) & \cdots & y_n(n-1)\end{vmatrix}.$$

Since the functions $y_j(t)$ are a fundamental set (that is, they are linearly independent), this determinant is different from zero and so the system (7.18) can always be solved.

7.4 Stability conditions

As we have seen, in the solution of higher-order difference equations there are no *conceptual* difficulties greater than those met in relation to second-order equations. The 'jump' in conceptual difficulty occurs when we pass from first- to second-order equations (complex roots, etc.). From the second-order on we do not think that the conceptual difficulties are greater. The greater difficulty of higher-order equations lies in the practical problem of how to find the roots of the characteristic equation. This is a problem in numerical analysis (nowadays easily solvable with the current computing equipment) which is outside the scope of this book.

This problem, in any case, is *not* of great importance for the economic theorist, who generally works with *qualitative* information only. In this connection it would be highly desirable to have conditions –of the kind of Eqs. (5.19)– to check the stability of the movement by means of inequalities involving the coefficients of the characteristic equation, i.e. to check whether the roots of the characteristic equation are all in modulus less than unity without solving the characteristic equation.

Another way of stating stability is that the roots all lie within the unit circle in the complex plane, that is, they are 'stable' roots. The reason why also the roots with unit modulus are to be excluded is that we want *asymptotic* stability, that is $\lim_{t\to+\infty} y(t) = 0$, where $y(t)$ is the general solution of the homogeneous equation (for a more detailed and rigorous treatment of the notion of stability see Part III, Chap. 21). Now, a root with unit modulus gives rise, in the solution, either to a constant term (root $= +1$) or to a constant-amplitude alternation (root $= -1$) or to a constant-amplitude oscillation (pair of complex conjugate roots with unit modulus). In each of these cases the time path, while not being divergent, is not stable in the sense defined above.

Such conditions exist, and can be given in two forms.

7.4.1 Necessary and sufficient stability conditions (Samuelson's form)

Given equation (7.4), form the following sums ($a_0 = 1$):

$$
\begin{aligned}
\bar{a}_0 &= \sum_{i=0}^{n} a_i,\\
\bar{a}_1 &= \sum_{i=0}^{n} a_i(n-2i),\\
&\dots\dots\dots\dots\dots\dots\dots\dots,\\
\bar{a}_r &= \sum_{i=0}^{n} a_i \sum_{k=0}^{n} \binom{n-i}{r-k}(-1)^k \binom{i}{k},\\
&\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots,\\
\bar{a}_n &= 1 - a_1 + a_2 - \dots\dots + (-1)^{n-1}a_{n-1} + (-1)^n a_n.
\end{aligned}
$$

Then the stability condition are

$$
\begin{aligned}
\bar{a}_0 &> 0,\\
\Delta_1 &> 0,\\
\Delta_2 &> 0,\\
&\dots\dots\dots,\\
\Delta_n &> 0,
\end{aligned}
$$

where the Δ's are the leading principal minors[1] of the matrix (of which only the first n rows and columns must be considered)

[1] Let us recall that, among all the $\binom{n}{r}$ principal minors of order r of a $n \times n$ matrix, the *leading* principal minor (also called the upper left-hand principal minor) is that formed by deleting the last $n-r$ columns and the last $n-r$ rows of the determinant of the matrix.

$$
\begin{bmatrix}
\bar{a}_1 & \bar{a}_3 & \bar{a}_5 & \dots\\
\bar{a}_0 & \bar{a}_2 & \bar{a}_4 & \dots\\
0 & \bar{a}_1 & \bar{a}_3 & \dots\\
0 & \bar{a}_0 & \bar{a}_2 & \dots\\
\dots & \dots & \dots & \dots\\
\dots & \dots & \dots & \dots
\end{bmatrix}.
$$

7.4.2 Necessary and sufficient stability conditions (Cohn-Schur form)

The following determinants (the broken lines are inserted only to bring out their symmetry):

$$
\left|\begin{array}{c|c}
a_0 & a_n\\
- & -\\
a_n & a_0
\end{array}\right|,
\quad
\left|\begin{array}{cc|cc}
a_0 & 0 & a_n & a_{n-1}\\
a_1 & a_0 & 0 & a_n\\
\hline
a_n & 0 & a_0 & a_1\\
a_{n-1} & a_n & 0 & a_0
\end{array}\right|, \dots,
$$

$$
\left|\begin{array}{cccc|cccc}
a_0 & 0 & \dots & 0 & a_n & a_{n-1} & \dots & a_{n-r+1}\\
a_1 & a_0 & \dots & 0 & 0 & a_n & \dots & a_{n-r+2}\\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots\\
a_{r-1} & a_{r-2} & \dots & a_0 & 0 & 0 & \dots & a_n\\
- & - & - & - & - & - & - & -\\
a_n & 0 & \dots & 0 & a_0 & a_1 & \dots & a_{r-1}\\
a_{n-1} & a_n & \dots & 0 & 0 & a_0 & \dots & a_{r-2}\\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots\\
a_{n-r+1} & a_{n-r+2} & \dots & a_n & 0 & 0 & \dots & a_0
\end{array}\right| \dots
$$

$$
\dots,
\left|\begin{array}{cccc|cccc}
a_0 & 0 & \dots & 0 & a_n & a_{n-1} & \dots & a_1\\
a_1 & a_0 & \dots & 0 & 0 & a_n & \dots & a_2\\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots\\
a_{n-1} & \underset{n-2}{a_n} & \dots & a_0 & 0 & 0 & \dots & a_n\\
- & - & - & - & - & - & - & -\\
a_n & 0 & \dots & 0 & a_0 & a_1 & \dots & a_{n-1}\\
a_{n-1} & a_n & \dots & 0 & 0 & a_0 & \dots & a_{n-2}\\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots\\
a_1 & a_2 & \dots & a_n & 0 & 0 & \dots & a_0
\end{array}\right|.
$$

must be all positive (in Eq. (7.4), $a_0 = 1$). For this formulation, see Chipman (1950). Further results are given in Barnett (1974, 1983).

The Schur conditions may seem easier to apply, since no previous transformations on the coefficients of the equation are required, and in fact some authors prefer them to the Samuelson conditions. However, the latter conditions, once the transformations are made, imply the expansion of smaller-order determinants (the maximum order of the determinants to expand is n in Samuelson form and $2n$ with the Schur's form, and this is a rather important thing in the economy of the computations). So Schur's form may not be more convenient than Samuelson's form, and this is why we have given both. In the following chapters we shall use the Samuelson conditions. But this is a choice that perhaps reflect personal tastes.

In either form the stability conditions become increasingly complicated as the order of the equation increases and correspondingly, their economic interpretation becomes more and more intricate. Indeed, there is little hope of extricating a clear economic meaning from the stability conditions when the equation is of order higher than the third.

For the third-order equation, the explicit form of the stability conditions is

$$\begin{aligned} 1+a_1+a_2+a_3 &> 0,\\ 3-a_1-a_2+3a_3 &> 0,\\ 1-a_1+a_2-a_3 &> 0,\\ 3+a_1-a_2-3a_3 &> 0,\\ (3+a_1-a_2-3a_3)(3-a_1-a_2+3a_3)&\\ -(1+a_1+a_2+a_3)(1-a_1+a_2-a_3)&\\ \equiv 8(-a_3^2+a_1a_3-a_2+1) &> 0. \end{aligned} \tag{7.19}$$

It can easily be seen that either the second or the fourth inequality can be eliminated, since one of them is clearly redundant[2]. However, we have given them both, since in economic applications it may be convenient sometimes to drop the second and sometimes to drop the fourth. The fifth condition can be further simplified to $-a_3^2+a_1a_3-a_2+1>0$.

It can be shown (Farebrother, 1973) that a simplified form of the necessary and sufficient stability conditions for a third-order equation is

$$\begin{aligned} 1+a_1+a_2+a_3 &> 0,\\ 1-a_1+a_2-a_3 &> 0,\\ 1-a_2+a_1a_3-a_3^2 &> 0,\\ a_2 &< 3. \end{aligned} \tag{7.20}$$

The first two inequalities in(7.20) may be replaced by the single inequality $1+a_2>|a_1+a_3|$. Simplified condition for a fourth-order equation are also given in Farebrother's paper.

Okuguchi and Irie (1990) have subsequently shown that the fourth inequality in (7.20) is redundant, hence the most simplified set of necessary and sufficient conditions for a third order equation turns out to be

[2]Call d_1, d_2, d_3, d_4, d_5 the five inequalities in (7.19). Inequality d_5 is not implied by the others and so is independent. Now, d_5, d_1, d_2, d_3 together imply d_4, and d_5, d_1, d_3, d_4 together imply d_2. On the contrary d_5, d_1, d_2, d_4 together do not imply d_3 and d_5, d_2, d_3, d_4 together do not imply d_1. This proves that the redundant inequality is either d_4 or d_2.

$$\begin{aligned} 1+a_1+a_2+a_3 &> 0,\\ 1-a_1+a_2-a_3 &> 0,\\ 1-a_2+a_1a_3-a_3^2 &> 0. \end{aligned} \tag{7.21}$$

In the particular case in which all the coefficients of Eq. (7.4) are non positive (i.e. $a_i \le 0$ for $i=1,2,.....n$), the equation has a simple positive root with the largest modulus which lies between 1 and $\sum_{i=1}^n |a_i|$ (Sato, 1970). In this case, therefore, a particularly simple necessary and sufficient condition is

$$\sum_{i=1}^n |a_i| < 1. \tag{7.22}$$

When the coefficients are arbitrary, condition (7.22) is no longer necessary, but remains a sufficient stability condition, whereas a *necessary* stability condition (Smithies, 1942) is

$$-\sum_{i=1}^n a_i < 1. \tag{7.23}$$

Another *necessary* stability condition (Farebrother, 1973) is

$$|a_j| < {}^nC_j, \qquad j=1,2,......,n, \tag{7.24}$$

where ${}^nC_j = n!/[j!(n-j)!]$.

Finally, consider the case in which all the coefficients of Eq. (7.4) are positive; then, a *sufficient* stability condition (generalized Kakeya theorem: Murata, 1977; see also Sato, 1970) is that

$$1 > a_1 \ge a_2 \ge ..... \ge a_{n-1} > a_n. \tag{7.25}$$

7.5 Exercises

7.5.1 Example

Let us solve the following difference equation

$$y_t - 1.1y_{t-1} - 3.8y_{t-2} + 4y_{t-3} = 0,$$

whose characteristic equation is

$$\lambda^3 - 1.1\lambda^2 - 3.8\lambda + 4. \tag{7.26}$$

A root is 2, as the student may check by straightforward substitution in (7.26). As we know from elementary algebra, Eq.(7.26) may be divided by $\lambda-2$, with remainder zero. We obtain

$$\lambda^3 - 1.1\lambda^2 - 3.8\lambda + 4 = (\lambda-2)(\lambda^2+0.9\lambda-2) = 0.$$

and so the remaining two roots of (7.26) are given by the solution of

$$\lambda^2 - 0.9\lambda - 2 = 0,$$

whose roots are

$$\frac{-0.9 \pm \sqrt{8.81}}{2} \simeq \begin{Bmatrix} 1.034 \\ -1.935 \end{Bmatrix}.$$

The general solution is then

$$y_t = A_1 2^t + A_2(1.034)^t + A_3(-1.935)^t, \tag{7.27}$$

where A_1, A_2, A_3 are arbitrary constants. The movement is divergent, with an explosive monotonic component given by the first two terms on the right-hand side of Eq. (7.27), on which an improper oscillation of increasing amplitude is superimposed.

7.5.2 Other exercises

(a) Find the general solution of the following difference equations:

(*i*) $y_t - 4y_{t-1} + 4.8y_{t-2} - 1.6y_{t-3} = 100,$
$y(0) = 400, \quad y(1) = 420, \quad y(2) = 450.$

(*ii*) $y_t - 3.6y_{t-1} + 4.05y_{t-2} - 1.35y_{t-3} = 100,$
$y(0) = 900, \quad y(1) = 910, \quad y(2) = 920.$

(b) Solve the difference equations and evaluate the arbitrary constants given the initial conditions:

(*i*) $y_{t+3} - y_{t+2} - 2y_{t+1} + 2y_t = 0 \quad y(0) = 0; y(1) = 2; y(2) = 4.$

(*ii*) $y_{t+4} + 5y_{t+2} + 4y_t = 0 \quad y(0) = 0; y(1) = 2; y(2) = -2; y(3) = 4.$

(*iii*) $y_t + y_{t-2} + 1/4y_{t-4} = 0 \quad y(0) = 0; y(1) = 5; y(2) = -3; y(3) = 1.$

7.6 References

Barnett, S., 1974, Application of the Routh Array to Stability of Discrete-Time Linear Systems.
Barnett, S., 1983, *Polynomials and Linear Control Systems,* Chap. 3, Sect. 3.2.2.
Baumol, W.J., 1970, *Economic Dynamics,* Chap. 10, Sect. 5; Chap. 12.
Chipman, J.S., 1950, The Multi-Sector Multiplier, Sect. VI.
Farebrother, R.W., 1973, Simplified Samuelson Conditions for Cubic and Quartic Equations.
Folsom, R.N., D.C. Boger and H.C. Mullikin, 1976, Stability Conditions for Linear Constant Coefficient Difference Equations in Generalized Differenced Form.
Goldberg, S., 1958, *Introduction to Difference Equations,* Chap. 3, Sect. 3.7.

Hall, S.G., 1984, On the Solution of High Order, Symmetric, Difference Equations.
Hyslop, J.M., 1959, *Infinite Series,* Chap. III, Sect. 18.
Murata, Y., 1977, *Mathematics for Stability and Optimization of Economic Systems,* Chap. 4, pp. 136-7.
Kenkel, J.L., 1974, *Dynamic Linear Economic Models,* Chap. 7, Sect. 7.2.
Luenberger, D.G., 1979, *Introduction to Dynamic Systems,* Chap. 8, Sects. 8.1-8.3.
Okuguchi, K. and K. Irie, 1990, The Schur and Samuelson Conditions for a Cubic Equation.
Samuelson, P.A., 1941a, Conditions that the Roots of a Polynomial be less than Unity in Absolute Value.
Samuelson, P.A., 1947, *Foundations of Economic Analysis,* Appendix B, Sect. 37.
Sato, R., 1970, A Further Note on a Difference Equation Recurring in Growth Theory.
Smithies, A., 1942, The Stability of Competitive Equilibrium, p. 269.
Turnbull, H.W., 1957, *Theory of Equations,* Chap. III.

Chapter 8

Higher-order Difference Equations in Economic Models

8.1 Inventory cycles

The importance of inventories has long been recognised, and—although inventory investment is a relatively small component of GDP—the shift from inventory accumulation to liquidation is a critical factor in the propagation of cyclical reversals in economic activity. Theoretical and empirical research on inventory cycles is actively going on (see, for example, Blinder and Maccini, 1991; Fiorito ed., 1994).

One of the seminal models for explaining inventory cycles is that of Metzler (1941), that played the same role as Samuelson's multiplier-accelerator model for fixed investment. For a continuous time version of this model see below, Chap. 17, Sect. 17.2, problem 5.

Let us observe that total current output is the sum of the output of consumption goods and the output of investment goods, the latter being assumed exogenous. The output of consumption goods is made up of two components:

(1) output to be currently sold, according to producers' *expectations* on sales, U_t;

(2) output to bring inventories to their *desired* level $\hat{Q}_t$.

Component (2) may, of course, be negative, which simply means that firms will produce *less* than expected sales, the difference being provided for, in their plans, by the desired decrease in inventories.

Naturally, expectations may be wrong, i.e. actual sales may be different from expected sales, the difference implying an *unintended* variation in inventories. Note that actual sales coincide with current *consumption demand* C_t, which is not the same as the output of consumption goods, since in Metzler's model the output of (as specified above) and the demand for consumption goods are allowed to be different.

Putting all this together, we have the basic equation of Metzler's model

$$Y_t = U_t + (\hat{Q}_t - Q_{t-1}) + I_0. \tag{8.1}$$

We must now specify the desired stock of inventories, which is determined by assuming that producers wish to maintain a constant ratio k between inventories and sales (the inventory accelerator). Since actual sales will be known only *ex post,* producers apply this ratio to *expected* sales to compute the desired level of inventories. Hence

$$\hat{Q}_t = kU_t. \tag{8.2}$$

To complete the model we have to specify expectations. Metzler considers various formulations, amongst which

$$U_t = C_{t-1} + \rho(C_{t-1} - C_{t-2}), \quad \rho > 0, \tag{8.3}$$

that relates expected sales to realized sales according to an extrapolative mechanism. Consumption demand depends on current income, i.e. no lag is assumed to exist in the consumption function:

$$C_t = bY_t. \tag{8.4}$$

In order to reduce the model to a single difference equation in one unknown function of time (Y_t), we observe that in any period t the actual inventory level Q_t is equal to the level that producers had planned for that period, i.e. $\hat{Q}_t$, *minus* the unintended variation in inventories (if any) occurring because of the difference between realized and expected sales, $C_t - U_t$. Thus

$$Q_{t-1} = \hat{Q}_{t-1} - (C_{t-1} - U_{t-1}) = (1+k)U_{t-1} - C_{t-1}. \tag{8.5}$$

After simple substitutions we obtain

$$Y_t - b[(1+k)(1+\rho)+1]Y_{t-1} + b(1+k)(1+2\rho)Y_{t-2} - (1+k)b\rho Y_{t-3} = I_0. \tag{8.6}$$

A particular solution is obtained by trying $\overline{Y}_t = \overline{Y}$, a constant, which yields

$$\overline{Y} = \frac{I_0}{1-b}, \tag{8.7}$$

which is the stationary equilibrium, given by the multiplier applied to the constant exogenous expenditure.

The homogeneous equation corresponding to Eq. (8.6) is

$$\lambda^3 - b[(1+k)(1+\rho)+1]\lambda^2 + b(1+k)(1+2\rho)\lambda - (1+k)b\rho = 0. \tag{8.8}$$

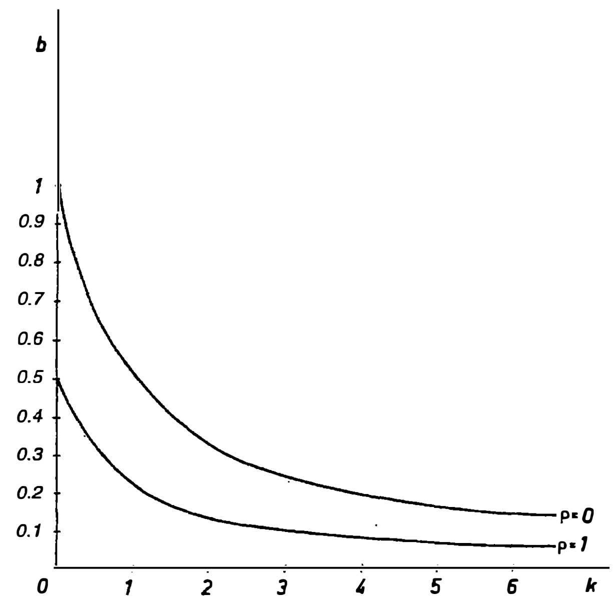


Figure 8.1: Metzler's inventory cycle: stability regions

Let us now apply the stability conditions (7.19) of Chap. 7, dropping the second inequality. We have

$$\begin{aligned}
&1 - b[(1+k)(1+\rho)+1] + b(1+k)(1+2\rho) - (1+k)b\rho &&> 0,\\
&1 + b[(1+k)(1+\rho)+1] + b(1+k)(1+2\rho) + (1+k)b\rho &&> 0,\\
&3 - b[(1+k)(1+\rho)+1] - b(1+k)(1+2\rho) + 3(1+k)b\rho &&> 0,\\
&-(1+k)^2b^2\rho^2 + b^2(1+k)\rho[(1+k)(1+\rho)+1] - b(1+k)(1+2\rho) + 1 &&> 0.
\end{aligned}$$

The first and second inequalities are certainly satisfied, since the marginal propensity to consume is smaller than unity and the expectation coefficient is non negative. The relevant inequalities are then the third and fourth, which, after simple manipulation, can be written as

$$\begin{aligned}
3 - b(2k+3) &> 0,\\
(1+k)(2+k)\rho b^2 - (1+k)(1+2\rho)b + 1 &> 0.
\end{aligned} \tag{8.9}$$

These inequalities involve three parameters: b, k, ρ. By letting ρ vary parametrically, a simple diagram can be drawn (see Fig. 8.1).

For $\rho = 0$, for example, the crucial stability condition becomes $b < 1/(1+k)$, i.e. all the combinations of b, k lying below the curve $b = 1/(1+k)$. Note that $\rho = 0$ means that producers hold static expectations, i.e. expected sales are equal to realized sales. For $\rho = 1$ the stability region becomes much smaller (see the diagram, and exercise 6), and such that the economic system is not likely to be stable. The introduction of a partial adjustment equation, however, stabilizes the model (see exercise 7).

8.2 Distributed lags and interaction between the multiplier and the accelerator

The basic idea of this model, due to Hicks (1950), is that, in any period, investment and consumption depend on the values of national income in the n preceding periods. Thus we are in the presence of distributed lag equations (see Chap. 3, Sect. 3.4).

As regards the investment function, the assumption is that the investment induced by a variation in income is not entirely carried out in a single period but it is spread over n successive periods. Thus, if ΔY is the variation in income and $k\,\Delta Y$ is the induced investment, given the acceleration coefficient k, a fraction $e_1(k\Delta Y)$ will be invested in the next period, another fraction $e_2(k\Delta Y)$ two periods after, and so on up to $e_n(k\Delta Y)$. Of course $e_1 + e_2 + \ldots\ldots + e_n = 1$. Call $e_i k = k_i$; thus $\sum_{i=1}^{n} k_i = k$. From all this it follows that total induced investment actually carried out in any period t consists of a part depending on ΔY_{t-1}, of another part depending on ΔY_{t-2}, and so on up the part depending on ΔY_{t-n}. Calling I_t' total induced investment carried out in period t, we have

$$I_t' = k_1(Y_{t-1} - Y_{t-2}) + k_2(Y_{t-2} - Y_{t-3}) + \ldots + k_n(Y_{t-n} - Y_{t-n-1}). \qquad (8.10)$$

Consumption is assumed to depend on the values of national income in the last n periods, i.e.

$$C_t = b_1 Y_{t-1} + b_2 Y_{t-2} + \ldots + b_n Y_{t-n}. \qquad (8.11)$$

where

$$b_1 + b_2 + \ldots + b_n = b.$$

Eq. (8.11) is based on the assumption that the variation in consumption depending on a variation in income is spread over n successive periods, as for induced investment.

With equations (8.10) and (8.11), the model gives rise to a difference equation of order $n+1$. Here we shall analyse the case in which investment is distributed over two successive periods and so is consumption, so that a third-order equation results. The equations of the model are then

$$\begin{aligned} Y_t &= C_t + I_t, && (8.12)\\ C_t &= b_1 Y_{t-1} + b_2 Y_{t-2}, && (8.13)\\ I_t &= I_t' + I_t'', && (8.14)\\ I_t' &= k_1(Y_{t-1} - Y_{t-2}) + k_2(Y_{t-2} - Y_{t-3}), && (8.15)\\ I_t'' &= A_0(1+g)^t, && (8.16)\end{aligned}$$

where $b_1 + b_2 = b, k_1 + k_2 = k$, and autonomous investment I_t'' is assumed to increase over time at the constant rate of growth g. After the usual substitutions we obtain the third-order equation

$$Y_t - (b_1 + k_1)Y_{t-1} - (k_2 + b_2 - k_1)Y_{t-2} + k_2 Y_{t-3} = A_0(1+g)^t. \qquad (8.17)$$

As a particular solution, let us try $\bar{Y}_t = Y_0(1+g)^t$, where Y_0 is an undetermined constant. Substituting in (8.17) and performing the usual manipulations—the details are left as an exercise—we obtain

$$\bar{Y}_t = \frac{A_0(1+g)^3}{(1+g)^3 - (b_1 + k_1)(1+g)^2 - (k_2 + b_2 - k_1)(1+g) + k_2}(1+g)^t. \qquad (8.18)$$

We assume that the denominator in (8.18) is positive, so that the particular solution is economically meaningful and can be interpreted as the equilibrium growth path of national income. The characteristic equation of the homogeneous part of (8.17) is

$$\lambda^3 - (b_1 + k_1)\lambda^2 - (k_2 + b_2 - k_1)\lambda + k_2 = 0. \qquad (8.19)$$

Before examining this equation, let us show that k_2 is presumably greater than 1. First of all, as Hicks points out, it is very likely than the greatest part of induced investment is concentrated not in the period immediately following the variation in income, but in the farthest periods. In our case the implication is that k_2 is greater than k_1. Now, empirical data allow us to assume that the value of the overall acceleration coefficient k is not smaller than 2. From $k_1 + k_2 = k \geq 2$ and from $k_2 > k_1$ it follows that $k_2 > 1$.

Now, if $k_2 > 1$, the deviations of national income from the trend, whatever their nature, are certainly unstable. To show this there is no need to apply the stability conditions to Eq. (8.19), since we can use the well known fact that in any n-th degree algebraic equation (written in such a way that the coefficient of the term of highest power is unity) *the product of the roots equals $(-1)^n$ times the constant term.* In our case we have

$$\lambda_1\lambda_2\lambda_3 = -k_2,$$

in the case of three real roots, and

$$\lambda_1 r^2 = -k_2,$$

in the case of one real root and a pair of complex conjugate roots with modulus r[1].

If we consider only absolute values, it is obvious that at least one root must be in absolute value greater than 1 since $k_2 > 1$. In other words $k_2 > 1$ is a sufficient (although *not* a necessary) *instability* condition. Thus two cases are possible:

(1) three real roots. Since the succession of the coefficient in Eq. (8.19) is $+--+$, two roots will be positive and one negative (monotonic movement plus an improper oscillation).

(2) one real root and two complex conjugate roots. Since the product of the roots is $-k_2 < 0$, and since $r^2 > 0$, it follows that the real root must be negative (an improper oscillation plus a proper oscillation).

Once we have found that the movement is unstable, the nature of the movement (monotonic, oscillatory, etc.) is of secondary importance, since the presence of an upper and lower limit (the *ceiling* and the *floor* in Hicks' terminology) will give rise in any case to a constant amplitude oscillation in relative terms. It is interesting to note that the presence of these limits transforms the model into a non-linear growth-cyclical model of the hard-bouncing oscillator type, which however can be treated by linear methods since it can be divided into several linear pieces connected by a switching function (see below, Exercises 1, 2 and 3).

8.3 Exercises

1. In Hicks' model, the upper limit B_t is given by full employment output, which is assumed to grow at the same rate as exogenous investment, i.e. $B_t = B_0(1+g)^t$, where $B_0 > Y_0$, since the ceiling is certainly higher than the trend. This is enough to show that output cannot 'crawl' along the ceiling when it hits it. First of all, note that when income 'hits' the ceiling there is a switch, namely

$$Y_t \begin{cases} = (b_1+k_1)Y_{t-1} + (k_2+b_2-k_1)Y_{t-2} - k_2Y_{t-3} + A_0(1+g)^t \\ \text{if } (b_1+k_1)Y_{t-1} + (k_2+b_2-k_1)Y_{t-2} - k_2Y_{t-3} + A_0(1+g)^t \leq B_0(1+g)^t; \\ = B_0(1+g)^t \\ \text{if } (b_1+k_1)Y_{t-1} + (k_2+b_2-k_1)Y_{t-2} - k_2Y_{t-3} + A_0(1+g)^t > B_0(1+g)^t. \end{cases}$$

Now, income can grow at the constant rate g only on the trend, since this latter has been found as the only exponential function of the type $Y_0(1+g)^t$ compatible with the non-homogeneous difference equation defining the movement of income over time. And since we have assumed that the denominator

[1] If $\alpha \pm i\theta$ is a pair of complex conjugate numbers then $(\alpha+i\theta)(\alpha-i\theta) = \alpha^2+\theta^2$ since $i^2=-1$. The modulus or absolute value of a complex number $\alpha+i\theta$ or $\alpha-i\theta$ is defined as $r = +(\alpha^2+\theta^2)^{1/2}$, hence the statement in the text follows.

in Eq. (8.18) is positive, we see that $(1+g)$ cannot be a root of the characteristic equation (8.19). Hence income must turn down toward the trend, since the constant rate of growth g is sustainable only there.

Prove formally that income cannot stay along the ceiling for more than three periods (Hint: assume that $Y_{t-1}, Y_{t-2}, Y_{t-3}$ have been equal to the corresponding values of the ceiling, i.e. $Y_{t-1} = B_0(1+g)^{t-1}$ etc. Use these values to determine Y_t through Eq. (8.17) and show that $Y_t < B_0(1+g)^t$.)

2. Once the descending phase derived in exercise 1 has begun, the explosiveness of the movement causes income to overshoot the trend and go further down. We must now turn to the explanation of the floor. This comes from an intrinsic non-linearity in the accelerator. When income decreases, net investment must be negative, according to the acceleration principle. But the absolute value of negative net investment (i.e., disinvestment in fixed capital) cannot exceed the absolute value of the physical depreciation of the capital stock. The maximum rate at which the capital stock can be reduced is obtained when capital goods are not replaced. Thus in the descending phase Eq. (8.16) holds only as long as induced investment is not greater than depreciation in absolute value. In the opposite case, Eq. (8.16) must be replaced by the equation $I'_t = -a_t$, where $a_t > 0$ is the absolute value of depreciation. This is another non-linear feature of the model of the switching function type, namely

$$I'_t \begin{cases} = k_1(Y_{t-1}-Y_{t-2}) + k_2(Y_{t-2}-Y_{t-3}) \\ \text{if } k_1(Y_{t-1}-Y_{t-2}) + k_2(Y_{t-2}-Y_{t-3}) \geq -a_t; \\ = -a_t \\ \text{if } k_1(Y_{t-1}-Y_{t-2}) + k_2(Y_{t-2}-Y_{t-3}) < -a_t. \end{cases}$$

Assume, following Hicks, that $a_t = a$, where a is a constant (for example, the depreciation corresponding to the capital stock in existence when the descending phase begins), and determine the floor (Hint: the floor is a particular solution of the difference equation resulting from the model (8.12)-(8.16) when $I'_t = -a$.).

3. With the data of exercise 2, show that in the descending phase the movement converges to the floor. Also show that—since the floor is a growing magnitude—the overall value of Y_t will sooner or later begin to increase, hence we are back into the initial model and a new ascending phase begins.

4. Consider the general form (i.e., $n > 3$) of Hicks' model and suppose that induced investment is evenly spread over n successive periods, i.e. $k_i = (1/n)k$. Show that the model is stable for n sufficiently great.

5. In Metzler's model, assume that producers want to maintain inventories at a fixed level, $\hat{Q}_t = Q_0$. Derive the equation describing the behaviour of income and examine the stability of the model.

6. Show that, when $\rho = 1$, the crucial stability condition becomes $b < b_1$, where b_1 is the smaller root of the equation $(1+k)(2+k)b^2-(1+k)(1+2)b+1 = 0$ (Hint: the first inequality in (8.9) is satisfied for $b < 3/(2k+3)$. The second can be treated as a second-degree inequality in b, which turns out to be satisfied for the values of b outside the interval of the roots b_1, b_2. Then show, by considering the parabola $f(b) = (1+k)(2+k)b^2-(1+k)(1+2)b+1$, that $f(3/(2k+3)) < 0$, hence $3/(2k+3)$ lies between b_1, b_2, etc.).

7. Suppose that inventory investment follows a partial adjustment equation, i.e. producers wish to eliminate only a fraction α of the discrepancy between the desired and actual stock of inventories. This means that the inventory investment component in Y_t is no longer $\hat{Q}_t - Q_{t-1}$, but $\alpha(\hat{Q}_t - Q_{t-1})$. Show that the model becomes stable for sufficiently low values of α.

8.4 References

Biederman, D.K., 1993, Permanent Income and Long-run Stability in a Generalized Multiplier/Accelerator Model.

Blinder, A.S. and L.J.Maccini, 1991, The Resurgence of Inventory Research: What Have We Learned?

Fiorito, R. (ed.), 1994, *Inventory, Business Cycles and Monetary Transmission.*

Hicks, J.R., 1950, *A Contribution to the Theory of the Trade Cycle.*

Kenkel, J.L., 1974, *Dynamic Linear Economic Models,* Chap. 9.

Metzler, L.A., 1941, The Nature and Stability of Inventory Cycles.

Chapter 9

Simultaneous Systems of Difference Equations

Two (or more) equations in which two (or more) unknown functions are involved, form a simultaneous system. For the system to be solvable, the number of equations must be equal to the number of unknowns, provided that the equations are independent and consistent.

9.1 First-order 2×2 systems in normal form

The simplest type of system is the following first-order system in 'normal' form

$$\begin{aligned} y_{t+1} &= a_{11}y_t + a_{12}z_t + g_1(t), \\ z_{t+1} &= a_{21}y_t + a_{22}z_t + g_2(t), \end{aligned} \tag{9.1}$$

where the coefficients a_{ij} are given constants and $g_1(t), g_2(t)$ are known functions. The system is first-order because only two adjacent points of time appear ($t+1$ and t; or t and $t-1$) and is called 'normal' because in each equation only one unknown function in turn appears evaluated at time $t+1$. It must be noted that the order of a system is equal to the degree of the characteristic equation (see below). Therefore, the expression 'first-order system' must be understood only in the sense defined above, and not as indicating the order of the system (actually the order of system (9.1) is 2).

System (9.1) is a non-homogeneous system. As in the case of single equations the general solution of (9.1) is obtained adding a particular solution to the general solution of the corresponding homogeneous system.

9.1.1 General solution of the homogeneous system: first method

Let us write the homogenous system corresponding to (9.1)

$$\begin{aligned} y_{t+1} &= a_{11}y_t + a_{12}z_t, \\ z_{t+1} &= a_{21}y_t + a_{22}z_t. \end{aligned} \tag{9.2}$$

A method of solving (9.2) consists of reducing it to a single equation in which only one unknown function appears; this reduction is always possible by suitable transformations. From the first equation of system (9.2) we obtain, provided that $a_{12} \neq 0$[1],

$$z_t = \frac{1}{a_{12}}y_{t+1} - \frac{a_{11}}{a_{12}}y_t, \tag{9.3}$$

whence

$$z_{t+1} = \frac{1}{a_{12}}y_{t+2} - \frac{a_{11}}{a_{12}}y_{t+1}. \tag{9.4}$$

Substituting (9.3) and (9.4) in the second equation of system (9.2), we have

$$\frac{1}{a_{12}}y_{t+2} - \frac{a_{11}}{a_{12}}y_{t+1} = a_{21}y_t + \frac{a_{22}}{a_{12}}y_{t+1} - \frac{a_{11}a_{22}}{a_{12}}y_t.$$

Multiplying both members by a_{12} and rearranging terms, we have

$$y_{t+2} - (a_{11} + a_{22})y_{t+1} - (a_{12}a_{21} - a_{11}a_{22})y_t = 0, \tag{9.5}$$

i.e.

$$y_t - (a_{11} + a_{22})y_{t-1} - (a_{12}a_{21} - a_{11}a_{22})y_{t-2} = 0. \tag{9.6}$$

Thus we have a second-order difference equation in which only one unknown function, y_t, appears. We solve it in the usual way, and to obtain z_t we have only to substitute in (9.3). Suppose for example that the roots of the characteristic equation of (9.6) are real and distinct. Then the general solution of (9.6) is, as we know,

$$y_t = A_1\lambda_1^t + A_2\lambda_2^t, \tag{9.7}$$

where A_1 and A_2 are arbitrary constants. Substituting (9.7) in (9.3) we have

$$z_t = \frac{A_1\lambda_1^{t+1} + A_2\lambda_2^{t+1}}{a_{12}} - \frac{a_{11}(A_1\lambda_1^t + A_2\lambda_2^t)}{a_{12}}.$$

Since $\lambda_1^{t+1} = \lambda_1\lambda_1^t, \lambda_2^{t+1} = \lambda_2\lambda_2^t$, we can collect terms and obtain

$$z_t = \frac{\lambda_1 - a_{11}}{a_{12}}A_1\lambda_1^t + \frac{\lambda_2 - a_{11}}{a_{12}}A_2\lambda_2^t. \tag{9.8}$$

[1]If $a_{12} = 0$, then we use the second equation to isolate y_t etc. (the steps are the same as those expounded in the text for z_t; note that when $a_{12} = 0, a_{21}$ must be different from zero, since, if also $a_{21} = 0$, we would no more have a simultaneous system, but two separate equations with no interdependency). Another method when either a_{12} or a_{21} are zero is to solve separately the equation where only one unknown function appears and to substitute the result in the other equation.

Eqs. (9.7) and (9.8) are the general solution of system (9.2); of course, if the roots of the characteristic equation (9.6) are real and equal or are complex, we shall proceed as explained in Chap. 5 to obtain y_t, and then substitute in (9.3) to obtain z_t. If such roots are real and equal, the solution for y_t is, as we know,

$$y_t = (A_1 + A_2t)\lambda^{*t}. \tag{9.9}$$

Substituting in (9.3) we have

$$\begin{aligned} z_t &= \tfrac{1}{a_{12}}[A_1 + A_2(t+1)]\lambda^{*t+1} - \tfrac{a_{11}}{a_{12}}(A_1 + A_2t)\lambda^{*t} \\ &= \lambda^{*t}\left[\tfrac{\lambda^*}{a_{12}}(A_1 + A_2 + A_2t) - \tfrac{a_{11}}{a_{12}}(A_1 + A_2t)\right]; \end{aligned}$$

therefore

$$z_t = \left[\frac{(\lambda^* - a_{11})A_1 + \lambda^*A_2}{a_{12}} + \frac{\lambda^* - a_{11}}{a_{12}}A_2t\right]\lambda^{*t}. \tag{9.10}$$

Using the fact that $\lambda^* = \frac{1}{2}(a_{11} + a_{22})$, Eq. (9.10) may also be written

$$z_t = \left[\frac{(a_{22} - a_{11})A_1 + (a_{11} + a_{22})A_2}{2a_{12}} + \frac{a_{22} - a_{11}}{2a_{12}}A_2t\right]\lambda^{*t}. \tag{9.11}$$

If, finally, the roots of the characteristic equation of (9.6) are complex, then, as we know, oscillations will take place according to the formula

$$y_t = r^t(A_1\cos\omega t + A_2\sin\omega t). \tag{9.12}$$

Substituting in (9.3) we have

$$\begin{aligned} z_t &= \frac{r^{t+1}[A_1\cos(\omega t + \omega) + A_2\sin(\omega t + \omega)] - a_{11}r^t(A_1\cos\omega t + A_2\sin\omega t)}{a_{12}} \\ &= r^t\left\{\frac{r[A_1(\cos\omega t\cos\omega - \sin\omega t\sin\omega) + A_2(\sin\omega t\cos\omega + \cos\omega t\sin\omega)]}{a_{12}}\right. \\ &\qquad \left. - \frac{a_{11}(A_1\cos\omega t + A_2\sin\omega t)}{a_{12}}\right\}, \end{aligned}$$

which gives

$$\begin{aligned} z_t = r^t\Bigg(&\frac{A_1r\cos\omega + A_2r\sin\omega - a_{11}A_1}{a_{12}}\cos\omega t \\ &+ \frac{A_2r\cos\omega - A_1r\sin\omega - a_{11}A_2}{a_{12}}\sin\omega t\Bigg). \end{aligned} \tag{9.13}$$

The method that we have illustrated is undoubtedly fairly simple and yields the required solution. There is, however, another method, which (although it may seem more complicated) has the advantage of being more direct, in the sense that it gives simultaneously the unknown functions y_t and z_t, without any need to reduce the system to a single equation in one unknown function. Such a method, moreover, can easily be generalized.

9.1.2 General solution of the homogeneous system: second (or direct) method

The alternative method consists in trying directly as a solution—by analogy with single equations—the functions $y_t = \alpha_1\lambda^t$, $z_t = \alpha_2\lambda^t$, where α_1, α_2 are constants not both zero. Substituting in (9.2) we have

$$\begin{aligned} \alpha_1\lambda^{t+1} &= a_{11}\alpha_1\lambda^t + a_{12}\alpha_2\lambda^t, \\ \alpha_2\lambda^{t+1} &= a_{21}\alpha_1\lambda^t + a_{22}\alpha_2\lambda^t, \end{aligned}$$

from which

$$\begin{aligned} \lambda^t\left[(a_{11}-\lambda)\alpha_1 + a_{12}\alpha_2\right] &= 0, \\ \lambda^t\left[a_{21}\alpha_1 + (a_{22}-\lambda)\alpha_2\right] &= 0. \end{aligned} \tag{9.14}$$

The functions that we have tried will be the solution of system (9.2) if, and only if, system (9.14) is satisfied for any t, i.e. (apart from the trivial case $\lambda = 0$) if, and only if,

$$\begin{aligned} (a_{11}-\lambda)\alpha_1 + a_{12}\alpha_2 &= 0, \\ a_{21}\alpha_1 + (a_{22}-\lambda)\alpha_2 &= 0. \end{aligned} \tag{9.15}$$

System (9.15) has the trivial solution $\alpha_1 = \alpha_2 = 0$, but we have excluded it from the beginning for obvious reasons. From elementary algebra we know that the necessary and sufficient condition for a linear homogeneous system to have non-trivial solutions, in addition to the trivial one, is that the determinant of the system be zero. In our case, then, it must be

$$\begin{vmatrix} a_{11}-\lambda & a_{12} \\ a_{21} & a_{22}-\lambda \end{vmatrix} = 0. \tag{9.16}$$

Expanding the determinant we have

$$\begin{aligned} &(a_{11}-\lambda)(a_{22}-\lambda) - a_{12}a_{21} \\ &= \lambda^2 - (a_{11}+a_{22})\lambda + (a_{11}a_{22} - a_{12}a_{21}) = 0. \end{aligned} \tag{9.17}$$

The determinantal equation (9.16) and its expanded form (9.17) are called the characteristic equation of the system of difference equations (9.2). Let us note that such an equation is the same as the characteristic equation of (9.6) above, and this is correct since the λ values must be the same whichever method is followed to solve the system of difference equations.

9.1.2.1 Unequal real roots

From the solution of (9.17) we obtain two values of λ; let us assume for the moment that they are real and distinct. Thus the determinant of system (9.15) equals zero for $\lambda = \lambda_1$ and for $\lambda = \lambda_2$; correspondingly, we shall have two solutions of that system. Let us call $[\alpha_1^{(1)}, \alpha_2^{(1)}]$ the solution that we obtain putting $\lambda = \lambda_1$ in (9.15) and $[\alpha_1^{(2)}, \alpha_2^{(2)}]$ the solution that we have putting $\lambda = \lambda_2$. For $\lambda = \lambda_1$ we have

$$\begin{aligned} (a_{11}-\lambda_1)\alpha_1^{(1)} + a_{12}\alpha_2^{(1)} &= 0, \\ a_{21}\alpha_1^{(1)} + (a_{22}-\lambda_1)\alpha_2^{(1)} &= 0. \end{aligned} \tag{9.18}$$

From elementary algebra we know that, since the determinant of the system is zero, we can fix the value of one of the unknowns arbitrarily[2] and then determine the value of the other (in other words, only the ratio between the two unknowns is determined). We choose to fix $\alpha_1^{(1)} = 1$ whence, from the first equation of (9.18),

$$\alpha_2^{(1)} = \frac{\lambda_1 - a_{11}}{a_{12}}.$$

Note that, from the second equation, $\alpha_2^{(1)} = a_{21}/(\lambda_1 - a_{22})$. The two values, however, are equal since λ_1 is a root of the characteristic equation, i.e.

$$(a_{11}-\lambda_1)(a_{22}-\lambda_1) - a_{12}a_{21} = 0,$$

so that

$$\frac{\lambda_1 - a_{11}}{a_{12}} = \frac{a_{21}}{\lambda_1 - a_{22}}.$$

In a similar way for $\lambda = \lambda_2$ we fix $\alpha_1^{(2)} = 1$ and obtain

$$\alpha_2^{(2)} = \frac{\lambda_2 - a_{11}}{a_{12}} = \frac{a_{21}}{\lambda_2 - a_{22}}.$$

Thus we have reached the result that $y_{1t} = \alpha_1^{(1)}\lambda_1^t$, $z_{1t} = \alpha_2^{(1)}\lambda_1^t$ is a solution of system (9.2) and $y_{2t} = \alpha_1^{(2)}\lambda_2^t$, $z_{2t} = \alpha_2^{(2)}\lambda_2^t$ is another solution. We can then combine them linearly with two arbitrary constants A_1, A_2 and obtain the general solution of system (9.2):

$$y_t = A_1\alpha_1^{(1)}\lambda_1^t + A_2\alpha_1^{(2)}\lambda_2^t, \tag{9.19}$$

$$z_t = A_1\alpha_2^{(1)}\lambda_1^t + A_2\alpha_2^{(2)}\lambda_2^t, \tag{9.20}$$

i.e., with the values of $\alpha_i^{(j)}, i, j = 1, 2$, found above,

$$y_t = A_1\lambda_1^t + A_2\lambda_2^t, \tag{9.21}$$

[2] Such arbitrariness does not give any trouble in the solution of the system (9.2), since it combines in a multiplicative way with the arbitrary constants A_1, A_2 which appear in the general solution. Thus we have chosen $\alpha_1^{(1)} = 1$ (and similarly we shall choose $\alpha_1^{(2)} = 1$) so that the solution of system (9.2) that we shall obtain will be immediately comparable, without any need of further manipulations, with (9.7) and (9.8) above.

$$z_t = A_1\frac{\lambda_1 - a_{11}}{a_{12}}\lambda_1^t + A_2\frac{\lambda_2 - a_{11}}{a_{12}}\lambda_2^t, \tag{9.22}$$

which are the same as Eq. (9.7) and Eq. (9.8). The student may check by direct substitution that (9.21) and (9.22) indeed satisfy system (9.2). The number of arbitrary constants appearing in the general solution of this system is two, since the system is reducible to a second-order equation, as we have seen above. In general, the number of arbitrary constants is equal to the *order* of the system , that is to the *degree* of its characteristic equation.

The solution of the system may also be written in the equivalent form, which sometimes appears in the literature,

$$\begin{aligned} y_t &= A_1\lambda_1^t + A_2\lambda_2^t, \\ z_t &= A_1'\lambda_1^t + A_2'\lambda_2^t, \end{aligned} \tag{9.23}$$

where the arbitrary constants A_1, A_2, A_1', A_2' are connected by the relations

$$\frac{A_1'}{A_1} = \frac{\lambda_1 - a_{11}}{a_{12}} = \frac{a_{21}}{\lambda_1 - a_{22}}, \quad \frac{A_2'}{A_2} = \frac{\lambda_2 - a_{11}}{a_{12}} = \frac{a_{21}}{\lambda_2 - a_{22}}. \tag{9.24}$$

Since the ratios $A_1'/A_1, A_2'/A_2$ are uniquely determined in terms of the λ's and of the coefficients of the system, the independent arbitrary constants are actually only two.

9.1.2.2 Equal real roots

If the characteristic equation has two real and equal roots $\lambda_1 = \lambda_2 = \lambda^*$, let us try

$$\begin{aligned} y_t &= (A_1 + A_2 t)\lambda^{*t}, \\ z_t &= (A_1' + A_2' t)\lambda^{*t}, \end{aligned} \tag{9.25}$$

where the arbitrary constants A_1, A_2, A_1', A_2' are related in some way. Substituting in system (9.2) we have

$$\begin{aligned} \left[A_1 + A_2(t+1)\right]\lambda^{*t+1} &= a_{11}(A_1 + A_2 t)\lambda^{*t} + a_{12}(A_1' + A_2' t)\lambda^{*t}, \\ \left[A_1' + A_2'(t+1)\right]\lambda^{*t+1} &= a_{21}(A_1 + A_2 t)\lambda^{*t} + a_{22}(A_1' + A_2' t)\lambda^{*t}. \end{aligned} \tag{9.26}$$

Dividing through by $\lambda^{*t} \neq 0$ and rearranging terms we obtain

$$\begin{aligned} &\left[A_2(\lambda^* - a_{11}) - a_{12}A_2'\right]t + \left[(\lambda^* - a_{11})A_1 + A_2\lambda^* - a_{12}A_1'\right] = 0, \\ &\left[(\lambda^* - a_{22})A_2' - a_{21}A_2\right]t + \left[(\lambda^* - a_{22})A_1' - a_{21}A_1 + A_2'\lambda^*\right] = 0. \end{aligned} \tag{9.27}$$

Eqs. (9.27) are identically satisfied if, and only if, the expressions in square brackets are all zero, i.e.

$$A_2(\lambda^* - a_{11}) - a_{12}A_2' = 0, \tag{9.28}$$
$$(\lambda^* - a_{11})A_1 + A_2\lambda^* - a_{12}A_1' = 0, \tag{9.29}$$
$$(\lambda^* - a_{22})A_2' - a_{21}A_2 = 0, \tag{9.30}$$
$$(\lambda^* - a_{22})A_1' - a_{21}A_1 + A_2'\lambda^* = 0, \tag{9.31}$$

whence

$$A_2' = \frac{\lambda^* - a_{11}}{a_{12}}A_2, \tag{9.32}$$
$$A_1' = \frac{(\lambda^* - a_{11})A_1 + \lambda^* A_2}{a_{12}}, \tag{9.33}$$
$$A_2' = \frac{a_{21}}{\lambda^* - a_{22}}A_2, \tag{9.34}$$
$$A_1' = \frac{a_{21}}{\lambda^* - a_{22}}A_1 - \frac{\lambda^*}{\lambda^* - a_{22}}A_2'. \tag{9.35}$$

Since λ^* is a root of the characteristic equation, $(\lambda^* - a_{11})/a_{12} = a_{21}/(\lambda^* - a_{22})$ and so (9.32) and (9.30) coincide. Using (9.32) and the fact that $(\lambda^* - a_{11})/a_{12} = a_{21}/(\lambda^* - a_{22})$,(9.35) can be written as

$$A_1' = \frac{\lambda^* - a_{11}}{a_{12}}A_1 - \frac{\lambda^*(\lambda^* - a_{11})}{a_{12}(\lambda^* - a_{22})}A_2, \tag{9.36}$$

which coincides with (9.33) if, and only if, $-(\lambda^* - a_{11}) = \lambda^* - a_{22}$, i.e. if, and only if, $\lambda^* = \frac{1}{2}(a_{11} + a_{22})$, which is indeed true if, and only if, λ^* is a double root of the characteristic Eq. (9.17). Thus (9.25) is indeed the (general) solution of the system; using (9.33) and (9.32) to express A_1', A_2' in terms of A_1, A_2 it can be seen that Eqs. (9.25) coincide with (9.9) and (9.10) above.

9.1.2.3 Complex roots

If the roots of the characteristic equation are complex, $\lambda_1, \lambda_2 = \alpha \pm i\theta$, then the solution can at first be written (the procedure is the same as for the case of distinct real roots) as

$$\begin{aligned} y_t &= B_1(\alpha + i\theta)^t + B_2(\alpha - i\theta)^t, \\ z_t &= \frac{(\alpha + i\theta) - a_{11}}{a_{12}}B_1(\alpha + i\theta)^t + \frac{(\alpha - i\theta) - a_{11}}{a_{12}}B_2(\alpha - i\theta)^t \\ &= \frac{B_1(\alpha + i\theta)^{t+1} + B_2(\alpha - i\theta)^{t+1}}{a_{12}} - \frac{a_{11}\left[B_1(\alpha + i\theta)^t + B_2(\alpha - i\theta)^t\right]}{a_{12}}, \end{aligned} \tag{9.37}$$

where B_1, B_2 are arbitrary complex conjugate constants. Using standard transformations of complex numbers (see above, Chap. 5, Sect. 5.1.3), we obtain

$$\begin{aligned} B_1(\alpha + i\theta)^t + B_2(\alpha - i\theta)^t &= r^t(A_1\cos\omega t + A_2\sin\omega t), \\ B_1(\alpha + i\theta)^{t+1} + B_2(\alpha - i\theta)^{t+1} &= r^{t+1}[A_1\cos(\omega t + \omega) + A_2\sin(\omega t + \omega)] \\ &= r^{t+1}[(A_1\cos\omega + A_2\sin\omega)\cos\omega t \\ &\quad + (A_2\cos\omega - A_1\sin\omega)\sin\omega t], \end{aligned} \tag{9.38}$$

where $A_1 \equiv (B_1 + B_2)$, $A_2 \equiv (B_1 - B_2)i$ are arbitrary real constants and r, ω are related to α, θ in the usual way. Substituting (9.38) in Eqs.(9.37) and collecting terms where necessary we obtain

$$\begin{aligned} y_t &= r^t(A_1 \cos \omega t + A_2 \sin \omega t), \\ z_t &= r^t \left(\frac{A_1 r \cos \omega + A_2 r \sin \omega - a_{11}A_1}{a_{12}} \cos \omega t \right. \\ &\quad \left. + \frac{A_2 r \cos \omega - A_1 r \sin \omega - a_{11}A_2}{a_{12}} \sin \omega t \right), \end{aligned} \tag{9.39}$$

which coincide with (9.12) and (9.14). This completes the exposition of the 'direct' method of solution.

9.1.3 Particular solution. Determination of the arbitrary constants

We can now turn to the problem of finding a particular solution of the non-homogeneous system. The general method of undetermined coefficients can be applied here too, and we shall illustrate it in the case in which $g_1(t), g_2(t)$ are two given constants, say b_1, b_2. Thus we have

$$\begin{aligned} y_{t+1} &= a_{11}y_t + a_{12}z_t + b_1, \\ z_{t+1} &= a_{21}y_t + a_{22}z_t + b_2. \end{aligned} \tag{9.40}$$

As a particular solution let us try $\overline{y}_t = \mu_1, \overline{z}_t = \mu_2$, where μ_1, μ_2 are undetermined constants. Substituting in Eq. (9.40) and rearranging terms we have

$$\begin{aligned} (a_{11} - 1)\mu_1 + a_{12}\mu_2 &= -b_1, \\ a_{21}\mu_1 + (a_{22} - 1)\mu_2 &= -b_2, \end{aligned} \tag{9.41}$$

whence

$$\mu_1 = \frac{-b_1(a_{22} - 1) + b_2 a_{12}}{(a_{11} - 1)(a_{22} - 1) - a_{12}a_{21}}, \quad \mu_2 = \frac{-b_2(a_{11} - 1) + b_1 a_{21}}{(a_{11} - 1)(a_{22} - 1) - a_{12}a_{21}}. \tag{9.42}$$

The method is successful only if the determinant of system (9.41) is different from zero, i.e.

$$\begin{vmatrix} a_{11} - 1 & a_{12} \\ a_{21} & a_{22} - 1 \end{vmatrix} \neq 0. \tag{9.43}$$

Note that this condition means that $+1$ is *not* a root of the characteristic equation of the homogeneous difference system (9.2). When this occurs, the

method fails; a way out is to try $\overline{y}_t = \mu_{11} + \mu_{12}t, \overline{z}_t = \mu_{21} + \mu_{22}t$, where μ_{ij} are undetermined constants.

Finally, the determination of the arbitrary constants that appear in the general solution can be made as usual by means of a number of additional conditions equal to the number of arbitrary constants. These conditions give the information that, for a given value of t (usually for $t = 0$, whence the name *initial conditions*), the values of the various functions are known values (e.g., that y_0, z_0 are known). Substituting in the general solution of the system under consideration, we obtain a system of linear equations which can be solved for the values of the arbitrary constants.

9.2 First order n×n systems in normal form

Let us first observe that a $n \times n$ system can always be reduced, by a procedure similar to that used in relation to the 2×2 system, to a single nth order equation in one unknown function. Incidentally, note that the converse is also true, i.e. a nth order equation can always be transformed into a first-order system in normal form having n equations in n unknown functions. To do this, new variables $y_{1t}, y_{2t}, \ldots, y_{nt}$ are defined such that $y_{t+1} = y_{1t}, y_{1t+1} = y_{2t}, \ldots, y_{n-2,t+1} = y_{n-1,t}$. Substituting in the given nth order equation in y, a first-order equation in $y_{n-1,t+1}, y_{n-1,t}, \ldots, y_{1t}, y_t$ is obtained, that—together with the equations defining the new variables—forms a first-order system in normal form (see Sect. 9.4.2, exercise C).

The reduction process is however time-consuming when the number of equations increases, hence it is preferable to apply the direct method of solution. If we consider for example the homogeneous system

$$\begin{aligned} y_{t+1} &= a_{11}y_t + a_{12}z_t + a_{13}w_t, \\ z_{t+1} &= a_{21}y_t + a_{22}z_t + a_{23}w_t, \\ w_{t+1} &= a_{31}y_t + a_{32}z_t + a_{33}w_t, \end{aligned}$$

we can immediately write its characteristic equation

$$D(\lambda) = \begin{vmatrix} a_{11} - \lambda & a_{12} & a_{13} \\ a_{21} & a_{22} - \lambda & a_{23} \\ a_{31} & a_{32} & a_{33} - \lambda \end{vmatrix} = 0. \tag{9.44}$$

Expanding the determinant we have a third-degree algebraic equation in the unknown λ, whose solution will give three values, $\lambda_1, \lambda_2, \lambda_3$. Since the stable or unstable behaviour over time of the solution depends exclusively on the roots $\lambda_1, \lambda_2, \lambda_3$, to analyse the stability of the system we can examine only the nature of such roots, without any need to compute the coefficients $\alpha_i^{(j)}$. For this purpose we can apply to the characteristic equation the stability conditions stated in Chap. 7, which allow us to check whether the roots of a polynomial are in absolute value less then unity without finding them explicitly.

With the help of a little matrix algebra, the direct method can easily be generalized to first-order systems of type (9.1) having any number of equations.

In general, a $n \times n$ first-order system in normal form has the typical matrix form

$$\mathbf{y}_{t+1} = \mathbf{A}\mathbf{y}_t + \mathbf{g}(t), \tag{9.45}$$

where $\mathbf{y}_t = [y_{1t}, y_{2t}, ..., y_{nt}]$ is the column vector of the unknown functions of time to be found,

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \tag{9.46}$$

is the square matrix of the given coefficients, and $\mathbf{g}(t) = [g_1(t), g_2(t), ..., g_n(t)]$ is a column vector of known functions of time.

Let us begin with the homogeneous system

$$\mathbf{y}_{t+1} = \mathbf{A}\mathbf{y}_t. \tag{9.47}$$

An immediate extension of the direct method of solution leads us to consider the possible solution $\mathbf{y}_t = \boldsymbol{\alpha}\lambda^t$, where $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, ..., \alpha_n]$ is a vector of constants not all zero. By substituting into (9.47) we get

$$\boldsymbol{\alpha}\lambda^{t+1} = \mathbf{A}\boldsymbol{\alpha}\lambda^t, \tag{9.48}$$

i.e.

$$\lambda^t[\mathbf{A} - \lambda\mathbf{I}]\boldsymbol{\alpha} = \mathbf{0}, \tag{9.49}$$

which will be identically satisfied if, and only if,

$$[\mathbf{A} - \lambda\mathbf{I}]\boldsymbol{\alpha} = \mathbf{0}. \tag{9.50}$$

System (9.50) will have a non-trivial solution ($\boldsymbol{\alpha} \neq \mathbf{0}$) if, and only if, its determinant is zero, namely

$$|\mathbf{A} - \lambda\mathbf{I}| = 0, \tag{9.51}$$

which is the determinantal form of the characteristic equation of the matrix $\mathbf{A}$. Expansion of this determinant yields an nth order polynomial equation in λ of the type

$$(-1)^n\lambda^n + c_1(-1)^{n-1}\lambda^{n-1} + ... + c_r(-1)^r\lambda^r + ... + c_{n-1}(-1)\lambda + c_n = 0 \tag{9.52}$$

or, multiplying through by $(-1)^n$ and taking into account that $(-1)^{2n} = 1$, $(-1)^{2n-1} = -1$,

$$\lambda^n - c_1\lambda^{n-1} + ... + c_r(-1)^{n+r}\lambda^r + ... + c_{n-1}(-1)^{n+1}\lambda + (-1)^n c_n = 0, \tag{9.53}$$

where

$$\begin{aligned} c_1 &= \textstyle\sum_{i=1}^n a_{ii}, \\ c_2 &= \text{sum of all second-order principal minors of the matrix } \mathbf{A}, \\ \dots &\dots \dots \\ c_r &= \text{sum of all } n!/r!(n-r)! \text{ principal minors of the } r\text{th order}, \\ \dots &\dots \dots \\ c_n &= |\mathbf{A}|. \end{aligned} \tag{9.54}$$

The solution of Eq. (9.53) will give n roots $\lambda_1, \lambda_2, ..., \lambda_n$, the characteristic roots (or latent roots or eigenvalues) of the matrix $\mathbf{A}$. To each eigenvalue λ_j a corresponding eigenvector (or latent vector or characteristic vector) $\boldsymbol{\alpha}^{(j)}$ can be associated through (9.50), as shown in Sect. 9.1.2 for the 2×2 case.

The solution of the difference equation system under consideration will then have the form

$$\begin{aligned} y_{1t} &= A_1\alpha_1^{(1)}\lambda_1^t + A_2\alpha_1^{(2)}\lambda_2^t + ... + A_n\alpha_1^{(n)}\lambda_n^t, \\ y_{2t} &= A_1\alpha_2^{(1)}\lambda_1^t + A_2\alpha_2^{(2)}\lambda_2^t + ... + A_n\alpha_2^{(n)}\lambda_n^t, \\ \dots &\dots \dots\dots\dots\dots \\ y_{nt} &= A_1\alpha_n^{(1)}\lambda_1^t + A_2\alpha_n^{(2)}\lambda_2^t + ... + A_n\alpha_n^{(n)}\lambda_n^t, \end{aligned} \tag{9.55}$$

if the eigenvalues are real and distinct; in the case of multiple and/or complex roots we proceed as shown above in relation to the 2×2 system.

To determine the vector of the arbitrary constants $\mathbf{a} = [A_1, A_2, ..., A_n]$ we need n additional conditions, for example $\mathbf{y}_t = \mathbf{y}_0$, where $\mathbf{y}_0$ is a vector of known values. From Eqs. (9.55) we then have, since $\lambda^0 = 1$,

$$\mathbf{y}_0 = \mathbf{V}\mathbf{a}, \tag{9.56}$$

where $\mathbf{V}$, the matrix of system (9.55), is the matrix whose columns are the characteristic vectors of the matrix $\mathbf{A}$. From elementary matrix algebra we know that the characteristic vectors associated with distinct characteristic roots are linearly independent, hence $\mathbf{V}$ is non-singular. Therefore

$$\mathbf{a} = \mathbf{V}^{-1}\mathbf{y}_0 \tag{9.57}$$

yields the required solution for the arbitrary constants.

In the case of multiple roots we can no longer use the property of linear independence, but it is enough to observe that the general solution will consist of the linear combination of n distinct solutions forming a fundamental set, so that the appropriate matrices will always be non singular.

9.2.1 Direct matrix solution

We have seen in Chap. 3 that the solution of the simple first-order scalar difference equation $y_{t+1} = ay_t$ is of the type $y_t = a^t C$, where C is an arbitrary constant. Analogously we would like to be able to say that the solution of the vector difference equation (9.47) is

$$\mathbf{y}_t = \mathbf{A}^t \mathbf{c}, \tag{9.58}$$

where $\mathbf{c}$ is a $(n \times 1)$ vector of arbitrary constants. This is indeed true, for by elementary matrix algebra the integer powers of a matrix are defined exactly as the integer powers of a scalar. Hence, by substitution in Eq. (9.47) we get

$$\mathbf{A}^{t+1}\mathbf{c} = \mathbf{A}(\mathbf{A}^t\mathbf{c}) = \mathbf{A}^{t+1}\mathbf{c}. \tag{9.59}$$

Since Eq. (9.59) is an identity, Eq. (9.58), which contains n arbitrary constants, is the general solution of the vector difference equation (9.47).

Of course the formal solution (9.58), apart from its use in computations, is of little analytical value unless a closed-form expression for $\mathbf{A}^t$ can be found. This is fairly easy if we assume that $\mathbf{A}$ has distinct latent roots. In this case we know from matrix algebra that $\mathbf{A}$ can be diagonalized by a similar transformation with the matrix of the associated characteristic vectors (the modal matrix $\mathbf{V}$), namely

$$\mathbf{\Lambda} = \mathbf{V}^{-1}\mathbf{A}\mathbf{V} \quad \text{or} \quad \mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}, \tag{9.60}$$

where

$$\mathbf{\Lambda} = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \lambda_n \end{bmatrix} \tag{9.61}$$

is the diagonal matrix of the roots of $\mathbf{A}$. We now note that

$$\begin{aligned} \mathbf{A}^2 &= (\mathbf{V\Lambda V}^{-1})(\mathbf{V\Lambda V}^{-1}) = \mathbf{V\Lambda}(\mathbf{V}^{-1}\mathbf{V})\mathbf{\Lambda V}^{-1} \\ &= \mathbf{V\Lambda}^2\mathbf{V}^{-1} \end{aligned}$$

since $\mathbf{V}^{-1}\mathbf{V} = \mathbf{I}$. In a similar way we have

$$\mathbf{A}^t = \mathbf{V\Lambda}^t\mathbf{V}^{-1}. \tag{9.62}$$

Thus the solution (9.58) can be written as

$$\mathbf{y}_t = \mathbf{A}^t\mathbf{c} = \mathbf{V\Lambda}^t\mathbf{V}^{-1}\mathbf{c} = \mathbf{V\Lambda}^t(\mathbf{V}^{-1}\mathbf{c}). \tag{9.63}$$

Given the arbitrariness of $\mathbf{c}$, we can set $\mathbf{V}^{-1}\mathbf{c} = \mathbf{a}$, a new vector of arbitrary constants. We finally arrive at the expression

$$\mathbf{y}_t = \mathbf{V\Lambda}^t\mathbf{a}, \tag{9.64}$$

that coincides with the closed-form eigenvalue-eigenvector solution previously found, Eq. (9.55). The equivalence between the solution (9.58 and (9.64) also shows that $\mathbf{A}^t$ tending to zero as $t \to \infty$ is equivalent to the roots of $\mathbf{A}$ being in absolute value all smaller than unity.

9.2.2 Stability conditions

Since the stable or unstable behaviour over time of the solution depends exclusively on the roots $\lambda_1, \lambda_2, \dots, \lambda_n$, to analyse the stability of the system we can examine only the nature of such roots, without any need to compute the elements of the eigenvectors, $\alpha_i^{(j)}$. For this purpose we can apply to the characteristic equation (9.53) the stability conditions stated in Chap. 7, which allow us to check whether the roots of a polynomial are in absolute value less then unity without finding them explicitly.

However, the expansion of the determinant is rather laborious if n is great and if, as is the case in theoretical work, we have to deal with symbolic rather than numeric values of the matrix coefficients. Hence it would be highly desirable to have stability conditions which can be applied directly to the coefficients a_{ij} of the system, *without having to expand the determinant.* Such conditions exist, and the most important are listed below (further results are contained in Murata, 1977, Chaps. 3 and 4). Since the proofs require the knowledge of some advanced matrix algebra, they are given in Sect. (9.2.2.2), that can be skipped without loss of continuity.

In what follows, by 'stability conditions' we mean, as usual, 'conditions for the roots of the characteristic equation (9.53), be they real and/or complex, to be less than unity in absolute value' (or, which is the same thing, for these roots to lie within the unit circle in the complex plane). It is important to note that, when these conditions are applied to economic models, one should pay attention to whether the condition being applied is necessary and sufficient, or only necessary, or only sufficient.

I. Let $a_{ij} \geq 0$ (the coefficients must be non-negative, with at least one strictly positive). Then necessary and sufficient conditions are that the leading principal minors of the matrix $[\mathbf{I} - \mathbf{A}]$ be all positive, i.e.

$$1 - a_{11} > 0, \quad \begin{vmatrix} 1 - a_{11} & -a_{12} \\ -a_{21} & 1 - a_{22} \end{vmatrix} > 0, \quad \begin{vmatrix} 1 - a_{11} & -a_{12} & -a_{13} \\ -a_{21} & 1 - a_{22} & -a_{23} \\ -a_{31} & -a_{32} & 1 - a_{33} \end{vmatrix} > 0,$$

$$\dots, \begin{vmatrix} 1 - a_{11} & -a_{12} & \dots & -a_{1n} \\ -a_{21} & 1 - a_{22} & \dots & -a_{2n} \\ \dots & \dots & \dots & \dots \\ -a_{n1} & -a_{n2} & \dots & 1 - a_{nn} \end{vmatrix} > 0. \tag{9.65}$$

II. Let $a_{ij} \geq 0$. Then a necessary and sufficient stability condition is that the matrix $[\mathbf{I} - \mathbf{A}]$ has a dominant diagonal in the extended sense, i.e. all the coefficients on the main diagonal are positive and there are positive numbers $h_1, h_2, ..., h_n$ such that

$$h_i(1 - a_{ii}) > \sum_{\substack{j=1 \\ j \neq i}}^{n} h_j a_{ij} \qquad \text{(row dominance)}, \qquad (9.66)$$

or

$$h_i(1 - a_{ii}) > \sum_{\substack{j=1 \\ j \neq i}}^{n} h_j a_{ji} \qquad \text{(column dominance)}. \qquad (9.67)$$

Equivalently,

$$h_i > \sum_{j=1}^{n} h_j a_{ij} \qquad \text{or} \qquad h_i > \sum_{j=1}^{n} h_j a_{ji}. \qquad (9.68)$$

Since conditions (I) and (II) are both necessary and sufficient, they are equivalent (see below, Sect. 9.2.1.2).

III. Let $a_{ij} > 0$ (the coefficients must be all positive). Form the n sums

$$S_j = \sum_{i=1}^{n} a_{ij}, \qquad j = 1, 2..., n. \qquad (9.69)$$

Then a set of sufficient stability conditions is that no S_j is greater than one and at least one of them is smaller than one.

IV. Let $a_{ij} \geq 0$. Then a set of sufficient stability conditions is that all S_j, as defined in Eq. (9.69), are smaller than one.

V. Let $a_{ij} \geq 0$ as in the previous case, but in addition let the matrix $\mathbf{A}$ be indecomposable. A matrix $\mathbf{A} \equiv [a_{ij}]$ is indecomposable (also called irreducible or connected) if there is no permutation of the indices (i.e., interchange of rows followed by the same interchange of columns, or viceversa) that reduces it to the form

$$\mathbf{A} = \begin{bmatrix} \mathbf{B} & \mathbf{C} \\ \mathbf{0} & \mathbf{D} \end{bmatrix}, \qquad (9.70)$$

where $\mathbf{B}, \mathbf{D}$ are square matrices (not necessarily of the same order). If such reduction is possible, the matrix is decomposable or reducible. An alternative

definition of indecomposability for a non-negative matrix $\mathbf{A}$ is that, for each pair of indices (i, j) there exists a set of indices $j_1, j_2, ..., j_l$ such that $a_{ij_1} a_{j_1 j_2} \cdots a_{j_l j} > 0$ (see Schwartz, 1961, pp. 19-20).

Now, in the case under consideration, a set of sufficient stability conditions is the same as in III above. Actually conditions V absorb conditions III, since a positive matrix is *a fortiori* indecomposable, but we have stated them separately for greater clarity.

VI. Let a_{ij} be arbitrary. Form the n sums of absolute values

$$|S_j| = \sum_{i=1}^{n} |a_{ij}|, \qquad j = 1, 2, ..., n. \qquad (9.71)$$

Then a set of sufficient stability conditions is that all $|S_j|$ are smaller than one. This condition absorbs condition IV, since $\sum_{i=1}^{n} |a_{ij}| = \sum_{i=1}^{n} a_{ij}$ when $a_{ij} \geq 0$, but we have stated them separately for greater clarity.

VII. Let a_{ij} be arbitrary, but in addition let the matrix $\mathbf{A}$ be indecomposable as defined above. Then a set of sufficient conditions is that no $|S_j|$ (as defined above) is greater than one and at least one of them is smaller than one. This condition actually absorbs condition V, but we have stated them separately for greater clarity.

VIII. Let a_{ij} be arbitrary. Form the matrix $\mathbf{A}^+ = [|a_{ij}|]$, namely the matrix whose elements are the absolute values of the corresponding elements of $\mathbf{A}$ (this operation is called *majorization*). Then a sufficient stability condition for the system whose matrix is $\mathbf{A}$ is that $[\mathbf{I} - \mathbf{A}^+]$ has a dominant positive diagonal as defined in II, or that the leading principal minors of $[\mathbf{I} - \mathbf{A}^+]$ are all positive.

IX. A necessary stability condition is that $|\sum_{i=1}^{n} a_{ii}| < n$.

X. A necessary stability condition is that the determinant of $\mathbf{A}$ be less than one in absolute value.

XI. A sufficient *in*stability condition is that all S_j, as defined in Eq. (9.69) are greater than one.

Note, finally, that conditions III, IV, V, VI, VII, XI may be phrased in terms of the sums

$$S_i = \sum_{j=1}^{n} a_{ij}, \qquad j, i = 1, 2..., n$$

and be equally true. This follows from the fact that a matrix and its transpose have the same characteristic roots.

9.2.2.1 A digression on not-wholly-unstable systems

By a 'not wholly unstable system' (or 'conditionally stable' system) we mean a system whose roots are partly greater and partly smaller than one in absolute value. Since the presence of even only one root greater than one in absolute value causes the overall movement of the system to be divergent, such a system should be classified as unstable.

However, if the initial conditions can somehow be selected, then it is clear that one can choose them in such a way that the arbitrary constant(s) associated with the unstable root(s) turn out to be zero, so that *the final solution will only contain the term(s) corresponding to the stable root(s).* Hence the system is stable conditional on the initial position.

This can be simply illustrated with reference to the 2×2 homogeneous system, whose solution, given in Eqs. (9.19)-(9.22) above, is reproduced here:

$$y_t = A_1\lambda_1^t + A_2\lambda_2^t,$$

$$z_t = A_1\alpha_1\lambda_1^t + A_2\alpha_2\lambda_2^t,$$

where α_1, α_2 are given by Eqs. (9.22).

Suppose now that $|\lambda_1| < 1, |\lambda_2| > 1$. We want to know the relations that the initial values must satisfy in order that $A_2 = 0$. For this purpose, consider the equations for the determination of A_1, A_2 :

$$\begin{aligned} y_0 &= A_1 + A_2, \\ z_0 &= A_1\alpha_1 + A_2\alpha_2, \end{aligned} \tag{9.72}$$

from which

$$A_1 = \frac{\alpha_2 y_0 - z_0}{\alpha_2 - \alpha_1}, \qquad A_2 = \frac{z_0 - \alpha_1 y_0}{\alpha_2 - \alpha_1}. \tag{9.73}$$

Then a necessary and sufficient condition for $A_2 = 0$ is

$$z_0 - \alpha_1 y_0 = 0, \text{ i.e. } z_0/y_0 = \alpha_1. \tag{9.74}$$

Alternatively, we can impose $A_2 = 0$ and use the first equation of system (9.72) to determine A_1 given y_0. From the second equation we then obtain the required value of z_0 as $z_0 = y_0\alpha_1$.

This procedure can easily be generalized to $n \times n$ systems. As shown below –see Eqs. (9.102)– the equations for the determination of the arbitrary constants in the case of a homogeneous system turn out to be

$$\sum_{j=1}^{n} A_j\alpha_i^{(j)} = y_i(0), \qquad i = 1, 2, ..., n. \tag{9.75}$$

Now assume that there are k stable and $(n-k)$ unstable roots $(0 < k < n)$. Without loss of generality we can order them so that the first k are stable. In (9.75) we then impose $A_{k+1} = A_{k+2} = ... = A_n = 0$ and use the first k equations to determine $A_1, A_2, ..., A_k$ *given* $y_1(0), y_2(0), ..., y_k(0)$. These values of $A_1, A_2, ..., A_k$ are then substituted into the last $n-k$ equations of (9.75), thus *determining* the required values of $y_{k+1}(0), y_{k+2}(0), ..., y_n(0)$.

All this reasoning might seem a curiosum. The initial conditions are given, and a given is a given. However, the appropriate choice of the initial conditions is often dictated by the nature of the economic model under consideration. As we shall see in Chap. 22 (which also requires the study of Chap. 18, where further treatment of conditional stability is given), it turns out that rational expectations models and optimal growth models belong to this category. Hence, in economics, a not-wholly-unstable system may actually turn out to be a stable system.

On the contrary, we might be interested in the opposite case, namely in keeping only one term in the final solution, and precisely the term associated with an unstable positive root. This is obviously the case if we are considering a growth model and want to find the conditions under which the model is capable of balanced growth, i.e. a state of growth in which the proportions that the variables bear to each other are constant. This, of course, is equivalent to the requirement that all variables actually (and not only asymptotically) grow at the same proportional rate. Thus balanced growth obtains if, and only if, in the solution of the homogeneous system only one term (containing a positive root greater than one) remains. Suppose that λ_j is such a root and consider the solution of Eqs. (9.75) by Cramer's rule

$$A_i = \frac{D_i}{D}, \text{where } D \equiv \begin{vmatrix} \alpha_1^{(1)} & \alpha_1^{(2)} & ... & \alpha_1^{(n)} \\ \alpha_2^{(1)} & \alpha_2^{(2)} & ... & \alpha_2^{(n)} \\ ... & ... & ... & ... \\ \alpha_n^{(1)} & \alpha_n^{(2)} & ... & \alpha_n^{(n)} \end{vmatrix}, \tag{9.76}$$

and D_i is obtained substituting the column of $y_i(0)$ in the place of the ith column of D.

Then, if we choose $y_1(0), y_2(0), ..., y_n(0)$ proportional to $\alpha_1^{(j)}, \alpha_2^{(j)}, ..., \alpha_n^{(j)}$, it follows that each D_i (except D_j) will have two proportional columns, and so $A_i = 0, i \neq j$.

This reasoning can also be applied to wholly unstable systems (i.e., systems whose roots are all unstable). In such a case we might be interested in keeping only the term containing the dominant positive root, because this term—if the matrix is positive or non-negative indecomposable—is associated with a positive characteristic vector.

9.2.2.2 Proof of the stability conditions

Conditions I through V, although they might be (and actually have been by some authors) proved independently, are essentially applications of corollaries of the theorems of Perron on positive matrices, of Frobenius on non-negative indecomposable matrices, and of the extension of the latter to arbitrary non-negative matrices. For proofs of these theorems and corollaries see, for example, Gantmacher (1959). This is the line that we shall follow below.

Here is then a sketch of proofs of the various conditions; for other proofs of some of them, see, for example, Baumol (1970), Fisher (1962), Solow (1952).

In what follows λ_M denotes the *dominant* characteristic root of $\mathbf{A}$, namely the root whose absolute value is not smaller than the absolute value of any other root.

To prove conditions I we start from the theorem according to which a necessary and sufficient condition for the real number ρ to be greater than the dominant root of a non-negative matrix $\mathbf{A}$, is that all the leading principal minors of the matrix $[\rho\mathbf{I} - \mathbf{A}]$ be positive (Gantmacher, 1959, pp. 85 and 88). Now, if we let $\rho = 1$, conditions I immediately follow. For a proof of the equivalence of conditions I and II see Murata (1977, p. 97).

Conditions III and V follow from the fact that for indecomposable non-negative matrices (and so *a fortiori* for positive matrices),

$$\min S_j \le \lambda_M \le \max S_j, \tag{9.77}$$

where the sign of equality to the left or right of λ_M holds *only* when $\min S_j = \max S_j$ (Gantmacher, 1959, p. 76. He uses row sums, but the same property holds for column sums, since a matrix and its transpose have the same eigenvalues). It follows that $\max S_j = 1$ is a sufficient condition for $\lambda_M < 1$, provided that $\min S_j < 1$.

Conditions IV follow from the fact that for decomposable non-negative matrices, conditions (9.71) continue to hold, but now the sign of equality to the left or right of λ_M may hold also when $\min S_j \neq \max S_j$. Hence $\max S_j = 1$ must be excluded to obtain a sufficient condition for $\lambda_M < 1$.

Note that from the above bounds for λ_M it follows that a sufficient *in*stability condition is that $\min S_j > 1$, and this is condition XI.

Conditions VI follow from the theorem (see, for example, McKenzie, 1960) according to which, if λ is any characteristic root of an arbitrary matrix $\mathbf{A}$, then

$$|\lambda| \le \max_j \sum_{i=1}^{n} |a_{ij}|. \tag{9.78}$$

More generally, conditions VI can be considered a particular case of the theorem (see, for example, Conlisk, 1973) according to which for an arbitrary matrix $\mathbf{A}$ a sufficient stability condition is $f(\mathbf{A}) < 1$, where $f(\cdot)$ is any matrix norm. Using the column sum norm, namely $\max_j \sum_{i=1}^{n} |a_{ij}|$, conditions VI follow.

To prove conditions VIII (see below for conditions VII), first observe that $\mathbf{A}^+$ is a non-negative matrix, to which conditions I and II can be applied . Then the sufficiency of conditions VIII as regards $\mathbf{A}$ follow from a theorem (see, for example, Taussky, 1964, p. 127) according to which the dominant root of an arbitrary matrix $\mathbf{A}$ is in absolute value not greater than the dominant root of $\mathbf{A}^+$. Note that this theorem also allows an immediate proof of conditions VII given conditions IV, as well as of conditions VI given conditions IV.

To prove conditions IX and X, let us recall from elementary equation theory that, in any polynomial equation such as Eq. (9.53), the coefficient of λ^{n-1} is equal to the sum of the roots multiplied by -1, while the constant term is equal to the product of the roots multiplied by $(-1)^n$. Thus we have $\sum a_{ii} = \sum \lambda_i$ and, taking absolute values, $|\sum a_{ii}| = |\sum \lambda_i|$. Now, $|\sum \lambda_i| \le \sum |\lambda_i|$ and, if $|\lambda_i| < 1$, then $\sum |\lambda_i| < n$, so that $|\lambda_i| < 1$ implies (although is not implied by) $|\sum a_{ii}| < n$, which is then a necessary (although not a sufficient) stability condition. This proves condition IX.

We also have $\det\mathbf{A} = \prod_i \lambda_i$ and, taking absolute values, $|\det\mathbf{A}| = |\prod_i \lambda_i|$. Now, $|\prod_i \lambda_i| = \prod_i |\lambda_i|$ and, if $|\lambda_i| < 1$, then $\prod_i |\lambda_i| < 1$. Hence $|\lambda_i| < 1$ implies (although is not implied by) $|\det\mathbf{A}| < 1$, which is then a necessary, although not a sufficient, stability condition.

Finally, condition XI has already been proved above, in conjunction with conditions IV.

9.2.3 Particular solution

In general, when the functions contained in $\mathbf{g}(t)$ belong to one of the standard types (see Chap. 3, Sect. 3.2), the method of undetermined coefficients can be applied with success to find a particular solution. We exemplify by letting $\mathbf{g}(t) = \mathbf{b}$, where $\mathbf{b}$ is a vector of constants. We then try

$$\overline{\mathbf{y}}_t = \overline{\mathbf{y}} \tag{9.79}$$

as a particular solution, where $\overline{\mathbf{y}}$ is vector of undetermined constants. Substitution in Eq. (9.45) yields

$$\overline{\mathbf{y}} = \mathbf{A}\overline{\mathbf{y}} + \mathbf{b},$$

whence

$$\bar{\mathbf{y}} = (\mathbf{I} - \mathbf{A})^{-1}\mathbf{b}. \tag{9.80}$$

This of course requires that $+1$ is not a characteristic root, for in the contrary case we would have $|\mathbf{I} - \mathbf{A}| = 0$. In such a case we try $\bar{\mathbf{y}}_t = \mu_1 + \mu_2 t$, where μ_1, μ_2 are vectors of undetermined constants, and so on.

In economic applications it is often the case that $\mathbf{A} \geq \mathbf{0}, \mathbf{b} \geq \mathbf{0}$, and only a non-negative particular solution is economically meaningful, i.e. we are interested in $\bar{\mathbf{y}} \geq \mathbf{0}$. This turns out to be true if the equilibrium is stable. In fact, the stability conditions for non-negative matrices (see Sect. 9.2.2) also guarantee that $(\mathbf{I} - \mathbf{A})^{-1} \geq \mathbf{0}$. This is a consequence of the theorem (see, for example, Gantmacher, Chap. III, Sect. 3) according to which, when $\mathbf{A} \geq \mathbf{0}$, we have $(\mathbf{I} - \rho\mathbf{A})^{-1} \geq \mathbf{0}$ if, and only if, $\rho > \lambda_{MAX}$. Now, if we let $\rho = 1$, the condition is certainly satisfied in the case of stability ($\lambda_{MAX} < 1$). If, in addition, the matrix is indecomposable, the result can be further strengthened (Gantmacher, Chap. III, Sect. 2) to $(\mathbf{I} - \rho\mathbf{A})^{-1} > \mathbf{0}$.

9.2.3.1 The operational method

We now examine the case in which $\mathbf{g}(t)$ contains arbitrarily given functions of time, or sequences whose functional form we do not know (see Chap. 3, Sect. 3.2.6). Thus we have

$$\mathbf{y}_{t+1} = \mathbf{A}\mathbf{y}_t + \mathbf{x}_t,$$

from which, by shifting the time subscripts backwards by one unit and defining $\mathbf{X}_t \equiv \mathbf{x}_{t-1}$, we obtain

$$\mathbf{y}_t = \mathbf{A}\mathbf{y}_{t-1} + \mathbf{X}_t. \tag{9.81}$$

We can now apply the operational calculus (see Sect. 3.2.6) to find a particular solution. If we introduce the lag operator L, we have

$$(\mathbf{I} - \mathbf{A}L)\mathbf{y}_t = \mathbf{X}_t, \tag{9.82}$$

from which we see that a particular solution is

$$\bar{\mathbf{y}}_t = (\mathbf{I} - \mathbf{A}L)^{-1}\mathbf{X}_t. \tag{9.83}$$

By using the well-known matrix series expansion

$$(\mathbf{I} - \mathbf{A}L)^{-1} = \mathbf{I} + \mathbf{A}L + (\mathbf{A}L)^2 + \ldots = \sum_{i=0}^{\infty} \mathbf{A}^i L^i, \tag{9.84}$$

we finally obtain

$$\bar{\mathbf{y}}_t = \sum_{i=0}^{\infty} \mathbf{A}^i L^i \mathbf{X}_t = \sum_{i=0}^{\infty} \mathbf{A}^i \mathbf{X}_{t-i}. \tag{9.85}$$

Note that $\mathbf{A}^i \to 0$ as $i \to \infty$ only if all the characteristic roots of $\mathbf{A}$ lie within the unit circle, i.e. are all smaller than one in absolute value (see above, Sect. 9.2.1). Hence $\bar{\mathbf{y}}_t$ as given in Eq. (9.85) is a bounded sequence only in such a case. In the contrary case we can use an alternative expansion. We first observe that

$$\mathbf{I} - \mathbf{A}L = -(\mathbf{A}L)\left[\mathbf{I} - (\mathbf{A}L)^{-1}\right], \tag{9.86}$$

from which, defining $\mathbf{H} \equiv \mathbf{A}^{-1}$ and letting $L^{-1} = F$, where F is the forward (or shift) operator,

$$(\mathbf{I} - \mathbf{A}L)^{-1} = -\left[\mathbf{I} - \mathbf{H}F\right]^{-1}\mathbf{H}F. \tag{9.87}$$

The series expansion of $[\mathbf{I} - \mathbf{H}F]^{-1}$ gives

$$[\mathbf{I} - \mathbf{H}F]^{-1} = \mathbf{I} + \mathbf{H}F + (\mathbf{H}F)^2 + \ldots, \tag{9.88}$$

so that, by substituting (9.88) into (9.87) we get

$$\begin{aligned}(\mathbf{I} - \mathbf{A}L)^{-1} &= -[\mathbf{I} + \mathbf{H}F + (\mathbf{H}F)^2 + \ldots]\mathbf{H}F = -[\mathbf{H}F + (\mathbf{H}F)^2 + \ldots] \\ &= -\sum_{i=1}^{\infty} \mathbf{H}^i F^i. \end{aligned} \tag{9.89}$$

Application of Eq. (9.89) yields the particular solution

$$\bar{\mathbf{y}}_t = -\sum_{i=1}^{\infty} \mathbf{H}^i F^i \mathbf{X}_t = -\sum_{i=1}^{\infty} \mathbf{H}^i \mathbf{X}_{t+i}, \tag{9.90}$$

which is a bounded sequence if the characteristic roots of $\mathbf{H} \equiv \mathbf{A}^{-1}$ lie within the circle of convergence. Since the roots of $\mathbf{A}^{-1}$ are reciprocal to the roots of $\mathbf{A}$, we shall apply the *backward* solution (9.85) in the case in which the roots of $\mathbf{A}$ are stable and the *forward* solution (9.90) in the opposite case.

It is also possible that the roots of $\mathbf{A}$ are partly stable and partly unstable. In this case we shall have to apply the backward solution to the terms containing the stable roots and the forward solutions to the terms containing the unstable roots. This can easily be done by using the diagonalization shown above, Sect. 9.2.1, provided that there are no multiple roots.

By diagonalizing the matrix $\mathbf{A}$ we obtain

$$\mathbf{y}_t = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}\mathbf{y}_{t-1} + \mathbf{X}_t, \tag{9.91}$$

where $\mathbf{V}$ is the modal matrix. We now premultiply through by $\mathbf{V}^{-1}$ and obtain

$$\mathbf{V}^{-1}\mathbf{y}_t = (\mathbf{V}^{-1}\mathbf{V})\mathbf{\Lambda}\mathbf{V}^{-1}\mathbf{y}_{t-1} + \mathbf{V}^{-1}\mathbf{X}_t = \mathbf{\Lambda}\mathbf{V}^{-1}\mathbf{y}_{t-1} + \mathbf{V}^{-1}\mathbf{X}_t. \tag{9.92}$$

Let us now define the new variables

$$\mathbf{z} \equiv \mathbf{V}^{-1}\mathbf{y}, \quad \mathbf{v} \equiv \mathbf{V}^{-1}\mathbf{X}. \tag{9.93}$$

In this way we obtain the diagonal system

$$\mathbf{z}_t = \mathbf{\Lambda}\mathbf{z}_{t-1} + \mathbf{v}_t, \tag{9.94}$$

that can be easily solved, since it consists of a set of n independent first-order difference equations of the type

$$z_{i,t} = \lambda_i z_{i,t-1} + v_{i,t}. \tag{9.95}$$

The solution will be

$$z_{i,t} = c_i\lambda_i^t + \overline{z}_{i,t}, \tag{9.96}$$

where c_i is an arbitrary constant, and $\overline{z}_{i,t}$ is a particular solution. Possible complex roots will be transformed into a trigonometric oscillation as shown in Chap. 5, Sect. 5.1.3.

To find each equation's particular solution $\overline{z}_{i,t}$ we shall apply the operational method treated in Chap. 3, Sect. 3.26, in relation to first-order equations, and *use the backward (forward) solution for the equations where the stable (unstable) roots appear.*

Once we have found the solution of the system in terms of the transformed variables, we can go back to the original variables by applying the inverse transformation

$$\mathbf{y} = \mathbf{V}\mathbf{z}. \tag{9.97}$$

This procedure can of course also be applied to the case in which all the characteristic roots are of the same type, a case in which, however, the procedure given at the beginning is simpler.

The operational method of finding a particular solution is quite general and could be applied to the standard functional forms of $\mathbf{g}(t)$, but in these cases the method of undetermined coefficients is simpler. To show this, consider, for example, the case $\mathbf{g}(t) = \mathbf{b}$. Then Eq. (9.85) gives

$$\overline{\mathbf{y}}_t = \sum_{i=0}^{\infty}\mathbf{A}^i\mathbf{b} = \left(\sum_{i=0}^{\infty}\mathbf{A}^i\right)\mathbf{b} = (\mathbf{I}-\mathbf{A})^{-1}\mathbf{b}, \tag{9.98}$$

since $\sum_{i=0}^{\infty}\mathbf{A}^i$ converges to $(\mathbf{I}-\mathbf{A})^{-1}$ when the characteristic roots of $\mathbf{A}$ are smaller than unity in absolute value. In the opposite case we can use the forward solution and get

$$\overline{\mathbf{y}}_t = -\sum_{i=1}^{\infty}\mathbf{H}^i\mathbf{b} = -\left(\sum_{i=0}^{\infty}\mathbf{H}^i - \mathbf{I}\right)\mathbf{b} = \left[-(\mathbf{I}-\mathbf{H})^{-1} + \mathbf{I}\right]\mathbf{b}, \tag{9.99}$$

where we have first added and subtracted $\mathbf{H}^0 = \mathbf{I}$, and then used the fact that $\sum_{i=0}^{\infty}\mathbf{H}^i = (\mathbf{I}-\mathbf{H})^{-1}$. Now, $\mathbf{I} - (\mathbf{I}-\mathbf{H})^{-1} = (\mathbf{I}-\mathbf{H})^{-1}(\mathbf{I}-\mathbf{H}) - (\mathbf{I}-\mathbf{H})^{-1} = -(\mathbf{I}-\mathbf{H})^{-1}\mathbf{H}$, and

$$\begin{aligned} -(\mathbf{I}-\mathbf{H})^{-1}\mathbf{H} &= (-\mathbf{I}+\mathbf{H})^{-1}\mathbf{H} = (\mathbf{H}^{-1}(\mathbf{H}-\mathbf{I}))^{-1} = \left(\mathbf{I}-\mathbf{H}^{-1}\right)^{-1} \\ &= (\mathbf{I}-\mathbf{A})^{-1}. \end{aligned} \tag{9.100}$$

Therefore, by substituting (9.100) into (9.99), the forward solution gives

$$\overline{\mathbf{y}}_t = -\sum_{i=1}^{\infty}\mathbf{H}^i\mathbf{b} = (\mathbf{I}-\mathbf{A})^{-1}\mathbf{b}, \tag{9.101}$$

which is the same result found with the backward solution and, of course, with Eq. (9.80).

9.2.4 Determination of the arbitrary constants

The arbitrary constants that appear in the general solution can be determined as usual by means of a number of additional conditions equal to the number of arbitrary constants. These conditions give the information that, for a given value of t (usually for $t = 0$, whence the name *initial conditions*), the values of the various functions are known values. Substituting in the general solution of the system under consideration, we obtain a system of linear equations which can be solved for the values of the arbitrary constants.

Consider, for example, the case in which the characteristic roots are all distinct. Then the equations for the determination of the arbitrary constants turn out to be

$$\sum_{j=1}^{n} A_j\alpha_i^{(j)} + \overline{y}_i(0) = y_i(0), \qquad i = 1, 2, ..., n, \tag{9.102}$$

where $\overline{y}_i(0) = 0$ if the system is homogeneous (see above Eq. (9.56)). It is a well-known theorem in matrix algebra that characteristic vectors associated with distinct characteristic roots are linearly independent. Therefore, the matrix $[\alpha_i^{(j)}]$ is non-singular and system (9.102) can be solved. Some complications may arise when there are multiple roots, but it suffices to observe that we are starting from a fundamental set of solutions, so that, according to general principles, the relevant matrices will always be non singular.

Note, finally, that system (9.102) can also be written as

$$\sum_{j=1}^{n} A_j\alpha_i^{(j)} = y_i(0) - \overline{y}_i(0), \qquad i = 1, 2, ..., n, \tag{9.103}$$

where the r.h.s. can be interpreted as giving the *initial deviations* of the system from its equilibrium solution $\overline{y}_i(0)$.

9.3 General systems

The systems that we have so far examined have the peculiarity that in each equation only one unknown function appears at two different points of time, which moreover are adjacent (t and $t+1$, or t and $t-1$). This is why they are called first-order and normal, as stated at the beginning of this chapter. But, in general, in each equation *each* unknown function might appear at different points of time, not necessarily adjacent.

For didactic purposes we distinguish between first-order systems not in normal form and higher-order systems.

9.3.1 First-order systems not in normal form

In this case, the different points of time are still two and adjacent, but in each equation two (or more) functions appear evaluated at time $t+1$:

$$\mathbf{A}\mathbf{y}_{t+1} + \mathbf{B}\mathbf{y}_t = \mathbf{g}(t), \tag{9.104}$$

where $\mathbf{A}, \mathbf{B}$ are $n \times n$ square matrices, $\mathbf{y} = [y_1, y_2, ..., y_n]$ is a $n \times 1$ vector of unknown functions of time, and $\mathbf{g}(t)$ is a vector of known functions. The corresponding homogeneous system is

$$\mathbf{A}\mathbf{y}_{t+1} + \mathbf{B}\mathbf{y}_t = \mathbf{0}. \tag{9.105}$$

To solve system (9.105) we can use the same direct method explained above in relation to first-order systems in normal form, namely we start by trying as a solution $y_{it} = \alpha_i \lambda^t$:

$$\mathbf{y}_t = [\alpha_1\lambda^t, \alpha_2\lambda^t, ..., \alpha_n\lambda^t] = \boldsymbol{\alpha}\lambda^t, \tag{9.106}$$

where the elements of the vector $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, ..., \alpha_n]$ are not all zero. Substituting in system (9.105) we have

$$\mathbf{A}\boldsymbol{\alpha}\lambda^{t+1} + \mathbf{B}\boldsymbol{\alpha}\lambda^t = \lambda^t(\mathbf{A}\lambda + \mathbf{B})\boldsymbol{\alpha} = \mathbf{0}, \tag{9.107}$$

which will be identically satisfied if, and only if

$$(\mathbf{A}\lambda + \mathbf{B})\boldsymbol{\alpha} = \mathbf{0}. \tag{9.108}$$

System (9.108), in turn, admits of a non-trivial solution for $\boldsymbol{\alpha}$ if, and only if, its determinant is zero, namely

$$|\mathbf{A}\lambda + \mathbf{B}| = 0. \tag{9.109}$$

The determinantal equation (9.109) and its expanded form are called the characteristic equation of system (9.105). The expansion of the determinant gives rise to a n-th order polynomial equation in the unknown λ. From this

point on, the procedure is the same as that explained in detail in relation to first-order systems in normal form. If, for example, the characteristic equation has n distinct real roots, to each root λ_j we can associate a vector $\alpha^{(j)} = [\alpha_1^{(j)}, \alpha_2^{(j)}, ..., \alpha_n^{(j)}]$ derived from the solution of set (9.108) when $\lambda = \lambda_j$. Then the solution of system (9.105) will have the usual form

$$\begin{aligned} y_{1t} &= A_1\alpha_1^{(1)}\lambda_1^t + A_2\alpha_1^{(2)}\lambda_2^t + ... + A_n\alpha_1^{(n)}\lambda_n^t, \\ y_{2t} &= A_1\alpha_2^{(1)}\lambda_1^t + A_2\alpha_2^{(2)}\lambda_2^t + ... + A_n\alpha_2^{(n)}\lambda_n^t, \\ &\ldots\ldots \\ y_{nt} &= A_1\alpha_n^{(1)}\lambda_1^t + A_2\alpha_n^{(2)}\lambda_2^t + ... + A_n\alpha_n^{(n)}\lambda_n^t, \end{aligned} \tag{9.110}$$

where $A_1, A_2, ..., A_n$ are arbitrary constants to be determined by means of n additional conditions in the usual way.

Alternatively, we can *reduce system* (9.105) *to an equivalent first-order system in normal form* by simple manipulations. If the matrix $\mathbf{A}$ is non-singular, we immediately obtain the normal form

$$\mathbf{y}_{t+1} = \mathbf{C}\mathbf{y}_t, \qquad \mathbf{C} = -\mathbf{A}^{-1}\mathbf{B}. \tag{9.111}$$

If $\mathbf{A}$ is singular, we must first reduce the number of equations and unknowns to r (the rank of $\mathbf{A}$) by suitable substitutions, and then express the remaining $y_{it+1}, i = 1, 2, ..., r$; $(r < n)$, in terms of the y_{it} as before, obtaining the normal form. Note that when $\mathbf{A}$ is singular the order of the system is no longer n, but r.

The reason why the transformation of a general first-order system into the normal form (9.111) might be preferable to the direct solution is of course that by so doing we can apply the stability conditions given above (see Sect. 9.2.2).

Finally, a particular solution of the non-homogeneous system (9.104) can be found by the method of undetermined coefficients in the usual way, and the standard procedure can be applied to the determination of the arbitrary constants.

9.3.2 Higher-order systems

As we said above, in each equation *each* unknown function might appear at different points of time, not necessarily adjacent. Let us begin with a simple example before going on to the general case.

9.3.2.1 An example

Let us consider the non-homogeneous system

$$\begin{aligned} a_3y_{t+3} + a_2y_{t+2} + \quad b_1z_{t+1} + b_0z_t &= g_1(t), \\ c_1y_{t+1} + c_0y_t + \quad d_2z_{t+2} \quad &= g_2(t), \end{aligned} \tag{9.112}$$

and its homogeneous counterpart

$$\begin{aligned} a_3 y_{t+3} + a_2 y_{t+2} \quad + b_1 z_{t+1} + b_0 z_t &= 0, \\ c_1 y_{t+1} + c_0 y_t \quad + d_2 z_{t+2} &= 0. \end{aligned} \tag{9.113}$$

To solve system (9.113) we can proceed as with first-order systems in normal form, namely either reduce it to a single difference equation or apply the direct method, trying as a solution $y_t = \alpha_1 \lambda^t, z_t = \alpha_2 \lambda^t$, where α_1, α_2 are not both zero. We then have

$$\begin{aligned} a_3 \alpha_1 \lambda^{t+3} + a_2 \alpha_1 \lambda^{t+2} + \; b_1 \alpha_2 \lambda^{t+1} + b_0 \alpha_2 \lambda^t &= 0, \\ c_1 \alpha_1 \lambda^{t+1} + c_0 \alpha_1 \lambda^t + \; d_2 \alpha_2 \lambda^{t+2} &= 0, \end{aligned}$$

from which

$$\begin{aligned} \lambda^t[(a_3\lambda^3 + a_2\lambda^2)\alpha_1 + \; (b_1\lambda + b_0)\alpha_2] &= 0, \\ \lambda^t[(c_1\lambda + c_0)\alpha_1 + \; d_2\lambda^2\alpha_2] &= 0. \end{aligned} \tag{9.114}$$

If $\alpha_1\lambda^t, \alpha_2\lambda^t$ is a solution, system (9.114) must be satisfied for any t and this is possible—apart from the trivial case $\lambda = 0$—if, and only if

$$\begin{aligned} (a_3\lambda^3 + a_2\lambda^2)\alpha_1 + \; (b_1\lambda + b_0)\alpha_2 &= 0, \\ (c_1\lambda + c_0)\alpha_1 + \; d_2\lambda^2\alpha_2 &= 0. \end{aligned} \tag{9.115}$$

System (9.115) will yield a non-trivial solution for α_1, α_2 if, and only if, its determinant is zero, namely

$$\begin{vmatrix} a_3\lambda^3 + a_2\lambda^2 & b_1\lambda + b_0 \\ c_1\lambda + c_0 & d_2\lambda^2 \end{vmatrix} = 0. \tag{9.116}$$

The polynomial form of the determinantal equation (9.116) is

$$a_3 d_2 \lambda^5 + a_2 d_2 \lambda^4 - c_1 b_1 \lambda^2 - (b_1 c_0 + b_0 c_1)\lambda - b_0 c_0 = 0. \tag{9.117}$$

From this point on, the procedure is the same as that explained above for first-order systems in normal form. If, for example, the characteristic equation (9.117) yields five real and distinct roots, to each of them we can associate a couple of values $\alpha_1^{(j)}, \alpha_2^{(j)}$, and the solution will have the form

$$\begin{aligned} y_t &= A_1\alpha_1^{(1)}\lambda_1^t + A_1\alpha_1^{(2)}\lambda_2^t + A_1\alpha_1^{(3)}\lambda_3^t + A_1\alpha_1^{(4)}\lambda_4^t + A_1\alpha_1^{(5)}\lambda_5^t, \\ z_t &= A_1\alpha_2^{(1)}\lambda_1^t + A_1\alpha_2^{(2)}\lambda_2^t + A_1\alpha_2^{(3)}\lambda_3^t + A_1\alpha_2^{(4)}\lambda_4^t + A_1\alpha_2^{(5)}\lambda_5^t, \end{aligned} \tag{9.118}$$

where $A_1, A_2, ..., A_5$ are arbitrary constants.

The stability of the system can be examined by applying the stability conditions explained in Sect. 7.3 to Eq. (9.117); see, however, below for a more general treatment of stability.

In the case of a non-homogeneous system, a particular solution can be found by the method of undetermined coefficients. The arbitrary constants, finally, can be determined given a sufficient number of additional conditions, for example that y_0, y_1, y_2 and z_0, z_1, are all known values.

9.3.2.2 The general case

The determinantal form of the characteristic equation of a general higher order system can be easily found if we use the forward (or shift) operator F already introduced in Sect. (9.2.3.1), namely the operator such that

$$y_{t+1} = F y_t, \qquad F^n y_t = y_{t+n}, \tag{9.119}$$

and the polynomial operator

$$P(F) = a_0 F^n + a_1 F^{n-1} + \ldots + a_{n-1} F + a_n. \tag{9.120}$$

Note that F^k simply means that the operator F has been applied k times in succession.

Then a general homogeneous difference equation system can be written as follows:

$$\begin{aligned} P_{11}(F)y_{1t} + \; P_{12}(F)y_{2t} \; + \ldots + \; P_{1n}(F)y_{nt} &= 0, \\ P_{21}(F)y_{1t} + \; P_{22}(F)y_{2t} \; + \ldots + \; P_{2n}(F)y_{2t} &= 0, \\ \ldots \quad \ldots \qquad \ldots \qquad \ldots \quad & \ldots \\ P_{n1}(F)y_{1t} + \; P_{n2}(F)y_{2t} \; + \ldots + \; P_{nn}(F)y_{nt} &= 0, \end{aligned} \tag{9.121}$$

where the $P_{ij}(F)$ are polynomial operators of the appropriate orders. For example, the homogeneous part of system (9.112) can be written as

$$\begin{aligned} (a_3F^3 + a_2F^2)y_t \; &+ \; (b_1F + b_0)z_t = 0, \\ (c_1F + c_0)y_t \; &+ \; d_2F^2 z_t = 0. \end{aligned} \tag{9.122}$$

Now, if we apply the usual direct method of solution to system (9.121), it can easily be seen that the determinantal form of the characteristic equation of this system can be written simply by considering the determinant whose elements are the polynomials obtained by replacing F with λ in the operators $P_{ij}(F)$. That is to say,

$$\begin{vmatrix} P_{11}(\lambda) & P_{12}(\lambda) & \ldots & P_{1n}(\lambda) \\ P_{21}(\lambda) & P_{22}(\lambda) & \ldots & P_{2n}(\lambda) \\ \ldots & \ldots & \ldots & \ldots \\ P_{n1}(\lambda) & P_{n2}(\lambda) & \ldots & P_{nn}(\lambda) \end{vmatrix} = 0 \tag{9.123}$$

is the characteristic equation of system (9.121).

9.3.2.3 Transformation of a higher-order system into a first-order system in normal form

A general higher-order homogeneous system can always be written as

$$\mathbf{M}_0\mathbf{y}_{t+k} + \mathbf{M}_1\mathbf{y}_{t+k-1} + ... + \mathbf{M}_k\mathbf{y}_t = \mathbf{0}, \tag{9.124}$$

where $\mathbf{y}_t$ is the n-dimensional vector $[y_{1t}, y_{2t}, ..., y_{nt}]$ of unknown functions of time, and $\mathbf{M}_i$, $i = 0, 1, ..., k$ are $n \times n$ matrices with zeros in the appropriate places. Shifting the time subscripts we can also write this system as

$$\mathbf{M}_0\mathbf{y}_t + \mathbf{M}_1\mathbf{y}_{t-1} + ... + \mathbf{M}_k\mathbf{y}_{t-k} = \mathbf{0}. \tag{9.125}$$

If we consider for example system (9.113) and shift all the time subscripts appropriately backwards (by 3 those in the first equation, and by 2 those in the second one) we get

$$\begin{aligned} a_3y_t + a_2y_{t-1} + b_1z_{t-2} + b_0z_{t-3} &= 0, \\ c_1y_{t-1} + c_0y_{t-2} + d_2z_t &= 0, \end{aligned}$$

namely

$$\begin{bmatrix} a_3 & 0 \\ 0 & d_2 \end{bmatrix}\begin{bmatrix} y_t \\ z_t \end{bmatrix} + \begin{bmatrix} a_2 & 0 \\ c_1 & 0 \end{bmatrix}\begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} 0 & b_1 \\ c_0 & 0 \end{bmatrix}\begin{bmatrix} y_{t-2} \\ z_{t-2} \end{bmatrix} + \begin{bmatrix} 0 & b_0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} y_{t-3} \\ z_{t-3} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \tag{9.126}$$

which is in the form (9.125).

We can assume that $\mathbf{M}_0$ is non singular—in the opposite case, we first reduce the system to the rank of $\mathbf{M}_0$, as explained in Sect. (9.3.1)—so that we have

$$\mathbf{y}_t = \mathbf{N}_1\mathbf{y}_{t-1} + \mathbf{N}_2\mathbf{y}_{t-2} + ... + \mathbf{N}_k\mathbf{y}_{t-k}, \tag{9.127}$$

where $\mathbf{N}_i \equiv -\mathbf{M}_0^{-1}\mathbf{M}_i, i = 1, 2, ..., k$. Eq. (9.127) is called the *distributed lag* form of the system, and can be brought into an equivalent first-order system in normal form by standard transformations (see, e.g. Frazer, Duncan, Collar, 1938, Sect. 5.5; Collar and Simpson, 1987, Sect. 6.2). These amount to defining the variables

$$\mathbf{x}_{1t} = \mathbf{y}_{t-1}, ..., \mathbf{x}_{rt} = \mathbf{x}_{r-1,t-1} \text{ for } r = 2, ..., k-1 \tag{9.128}$$

and substituting them into (9.127), which then becomes

$$\mathbf{y}_t = \mathbf{N}_1\mathbf{y}_{t-1} + \mathbf{N}_2\mathbf{x}_{1,t-1} + \mathbf{N}_3\mathbf{x}_{2,t-1} + ... + \mathbf{N}_k\mathbf{x}_{k-1,t-1}. \tag{9.129}$$

System (9.129) can be written as a first-order system in normal form, namely

$$\mathbf{Y}_t = \mathbf{N}\mathbf{Y}_{t-1}, \tag{9.130}$$

where

$$\mathbf{Y}_t \equiv \begin{bmatrix} \mathbf{y}_t \\ \mathbf{x}_{1t} \\ \vdots \\ \mathbf{x}_{k-1,t} \end{bmatrix} \tag{9.131}$$

and

$$\mathbf{N} \equiv \begin{bmatrix} \mathbf{N}_1 & \mathbf{N}_2 & ... & \mathbf{N}_{k-2} & \mathbf{N}_{k-1} & \mathbf{N}_k \\ \mathbf{I} & \mathbf{0} & ... & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & ... & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ ... & ... & ... & ... & ... & ... \\ \mathbf{0} & \mathbf{0} & ... & \mathbf{0} & \mathbf{I} & \mathbf{0} \end{bmatrix}. \tag{9.132}$$

For example, take system (9.126) and first put it into the form (9.127). It is easy to check that

$$\mathbf{N}_1 = \begin{bmatrix} -a_2/a_3 & 0 \\ -c_1/d_2 & 0 \end{bmatrix}, \mathbf{N}_2 = \begin{bmatrix} 0 & -b_1/a_3 \\ -c_0/d_2 & 0 \end{bmatrix}, \mathbf{N}_3 = \begin{bmatrix} 0 & -b_0/a_3 \\ 0 & 0 \end{bmatrix}. \tag{9.133}$$

Hence the matrix $\mathbf{N}$ turns out to be

$$\mathbf{N} = \left[\begin{array}{cc|cc|cc} -a_2/a_3 & 0 & 0 & -b_1/a_3 & 0 & -b_0/a_3 \\ -c_1/d_2 & 0 & -c_0/d_2 & 0 & 0 & 0 \\ \hline 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{array}\right]. \tag{9.134}$$

The reader can check as an exercise that the characteristic equation of matrix (9.134), $|\mathbf{N}-\lambda\mathbf{I}| = 0$, gives rise to the polynomial equation

$$\lambda\left(\lambda^5 + \frac{a_2}{a_3}\lambda^4 - \frac{c_1b_1}{a_3d_2}\lambda^2 - \frac{b_1c_0 + b_0c_1}{a_3d_2}\lambda - \frac{b_0c_0}{a_3d_2}\right) = 0$$

which, excluding the trivial root $\lambda = 0$, coincides with Eq. (9.117).

9.3.2.4 Stability conditions for higher-order systems

The transformation of a higher-order system into an equivalent first-order system in normal form might seem a roundabout and complicated way of solving it. This is certainly true if one is interested in finding the explicit form of the solution. But it is no longer true if one is interested in checking the stability of the system without solving it. In such a case the transformation explained in the previous section gives us the possibility of applying to the matrix **N** –defined in (9.132)– the stability conditions valid for first-order systems in normal form (see above, Sect. (9.2.2).

Among these conditions the easiest to apply are undoubtedly the row-sum conditions on the (absolute values of the) elements of the matrix **N** (see, for example, Bear, 1966; Conlisk, 1973; Fujimoto and Indelli, 1986). It is, in fact, easy to see that these row sums are equal to 1 for the last $(k-1)n$ rows of **N**, while for the first n rows they equal the sum of the row sums of the (absolute values of the) elements of the matrices $\mathbf{N}_1, \mathbf{N}_2, ..., \mathbf{N}_k$.

9.4 Exercises

9.4.1 Example

Consider the system

$y_{t+1} = 0.70y_t + 0.30z_t,$
$z_{t+1} = 0.10y_t + 0.75z_t + 15,$
$y_0 = 225, z_0 = 25.$

We observe that this system is a first-order system in normal form, whose matrix

$$\mathbf{A} \equiv \begin{bmatrix} 0.70 & 0.30 \\ 0.10 & 0.75 \end{bmatrix}$$

is a positive matrix. Since the row sums are 1 and 0.85, we know by conditions (9.69) that both characteristic roots will be stable. The characteristic equation

$$\begin{vmatrix} 0.70-\lambda & 0.30 \\ 0.10 & 0.75-\lambda \end{vmatrix} = \lambda^2 - 1.45 + 0.495 = 0$$

has, in fact, the roots 0.90 and 0.55. Using Eqs. (9.21)-(9.22) we can write the general solution of the homogeneous part of our system as

$$\begin{aligned} y_t &= A_1(0.90)^t + A_2(0.55)^t, \\ z_t &= A_1\tfrac{2}{3}(0.90)^t + A_2(-0.5)(0.55)^t. \end{aligned}$$

By letting $y_{t+1} = y_t = \overline{y}, z_{t+1} = z_t = \overline{z}$, we obtain

$$\begin{aligned} -0.30\overline{y} + 0.30\overline{z} &= 0, \\ -0.10\overline{y} + 0.25\overline{z} &= 15, \end{aligned}$$

from which we get the particular solution of the non-homogeneous system, $\overline{y} = \overline{z} = 100$.

Given the initial conditions, we have the system

$$\begin{aligned} A_1 + A_2 + 100 &= 225, \\ \tfrac{2}{3}A_1 - 0.5A_2 &= 25, \end{aligned}$$

from which $A_1 = 75, A_2 = 50$.

9.4.2 Other exercises

A) Apply the stability conditions explained in Sect. (9.2.2) to check the stability of the following system

$$\mathbf{y}_{t+1} = \mathbf{A}\mathbf{y}_t, \qquad \mathbf{A} = \begin{bmatrix} 1/2 & 0 & 1/2 \\ 1 & 1/2 & 1 \\ 1/2 & 0 & 1/4 \end{bmatrix},$$

without solving it.

B) Solve the following systems:

1) $-1.5y_{1t+1} + 7.5y_{1t} - 9y_{2t} = 12,\ 3y_{1t} - 3y_{2t+1} = 3.$
(Hint: to put it in normal form, divide the first equation by 1.5 etc.)
2) $y_{1t+1} = 1.5y_{1t} - 0.5y_{2t},\ y_{2t+1} = 0.11y_{1t} + 0.85y_{2t}.$
3) $y_{1t+1} - 2y_{2t} + y_{3t+1} = 2,\ y_{1t} - 3y_{2t+1} + y_{2t} = 1,\ 2y_{2t+1} - 4y_{3t+1} + 2y_{3t} = 4.$
4) $y_{1t+1} - y_{1t} + 2y_{2t+1} = 0,\ y_{2t+1} - y_{2t} - 2y_{1t} = a^t.$

C) Consider the nth order difference equation

$$y_{t+n} + a_1 y_{t+n-1} + a_2 y_{t+n-2} + ... + a_{n-2}y_{t-2} + a_{n-1}y_{t+1} + a_n y_t = g(t)\cdot$$

and introduce the new variables

$$\begin{aligned} y_{t+1} &= y_{1,t}, \\ y_{1,t+1} &= y_{2,t}, \\ \vdots\quad & \quad \vdots \\ y_{n-2,t+1} &= y_{n-1,t}. \end{aligned}$$

Then the equation can be written as

$$y_{n-1,t+1} = -a_1 y_{n-1,t} - a_2 y_{n-2,t} - ... - a_{n-2}y_{2,t} - a_{n-1}y_{1,t} - a_n y_t + g(t).$$

Show that it is equivalent to the first-order system in normal form

$$\begin{bmatrix} y_{t+1} \\ y_{1,t+1} \\ y_{2,t+1} \\ \cdots \\ y_{n-2,t+1} \\ y_{n-1,t+1} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ a_n & a_{n-1} & a_{n-2} & \cdots & a_2 & a_1 \end{bmatrix} \begin{bmatrix} y_t \\ y_{1,t} \\ y_{2,t} \\ \cdots \\ y_{n-2,t} \\ y_{n-1,t} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ \cdots \\ 0 \\ g(t) \end{bmatrix}.$$

9.5 References

Baumol, W.J., 1970, *Economic Dynamics,* Chaps. 15 and 16.
Bear, D.V.T., 1966, Distributed Lags and Economic Theory.
Brauer, A., 1964, On the Characteristic Roots of Non-Negative Matrices.
Collar, A.R. and A. Simpson, 1987, *Matrices and Engineering Dynamics,* Chap. 6, Sect. 6.2.
Conlisk, J., 1973, Quick Stability Checks and Matrix Norms.
Fisher, F.M., 1962, An Alternate Proof and Extension of Solow's Theorem on Nonnegative Square Matrices.
Fisher, F.M., 1965, Choice of Units, Column Sums, and Stability in Linear Dynamic Systems with Nonnegative Square Matrices.
Frazer, R.A., W.J. Duncan and A.R. Collar, 1939, *Elementary Matrices and Some Applications to Dynamics and Differential Equations,* Chap. 5, Sect. 5.5.
Fujimoto, T. and P. Indelli, 1986, A Note on Quick Stability Checks for Linear Difference Equations.
Gantmacher, F.R., 1959, *Applications of the Theory of Matrices,* Chap. III, Sects. 1, 2, 3.
Goldberg, S. 1959, *Introduction to Difference Equations,* Chap. 4, Sect. 4.5.
Hadar, J., 1965, A Note on Dominant Diagonals in Stability Analysis.
Kenkel, J.L., 1974, *Dynamic Linear Economic Models,* Chap. 8.
Luenberger, D.G., 1979, *Introduction to Dynamic Systems,* Chap. 6.
McKenzie, L., 1960, Matrices with Dominant Diagonals and Economic Theory, p. 49 and p. 58.
Murata, Y., 1977, *Mathematics for Stability and Optimization of Economic Systems,* Chap. 3, Sects. 3.3, 3.5; Chap. 4, Sect. 4.1-4.3.
Samuelson, P.A., 1947, *Foundations of Economic Analysis,* Appendix B, Sects. 30-35.
Schwartz, J.T., 1961, *Lectures on the Mathematical Method in Analytical Economics,* Lecture 1, Sect. 2.
Solow, R.M., 1952, On the Structure of Linear Models.
Takayama, A., 1985, *Mathematical Economics*, Chap. 4, Sects. A-C.
Tarr, D.G., 1976, Distributed Lags, Morishima Matrices, and the Stability of Economic Models.

Taussky, O., 1964, On the Variation of the Characteristic Roots of a Finite Matrix under Various Changes of its Elements.
Varian, H. R., 1992, *Microeconomic Analysis.*

Chapter 10

Simultaneous Difference Systems in Economic Models

Actually all economic applications, even the simplest ones, are simultaneous systems. The fact that a lot of them are usually examined as giving rise to a single equation is because the reduction of the system to a single equation suggests itself quite naturally (i.e. only direct substitutions are required, without any previous manipulations) and is actually easier than simultaneous methods. For example, the homogeneous part of the multiplier model

$$\begin{aligned} &C_t - bY_{t-1} = 0, \\ &I_t = I_0, \\ &C_t + I_t - Y_t = 0, \end{aligned}$$

has the characteristic equation

$$\begin{vmatrix} \lambda & 0 & -b \\ 0 & 1 & 0 \\ 1 & 1 & -1 \end{vmatrix} = -\lambda + b = 0,$$

whence $\lambda = b$, and so on. However, direct substitution from the first two equations into the third is an obvious and simpler alternative. The same considerations apply to the models treated in the previous chapters. On the other hand there are some economic models, in which when all *direct* substitution have been performed, two or more simultaneous equations still remain, to which simultaneous method can be profitably applied.

10.1 Cournot oligopoly

The classic oligopoly model by Cournot (1838) is still present in all microeconomics textbooks, where it is shown that it can be interpreted in terms of modern game theory (see, for example, Varian, 1992, Chap. 16). It is also possible to give a dynamic interpretation to the model, following what

Cournot had in mind when he wrote on the stability of his model. This amounts to assuming a learning process in which each firm observes the other firms' choice of output, thereby refining its own belief on their behaviour. More precisely, in any period t each firm i observes the other firms' outputs and assumes that these quantities will remain unchanged in period $t+1$. Hence, given the market demand curve, firm i will choose its profit-maximising output consistent with this belief.

The surprising result—first found by Theocharis (1960)—is that, when we formalise this dynamic behaviour, the path of outputs is no longer stable when the number of firms is greater than two.

Let us consider a market with n oligopolistic firms producing a homogeneous output, and with a linear market demand curve,

$$p_t = a - b\sum_{i=1}^{n} x_{it}, \tag{10.1}$$

where $a > 0, b > 0$ and x_{it} is the actual output of firm i at time t. We also assume linear cost curves C_i for each firm, so that marginal cost c_i is constant.

Given the behaviour assumption made above, each firm has an ex ante market price based on the belief that the other firms' outputs will remain unchanged, namely

$$p^i_{t+1} = a - b\left(x_{it+1} + \sum_{j\neq i}^{n} x_{jt}\right), \tag{10.2}$$

on the basis of which firm i determines x_{it+1} so as to maximise its expected profit $\pi^i_{t+1} = p^i_{t+1}x_{it+1} - C_i = ax_{it+1} - bx^2_{it+1} - bx_{it+1}\sum^n_{j\neq i} x_{jt} - C_i$. The first-order conditions for an interior maximum are $\partial\pi^i_{t+1}/\partial x_{it+1} = 0$, namely

$$\begin{gathered}(a - c_i) - 2bx_{it+1} - b\sum_{j\neq i}^{n} x_{jt} = 0,\\ \text{or}\\ x_{it+1} = -\frac{1}{2}\sum_{j\neq i}^{n} x_{jt} + \frac{a - c_i}{2b}.\end{gathered} \tag{10.3}$$

The second order conditions $\partial^2\pi^i_{t+1}/\partial x^2_{it+1} < 0$ are certainly satisfied since we have assumed $b > 0$. Hence Eqs. (10.3) yield the required solution. Let us begin with the case of duopoly, in which the first-order conditions give rise to the non-homogeneous difference system

$$\begin{aligned} x_{1t+1} &= -\frac{1}{2}x_{2t} + \frac{a - c_1}{2b}, \\ x_{2t+1} &= -\frac{1}{2}x_{1t} + \frac{a - c_2}{2b}. \end{aligned} \tag{10.4}$$

The characteristic equation of the homogeneous part of system (10.4) is

$$\begin{vmatrix} -\lambda & -1/2 \\ -1/2 & -\lambda \end{vmatrix} = \lambda^2 - \frac{1}{4} = 0,$$

from which $\lambda_1, \lambda_2 = \pm\frac{1}{2}$. Thus the movement will be convergent toward the equilibrium solution, which is obtained by letting $x_{1t+1} = x_{1t} = \overline{x}_1, x_{2t+1} = x_{2t} = \overline{x}_2$ in system (10.4). By simple calculations we have $\overline{x}_1 = (a - 2c_1 + c_2)/3b, \overline{x}_2 = (a - 2c_2 + c_1)/3b$.

This proves Cournot's intuition that his equilibrium was stable, i.e. "if either of the producers, misled as to his true interest, leaves it temporarily, he will be brought back to it by a series of reactions, constantly declining in amplitude" (p. 81 of the English translation). But it also proves the danger of believing that what holds for two-dimensional systems automatically holds for n-dimensional systems. Let us consider the matrix of system (10.3), that is

$$\mathbf{A} = \begin{bmatrix} 0 & -\frac{1}{2} & \dots & -\frac{1}{2} \\ -\frac{1}{2} & 0 & -\frac{1}{2} & -\frac{1}{2} \\ \dots & \dots & \dots & \dots \\ -\frac{1}{2} & -\frac{1}{2} & \dots & 0 \end{bmatrix}, \tag{10.5}$$

and its characteristic equation

$$|\mathbf{A} - \lambda\mathbf{I}| = \begin{vmatrix} -\lambda & -\frac{1}{2} & \dots & -\frac{1}{2} \\ -\frac{1}{2} & -\lambda & -\frac{1}{2} & -\frac{1}{2} \\ \dots & \dots & \dots & \dots \\ -\frac{1}{2} & -\frac{1}{2} & \dots & -\lambda \end{vmatrix} = 0. \tag{10.6}$$

It can easily be checked that $\lambda = -(n-1)/2$ is a characteristic root. In fact, consider the characteristic determinant, and add for example the first $(n-1)$ rows to the last row (we know from elementary rules on determinants that this operation leaves the value of the determinant unchanged). All the elements of the last row turn out to be $-\lambda - (n-1)/2$. Now if we let $\lambda = -(n-1)/2$, we obtain a determinant with a row of zeros, which implies that the determinant is zero. Hence $\lambda = -(n-1)/2$ is a characteristic root.

It follows that for $n = 3$ the system will have a root equal to -1 (constant-amplitude improper oscillations), while for $n > 3$ these oscillations will be explosive. Hence the system is stable only for $n = 2$.

These results are of course subject to modification with the introduction of adjustment lags, different assumptions on the formation of expectations, etcetera. For a survey of such studies see Okuguchi (1976); see also Okuguchi and Szidarovszky (1990).

10.2 Multiplier effects in an open economy

In section 3.2, we examined, among others, the dynamics of the foreign trade multiplier in the case in which exports are assumed to be exogenous. Actually exports may depend indirectly on national income. Suppose, for example, that in a country (call it country 1) a variation in income occurs. Then also imports, which depend on income, change and this change means a change in the exports of the rest of the world (that for simplicity sake, we consider as a single country—call it country 2) to country 1. The change in exports in country 2 causes a change in income in that country and consequently country 2 imports from country 1 change. This means a change in exports of country 1 and consequently in this country's income, and so on.

This chain of events is known under the name of 'foreign repercussions', and the multiplier which takes account of them is called multiplier with foreign repercussions; for terminological convenience we shall call 'foreign trade multiplier without repercussions' the foreign trade multiplier treated in section 3.2.

Let us begin with the static model, which is made up of the following equations:

Country 1	Country 2
$C_1 = b_1Y_1,$	$C_2 = b_2Y_2,$
$I_1 = I_{01} + h_1Y_1,$	$I_2 = I_{02} + h_2Y_2,$
$M_1 = M_{01} + m_1Y_1,$	$M_2 = M_{02} + m_2Y_2,$
$X_1 = M_2,$	$X_2 = M_1,$
$Y_1 = C_1 + I_1 + X_1 - M_1.$	$Y_2 = C_2 + I_2 + X_2 - M_2.$

The equations express, in this order, the consumption function, the investment function, the import function, the fact that the exports of one country are the same as the imports of the country and the determination of national income in an open economy. I_0 and M_0 are autonomous components. Subscripts 1 and 2 respectively indicate country 1 and country 2. When we use b, h, m without subscripts, it will be because we refer generically to both countries.

Substituting, for both countries, from the first four equations into the fifth and rearranging terms, we have

$$\begin{aligned}(1 - b_1 - h_1 + m_1)Y_1 - m_2Y_2 &= I_{01} + M_{02} - M_{01},\\ -m_1Y_1 + (1 - b_2 - h_2 + m_2)Y_2 &= I_{02} + M_{01} - M_{02},\end{aligned} \tag{10.7}$$

from which

$$\begin{aligned}Y_1 &= \frac{(1 - b_2 - h_2)(I_{01} + M_{02} - M_{01}) + m_2(I_{01} + I_{02})}{(1 - b_1 - h_1 + m_1)(1 - b_2 - h_2 + m_2) - m_1m_2},\\ Y_2 &= \frac{(1 - b_1 - h_1)(I_{02} + M_{01} - M_{02}) + m_1(I_{01} + I_{02})}{(1 - b_1 - h_1 + m_1)(1 - b_2 - h_2 + m_2) - m_1m_2},\end{aligned} \tag{10.8}$$

and, considering the variations

$$\begin{aligned}\Delta Y_1 &= \frac{(1 - b_2 - h_2)(\Delta I_{01} + \Delta M_{02} - \Delta M_{01}) + m_2(\Delta I_{01} + \Delta I_{02})}{(1 - b_1 - h_1 + m_1)(1 - b_2 - h_2 + m_2) - m_1m_2},\\ \Delta Y_2 &= \frac{(1 - b_1 - h_1)(\Delta I_{02} + \Delta M_{01} - \Delta M_{02}) + m_1(\Delta I_{01} + \Delta I_{02})}{(1 - b_1 - h_1 + m_1)(1 - b_2 - h_2 + m_2) - m_1m_2}.\end{aligned} \tag{10.9}$$

Now the dynamics. The assumptions are the same as in Sect. 4.2, i.e. that in both countries C_t, I_t and M_t depend on Y_{t-1}. After the usual substitutions we have the difference system

$$\begin{aligned}Y_{1_t} &= (b_1 + h_1 - m_1)Y_{1_{t-1}} + m_2Y_{2_{t-1}} + (I_{01} + M_{02} - M_{01}),\\ Y_{2_t} &= m_1Y_{1_{t-1}} + (b_2 + h_2 - m_2)Y_{2_{t-1}} + (I_{02} + M_{01} - M_{02}).\end{aligned} \tag{10.10}$$

A particular solution of system (10.10) is obtained trying $Y_{1_t} = Y_{1_{t-1}} = \bar{Y}_1, Y_{2_t} = Y_{2_{t-1}} = \bar{Y}_2$, where $\bar{Y}_1, \bar{Y}_2$ are constant; the values that we obtain are the same as the static equilibrium values (10.8).

The characteristic equation of the homogeneous form of system (10.10) is

$$\begin{vmatrix} (b_1 + h_1 - m_1) - \lambda & m_2 \\ m_1 & (b_2 + h_2 - m_2) - \lambda \end{vmatrix} = 0. \tag{10.11}$$

Now since $b + h > m$ (see Chap. 4, Sect. 4.2, for the explanation), the coefficients are positive, so that we can apply the stability conditions expounded in Chap. 9, Sect. (9.2.2). From condition I we obtain the following necessary and sufficient conditions:

$$\begin{aligned}&1 - b_1 - h_1 + m_1 > 0,\\ &(1 - b_1 - h_1 + m_1)(1 - b_2 - h_2 + m_2) - m_1m_2 > 0.\end{aligned} \tag{10.12}$$

If we want only sufficient conditions, we can apply conditions IV and obtain

$$\begin{aligned}b_1 + h_1 &< 1,\\ b_2 + h_2 &< 1,\end{aligned} \tag{10.13}$$

From (10.12) and (10.13) we deduce the following conclusions:

(1) a necessary (but not sufficient) stability condition is that $1 - b_1 - h_1 + m_1$ and $1 - b_2 - h_2 + m_2$ are both positive;

(2) a sufficient (but not necessary) stability condition is that $b_1 + h_1$ and $b_2 + h_2$ are both less than 1;

(3) if $b_1 + h_1 > 1$ and also $b_2 + h_2 > 1$, the model is unstable (this follows from cond. XI);

(4) if one of the quantities $b_1 + h_1, b_2 + h_2$ is smaller than unity and the other greater than unity, the model can be stable or unstable according to the magnitude of m_1 and m_2.

To appreciate the economic meaning of these conclusions, let us recall from Sect. 4.2 that $b + h - m < 1$ is the stability condition for the foreign trade multiplier without repercussions and that $b + h < 1$ is the stability condition for the closed economy multiplier. Thus we can say:

(1) a necessary (but not sufficient) stability condition for the multiplier with foreign repercussions is that in both countries the foreign trade multiplier without repercussions is stable;

(2) a sufficient (but not necessary) stability condition for the multiplier with foreign repercussions is that for both countries in isolation the closed economy multiplier is stable;

(3) if in both countries in isolation the closed economy multiplier is unstable, the foreign trade multiplier with repercussions is unstable;

(4) if, when each country is considered in isolation, in one of them in closed economy multiplier is unstable whereas in the other it is stable, the foreign trade multiplier with foreign repercussions may be stable or unstable.

The multiplier model with foreign repercussions can be easily extended to any number of countries. Call $m_{ji}, i \neq j$, the (partial) propensity of country i to import from country j,

$$m_i = \sum_{\substack{j=1 \\ j \neq i}}^{n} m_{ji}$$

being the (total) propensity to import of country i. Similarly,

$$M_{0i} = \sum_{\substack{j=1 \\ j \neq i}}^{n} M_{0ji},$$

where M_{0ji} is autonomous imports of country i from country j. Then for any country $i = 1, 2, \ldots\ldots n$, we have

$$\begin{aligned} Y_{i_t} &= C_{i_t} + I_{i_t} + X_{i_t} - M_{i_t} \\ &= (b_i + h_i - m_i)Y_{i_{t-1}} + \textstyle\sum_{\substack{k=1 \\ k \neq i}}^{n} m_{ik} Y_{k_{t-1}} + I_{0i} + \sum_{\substack{k=1 \\ k \neq i}}^{n} M_{0ik} - M_{0i}. \end{aligned} \tag{10.14}$$

The characteristic equation of the homogeneous part of system (10.14) is

$$\begin{vmatrix} (b_1 + h_1 - m_1) - \lambda & m_{12} & \ldots & m_{1n} \\ m_{21} & (b_2 + h_2 - m_2) - \lambda & \ldots & m_{2n} \\ \ldots\ldots & \ldots\ldots & \ldots & \ldots\ldots \\ m_{n1} & m_{n2} & \ldots & (b_n + h_n - m_n) - \lambda \end{vmatrix} = 0.$$

Applying the stability conditions, we can deduce conclusions similar to those found for the two-country case. For example, $b_i + h_i < 1$, all i, is a sufficient stability condition, whereas if $b_i + h_i > 1$, all i, then the model is unstable. Moreover, the necessary and sufficient stability conditions also guarantee that a static equilibrium solution for the non-homogeneous system exists and is economically meaningful. A particular solution of system (10.14) is $\bar{\mathbf{Y}} = [\mathbf{I} - \mathbf{A}]^{-1}\nu$, where ν is the column vector of the autonomous terms and $\mathbf{A}$ is the matrix (obviously non-negative) of the coefficients of the difference system. It is well known (see, e.g., Solow, 1952) that the positivity of the leading principal minors of $[\mathbf{I} - \mathbf{A}]$ guarantees that $[\mathbf{I} - \mathbf{A}]^{-1}$ is non-negative.

10.3 Exercises

1. Show that another root of Eq. (10.6) is (1/2) with multiplicity $(n - 1)$. Hence deduce the general solution of system (10.3) (Hint: see McManus, 1962).

2. A country i imports directly from another country j when $m_{ji} > 0$. More generally, a country i is said to import *indirectly* from country j when country i imports directly from country j_1 which in turn imports directly from country j_2 ... which in turn imports directly from country j.

Show that, when each country imports directly or indirectly from all other countries, a sufficient stability condition is $b_i + h_i \leq 1$, with at least one strict inequality (Hint: use the alternative definition of indecomposability for a non-negative matrix $\mathbf{A}$, according to which for each pair of indices (i, j) there exists a set of indices $j_1, j_2, \ldots, j_l$ such that $a_{ij_1} a_{j_1 j_2} \cdots a_{j_l j} > 0$.)

3. Some interpretations of growth in an open economy are based on 'export led' models (see, for example, Gandolfo, 1995, Chap. 17). An element of these interpretations is the assumption that exports influence investment positively. Accordingly, replace the investment functions in the multiplier model with the functions $I_{1_t} = h_1 Y_{1_{t-1}} + h_3 X_{1_t}$, $I_{2_t} = h_2 Y_{2_{t-1}} + h_4 X_{2_t}$, $h_3 > 0, h_4 > 0$. Show that, for suitable values of h_3 and h_4, the system is capable of growth (Hint: for this result one must have a positive root greater than 1).

10.4 References

Cournot, A., 1899 (1838), *Researches into The Principles of the Wealth of Nations,* Chap. VII.

Gandolfo, G., 1995, *International Economics II - International Monetary Theory and Open-economy Macroeconomics,* Chaps. 13 and 17.

Machlup, F., 1943, *International Trade and the National Income Multiplier.*
McManus, M., 1962, Dynamic Cournot-type Oligopoly Models-A Correction.
Metzler, L.A., 1942, Underemployment Equilibrium in International Trade.
Metzler, L.A., 1950, A Multiple Region Theory of Income and Trade.
Okuguchi, K., 1976, *Expectations and Stability in Oligopoly Models.*
Okuguchi, K. and F. Szidarovszky, 1988, A Linear Oligopoly Model with Adaptive Expectations: Stability Reconsidered.
Okuguchi, K. and F. Szidarovszky, 1990, *The Theory of Oligopoly with Multi-Product Firms.*
Solow, R.M., 1952, On the Structure of Linear Models.
Theocharis, R.D., 1960, On the Stability of the Cournot Solution on the Oligopoly Problem.
Varian, H. R., 1992, *Microeconomic Analysis,* Chap. 16.

Part II

LINEAR DIFFERENTIAL EQUATIONS

Chapter 11

Differential Equations: General Principles

11.1 Definitions

An ordinary differential equation is a functional equation involving one or more of the derivatives y', y'', y''', etc. of an unknown function of time $y = f(t)$, which obviously is differentiable.

Before going on, it is as well to point out that we have adopted the prime notation (y' etc.) rather than the dot notation ($\dot{y}$ etc.) to denote time derivatives because the dot notation becomes awkward when one has to deal with higher-order derivatives.

We have called the equation an *ordinary* differential equation since the unknown function is a function of one argument; if the independent variables were more than one, partial derivatives would appear in differential equation, which would be partial differential equation (we shall not treat this type of functional equation).

The order of a differential equation is given by the highest derivative appearing in the equation.

After what we said in the Introduction, it will be clear that to solve (or to 'integrate') a differential equation means to find the unknown function that satisfies the relationship expressed in the equation.

Let us begin, as usual, by a simple example. Consider the differential equation $y' = a$, where a is a constant. From elementary integral calculus it follows that $y = at + b$; let us remember, incidentally, that integration—the 'inverse' operation of differentiation—represents the solution of a differential equation. We note that in the solution an arbitrary constant b appears (the other constant a is known, since it appear in the differential equation). This is not surprising, since we know that differentiation eliminates such constant, so that from $y' = a$. Let us now consider the second order differential equation $y'' = a$. Performing two successive integration we obtain $y = \frac{1}{2}at^2 + bt + c$;

now, two arbitrary constants, b and c, appear in the solution. It is easy to check that, if we differentiate the function y twice, such arbitrary constants disappear one after the other, so that from $y = \frac{1}{2}at^2 + bt + c$ we obtain, in fact, $y'' = a$, that is the differential equation from which we started.

We shall see later on how the arbitrary constant(s) can be determined through additional conditions; what interests us here is to note that we can induce, from the reasoning above, the following important theorem:

Theorem 11.1 *The general solution of a differential equation of order* n *is a function of t which involves exactly* n *arbitrary constants.*[1]

11.2 Linear differential equations with constant coefficients

As we said in the Introduction, the differential equations most widely used in economic dynamics are linear and with constant coefficients, which are also the easiest to handle from the mathematical point of view. In this second part of the book we shall use the expression 'differential equation' (or even, when the meaning is clear from the context, simply 'equation') in the sense of 'ordinary differential equations, linear with constant coefficients'.

The general form of an n-th order differential equation is

$$a_0 y^{(n)} + a_1 y^{(n-1)} + ... + a_{n-1}y' + a_n y = g(t), \tag{11.1}$$

where $y^{(n)}, y^{(n-1)}$, etc., indicate the derivatives of the order $n, n-1$, etc.; the a's are given constants and $g(t)$ is a known function. Some a's may be zero, but of course a_0 must be different from zero if the equation is of order n.

Eq. (11.1) is called a *non-homogeneous* equation; the corresponding homogeneous form is

$$a_0 y^{(n)} + a_1 y^{(n-1)} + ... + a_{n-1}y' + a_n y = 0. \tag{11.2}$$

The reason for dealing with these two forms separately is that the solution of Eq. (11.1) can be obtained in a relatively simple manner when the solution of Eq. (11.2) is known.

[1] We have implicitly assumed that this function exists and is unique (apart from the arbitrary constants). Actually this assumption could be proved by means of an existence and uniqueness theorem, but we shall not treat such theorems. All types of equations considered this book are well behaved, in the sense that their solution exists and is unique. In general, it may be observed that the properties that the 'well behaved' functions used in economic theory are assumed to have, are usually more than enough to saticfy the requirements of any existence and uniqueness theorem.

11.2.1 The homogeneous equation

The following theorems are fundamental in the theory of homogeneous differential equations:

Theorem 11.2. *If $y_1(t)$ is a solution of (11.2), then $Ay_1(t)$—where A is an arbitrary constant—is also a solution.*

The proof is simple. Assume that $y_1(t)$ satisfies Eq. (11.2), and substitute $Ay_1(t)$ in the same equation, obtaining

$$a_0 Ay_1^{(n)} + a_1 Ay_1^{(n-1)} + ... + a_{n-1}Ay_1' + a_n Ay_1 = 0;$$

therefore

$$A[a_0 y_1^{(n)} + a_1 y_1^{(n-1)} + ... + a_{n-1}y_1' + a_n y_1] = 0. \tag{11.3}$$

If $Ay_1(t)$ has to be a solution, Eq. (11.3) must be satisfied. Since $y_1(t)$ is a solution of (11.2), the expression in square brackets vanishes, and so Eq. (11.3) is indeed identically satisfied. This proves the theorem.

Before going on to the next theorem, it is as well to recall the notion of *linearly independent* functions.

Given n functions $y_1(t), y_2(t), ..., y_n(t)$, they are said to be *linearly dependent* if n constants $A_1, A_2, ..., A_n$ exist, which do not all vanish, and such that the equation

$$A_1y_1(t) + A_2y_2(t) + ... + A_my_m(t) = 0 \tag{11.4}$$

is identically satisfied for all admissible values of t. If, on the contrary, this equation can be identically satisfied only with $A_1 = A_2 = ... = A_n = 0$, the functions are *linearly independent.*

Theorem 11.3 *If $y_1(t), y_2(t)$ are two distinct (i.e., linearly independent) solutions of the homogenous equation ($n > 1$), then $A_1y_1(t) + A_2y_2(t)$ is also a solution for any two arbitrary constants A_1, A_2.*

The proof is similar to that of Theorem 11.2 and is left as an exercise.

Theorem 11.3—called the *superposition theorem*—can easily be extended to any number $k \leq n$ of distinct solutions of Eq. (11.2), and gives us the procedure to obtain the general solution of the homogeneous equation. This procedure consists in finding n distinct solutions $y_1(t), y_2(t), ..., y_n(t)$ and combining them linearly, as stated in Theorem 11.4:

Theorem 11.4 *The general solution of Eq. (11.2) is given by*

$$f(t; A_1, A_2, ..., A_n) = A_1y_1(t) + A_2y_2(t) + ... + A_ny_n(t), \tag{11.5}$$

where $y_1(t), y_2(t), ..., y_n(t)$ are n *linearly independent solutions of Eq. (11.2), and $A_1, A_2, ..., A_n$ are arbitrary constants.*

The proof is straightforward: by Theorem 11.3, the function (11.5) is a solution of the differential equation (11.2). Since this function contains exactly n arbitrary constants, we can conclude—from Theorem 11.1—that it is the general solution of Eq. (11.2). The practical problem of how to find the n functions $y_1(t)$, $y_2(t)$, ..., $y_n(t)$ will be tackled in the following chapters; for the moment we observe that, given a homogenous equation of order n, a set of n linearly independent solution is called a *fundamental set.* The condition for a set of n solutions to form a fundamental set is contained in Theorem 11.5.

Theorem 11.5 *Let* $y_1(t), y_2(t), ..., y_n(t)$ *be n solutions of Eq. (11.2). They are linearly independent (i.e. form a fundamental set) if, and only if, the following determinant* (called the Wronski determinant or *Wronskian*)

$$W(t) = \begin{vmatrix} y_1(t) & y_2(t) & \dots & y_n(t) \\ y_1'(t) & y_2'(t) & \dots & y_n'(t) \\ \dots & \dots & \dots & \dots \\ y_1^{(n-1)}(t) & y_2^{(n-1)}(t) & \dots & y_n^{(n-1)}(t) \end{vmatrix} \tag{11.6}$$

is different from zero for all admissible values of t.

To prove this theorem, consider Eq. (11.4): since this relation is satisfied identically, it may be differentiated any number of times up to $n-1$. Hence we can write the following linear system

$$\begin{array}{lllll} A_1y_1(t) & +A_2y_2(t) & +...+ & A_ny_n(t) & = 0, \\ A_1y_1'(t) & +A_2'y_2(t) & +...+ & A_ny_n'(t) & = 0, \\ \dots & \dots & \dots & \dots & \dots \\ A_1y_1^{(n-1)}(t) & +A_2y_2^{(n-1)}(t) & +...+ & A_ny_n^{(n-1)}(t) & = 0, \end{array} \tag{11.7}$$

which is a system of homogeneous linear equations whose determinant is $W(t)$. According to a well-known theorem in elementary algebra, when $W(t) \neq 0$, system (11.7) admits only the null solution, i.e. Eq. (11.4) holds true if, and only if, $A_1 = A_2 = ... = A_n = 0$, which is the definition of linearly independent functions (see above). On the contrary, $W(t) = 0$ is the necessary and sufficient condition for system (11.7) to possess non-trivial solutions (i.e., solutions with at least one non-zero A_i, $i = 1, 2, ..., n$), which is the definition of linearly dependent functions.

11.2.2 The non-homogeneous equation

We have so far dealt with the homogeneous differential equation. We now prove the basic theorem concerning the solution of the non-homogeneous equation.

Theorem 11.6. *If* $f(t; A_1, A_2, ..., A_n)$ *is the general solution of (11.2)—where* $A_1, A_2, ..., A_n$ *are arbitrary constants—and* $\bar{y}(t)$ *is any particular solution of (11.1), i.e. any function that satisfies (11.1), then*

$$y(t) = \bar{y}(t) + f(t; A_1, A_2, ..., A_n) \tag{11.8}$$

is the general solution of (11.1).

This theorem can be proved by direct substitution of $y(t)$ in Eq.(11.1) and checking that this equation is satisfied. Since $y(t)$ contains exactly n arbitrary constants, it is the general solution of Eq. (11.1).

The general solution of the homogenous equation is thus only a part of the general solution of the non-homogenous equation, and so it is not 'general' with respect to the latter. This means that the expression 'general solution' must always be qualified. As a matter of terminology, note the following : (1) some authors use the world 'integral' (particular or general) instead of 'solution' but with the same meaning; (2) the expression 'particular solution' is also used (a) in the sense of a solution obtained from the general solution by giving specific values to the arbitrary constants, and (b) in the sense of any single non-general solution of the homogenous equation (i.e., to indicate any one of $y_1(t), y_2(t)$, etc.); (3) the expression 'complementary function' is used to indicate the general solution of the homogenous equation when considered as part of the general solution of the non-homogeneous equation, and the expression 'reduced equation' is used to indicate the homogeneous part of a non-homogeneous equation, i.e. the corresponding homogeneous equation obtained putting $g(t) \equiv 0$ in the course of the procedure to solve a non-homogeneous equation. To avoid confusion, we shall not adopt these uses.

Theorem 11.6 contains the method to follow for solving the non homogeneous equation:

(a) find a particular solution $\bar{y}(t)$ of the non-homogeneous equation;

(b) put $g(t) \equiv 0$ and solve the resulting homogeneous equation (often called the 'reduced' equation) by using theorem 11.4;

(c) add the two results.

Steps (a) and (b) can be taken in any order; step (c) gives the general solution of the non-homogeneous equation.

The particular solution of the non-homogenous equation will depend, *ceteris paribus*, on the form of the given function $g(t)$. This suggest the following general approach: *to find a particular solution of the non-homogeneous equation, try a function having the same form of $g(t)$ but with undetermined constant(s) (e.g., if $g(t)$ is a constant, try an undetermined constant; if it is an exponential function, try the same exponential function with an undetermined multiplicative constant, and so on). Substitute this function in the non-homogeneous equation and determine the coefficient(s) so that the equation is satisfied.*

This method—called method of *undetermined coefficients*—will be expounded in detail in the following chapter, where we shall also examine the cases in which it cannot be applied.

It is interesting to note, from the economic point of view, that in the general solution of the non-homogeneous equation the particular solution $\bar{y}(t)$ may usually be interpreted as the *equilibrium state* of the variable y (a stationary equilibrium or a moving equilibrium according to whether $\bar{y}(t)$ is a constant or a function of t). The component $f(t; A_1, A_2, ..., A_n)$ in Eq. (11.8) may then be interpreted as giving the *deviations* from the equilibrium. Of course, from the mathematical point of view it is always true that $y(t) - \bar{y}(t) = f(t; A_1, A_2, ..., A_n)$, independently of the possibility of giving an economic interpretation to the particular solution $\bar{y}(t)$.

The practical problem of how to find the n functions $y_1(t), y_2(t), ..., y_n(t)$ and—if the equation is non-homogeneous—the function $\bar{y}(t)$, will be tackled in the following chapters.

11.3 Determination of the arbitrary constants

The problem remains of how to determine the arbitrary constants A_i. To do this we need an adequate number of additional conditions. This need derives from the fact that the solution—namely Eq. (11.5) or Eq. (11.8) as the case may be—of the differential equation under consideration gives solely the *form* of the function $y(t)$. Hence, to determine the n arbitrary constants, n additional conditions are needed. These usually specify that the function and its derivatives take on known values at a certain point in time, normally at the initial point $t = 0$. When only one point of time is involved, the additional conditions are called *initial conditions*. Other side conditions are possible, for example that the function $y(t)$ passes through n different given points in the (t, y) plane, say $y(t_j^*) = y_j^*$, $j = 1, 2, ..., n$ (in this case we speak of *boundary conditions*). Initial conditions are easier to deal with than boundary conditions, hence the side conditions are usually taken to be initial conditions, unless the nature of the problem requires boundary conditions.

Let us then consider the initial conditions. We are given $y(t) = y(0)$, $y'(t) = y'(0), ..., y^{(n-1)}(t) = y^{(n-1)}(0)$ for $t = 0$, where $y(0), y'(0), ..., y^{(n-1)}(0)$ are known values. Substituting these values in the general solution, we obtain a system of n linear equations in the n unknowns $A_1, A_2, ..., A_n$. Consider for example the general solution of Eq. (11.1)

$$y(t) = A_1 y_1(t) + A_2 y_2(t) + ... + A_n y_n(t) + \bar{y}(t),$$

where $y_1(t), y_2(t), y_n(t)$ are n distinct solutions of the corresponding homogeneous equation. Since this holds identically, we can differentiate $n-1$ times;

then by letting $t = 0$ and substituting the given values $y(0), y'(0), ..., y^{(n-1)}(0)$ we obtain, after rearranging terms,

$$\begin{aligned} A_1 y_1(0) + A_2 y_2(0) + ... + A_n y_n(0) &= y(0) - \bar{y}(0), \\ A_1 y_1'(0) + A_2 y_2'(0) + ... + A_n y_n'(0) &= y'(0) - \bar{y}'(0), \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots & \quad \cdots\cdots\cdots\cdots \\ A_1 y_1^{(n-1)}(0) + A_2 y_2^{(n-1)}(0) + ... + A_n y_n^{(n-1)}(0) &= y^{(n-1)}(0) - \bar{y}^{(n-1)}(0), \end{aligned} \tag{11.9}$$

where of course $\bar{y}(0), \bar{y}'(0), ... \bar{y}^{(n-1)}(0)$ are absent if we consider the solution of the homogeneous equation (11.2). System (11.9) is a linear system whose determinant is

$$W(0) = \begin{vmatrix} y_1(0) & y_2(0) & ... & y_n(0) \\ y_1'(0) & y_2'(0) & ... & y_n'(0) \\ ... & ... & ... & ... \\ y_1^{(n-1)}(0) & y_2^{(n-1)}(0) & ... & y_n^{(n-1)}(0) \end{vmatrix}.$$

It is easy to see that $W(0)$ coincides with the Wronskian $W(t)$—as defined in Eq. (11.6)—for $t = 0$. Since the functions $y_1(t), y_2(t), ..., y_n(t)$ form a fundamental set, $W(t)$ is different from zero for any t, and so also for $t = 0$. It follows that $W(0) \neq 0$. Thus system (11.9) can always be solved.

We now have enough general principles to pass on to a detailed treatment of the differential equations of the various orders, but before doing that we would like to point out the great similarity that exists between differential and difference equations, so much so that the general theorems are the same. However, some dissimilarities also exist, which give rise to differences in the economic interpretation. The principal formal dissimilarity is that in differential equations t varies continuously and also the function $y = f(t)$ is continuous (that it must be continuous is obvious, since it is differentiable), whereas in difference equations t varies discontinuously over a set of equispaced values, and so the solution $y = f(t)$ is a function which is defined only corresponding to these values of t.

This implies that it is *not* immaterial whether we use differential or difference equations in the formalization of an economic problem. If we think that certain dynamic economic phenomenon takes place in a continuous way and without discontinuous lags, then the appropriate mathematical tool to use is differential equations, while if we think that it takes place in a discontinuous way, then the appropriate tool is difference equations. Unfortunately this sharp distinction is not so often possible. And when it is not possible (or if we do not want to make it), we must use more complex mathematical tools, e.g. mixed difference-differential equations (see Chap. 27, where the reader will also find a general discussion of continuous vs discrete time in economics). Thus, if we want to avoid these more complicated mathematical tools, we must give a judgement, from the economic point of view, as to which aspects are prevalent. If we think that the economic problem under consideration is mainly continuous and without discontinuous lags, then we shall use differential equations; if we think that in such phenomenon discontinuous lags, etc., are the main characteristic, then we shall use difference equations. It is important that this judgment be given from the beginning and clearly, since the use of the one rather than the other tool in formalizing an economic problem may give different economic results.

11.4 References

Allen, R.G.D., 1959, *Mathematical Economics,* Chap. 5, Sects. 5.1, 5.2.

Baumol, W.J., 1970, *Economic Dynamics,* Chap. 14, Sect. 1.

Ince, E.L., 1956 (1926), *Ordinary Differential Equations,* Chap. I, Sects. 1.1, 1.2; Chap. V, Sects. 5.1-5.3.

Ince, E.L., 1959, *Integration of Ordinary Differential Equations,* Chap. I, Sects. 1, 2; Chap. II, Sect. 37.

Chapter 12

First-order Differential Equations

The general form of these equations is

$$a_0 y' + a_1 y = g(t), \qquad a_0 \neq 0, \tag{12.1}$$

where a_0, a_1 are given constants and $g(t)$ is a known function. The constant a_0 must be different from zero, since if it were equal to zero the equation would not be a differential equation; on the contrary a_1 may be zero, in which case we have

$$y' = \frac{1}{a_0} g(t). \tag{12.2}$$

To solve Eq. (12.2) means to find a function such that its derivative equals the right-hand side of (12.2). This is fairly easy: integrating both sides we have

$$y(t) = \frac{1}{a_0} \int g(t) \mathrm{d}t + C,$$

where C is an arbitrary constant. Thus the solution of (12.2) can be found by the usual formulae and rules of integration. Indeed, as we have already noted in Chap. 11, the operation of integration (the inverse operation of differentiation) represents the solution of the simplest type of first order differential equation, expressed by (12.2).

12.1 Solution of the homogeneous equation

Let us go back to (12.1), assuming now that both a_0 and a_1 are different from zero. We begin by studying the corresponding homogeneous equation:

$$a_0 y' + a_1 y = 0, \tag{12.3}$$

which can be written as

$$y' + by = 0, \qquad \text{where } b \equiv a_1/a_0. \tag{12.4}$$

From (12.4) we have $y' = -by$. Thus we must find a function such that its derivative is equal to the function itself, multiplied by the constant $-b$. To solve this problem we proceed by degrees. Let us first consider the particular case in which $b = -1$, so that

$$y' = y. \tag{12.5}$$

From elementary differential calculus we know that the only function which is equal to its derivative for any value of the argument is e^t. This suggests the idea of trying, as a solution to (12.4), a function of the type $e^{\lambda t}$, where λ is a constant to be determined. Substituting in (12.4) (remember that $\mathrm{d}e^{\lambda t}/\mathrm{d}t = \lambda e^{\lambda t}$) we have

$$\lambda e^{\lambda t} + be^{\lambda t} = 0,$$

i.e.

$$e^{\lambda t}(\lambda + b) = 0. \tag{12.6}$$

If the function $e^{\lambda t}$ is a solution of (12.4), Eq.(12.6) must be satisfied for any t, and this is possible if, and only if,

$$\lambda + b = 0, \tag{12.7}$$

which is called the *characteristic (*or *auxiliary) equation* of the differential equation (12.4).

From (12.7) we obtain $\lambda = -b$ and so, according to the general principles expounded in Chap. 11, the general solution of (12.4) is

$$y = Ae^{-bt}, \tag{12.8}$$

where A is an arbitrary constant. The reader may check as an exercise that (12.8) does satisfy (12.4) for any t.

Another way for obtaining the same result is the following. Write (12.4) in the form

$$y'/y = -b \tag{12.9}$$

and observe that $y'/y = \mathrm{d}\log_e y/\mathrm{d}t$, so that integrating both sides of (12.9) we have

$$\log_e y = -bt + C, \tag{12.10}$$

where C is an arbitrary constant. From (12.10) it follows that

$$y(t) = e^{-bt+C} = e^C e^{-bt} \tag{12.11}$$

and putting $e^C = A$ we obtain (12.8).

In order to determine the arbitrary constant we need an additional condition. In fact, the solution of a differential equation gives only the form of the unknown function but not its position in the (t, y) Cartesian plane.

As soon as the function is constrained to pass through a given point, say (t^*, y^*), its position, which here depends on one arbitrary constant only, is determined and the arbitrariness of the constant disappears. More formally, let it be known that $y(t) = y^*$ for $t = t^*$, where t^* and y^* are given values. Substituting this values in (12.8) we have

$$y^* = Ae^{-bt^*}, \tag{12.12}$$

and so

$$A = y^*/e^{-bt^*}. \tag{12.13}$$

In economic problems the value of y in the initial period is usually assumed to be known, at least in principle, i.e. $y(t) = y_0$ for $t = 0$; it follows from (12.13) that $A = y_0$. In this case we speak of *initial condition.*

The behaviour over time of the function Ae^{-bt} depends on the *sign* of the parameter b. When b is negative (positive), then $-b$ is positive (negative) and e^{-bt} increases (decreases) monotonically as t increases. Thus the stability condition is $-b < 0$. Fig. 12.1 illustrates the various cases.

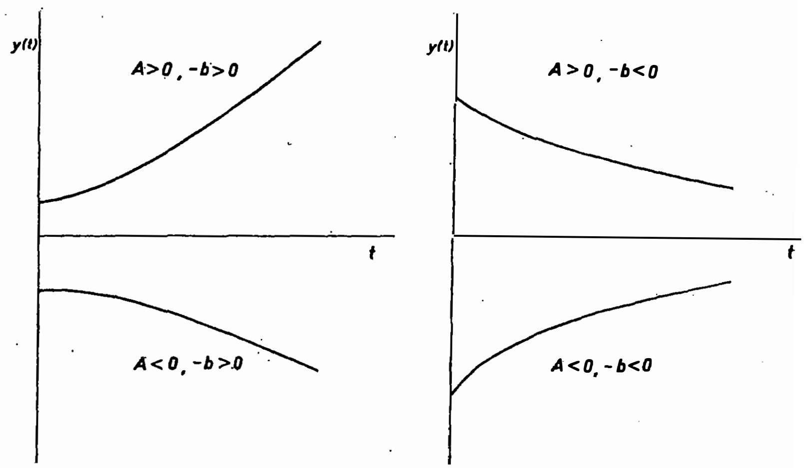


Figure 12.1: Solutions to a first-order differential equation

If we compare the solution of the homogeneous first-order differential equation with that of the homogeneous difference equation of the same order, we see that the former admits of a smaller variety of time paths than the latter. In particular, alternating movements (i.e. 'improper oscillations', as we termed them in Chap. 3) cannot occur with a first-order differential equation, and this is true for differential equations of any order. Another consideration, also true generally, is that in differential equations the qualitative behaviour of the solution depends only on the *sign* of a certain parameter, whereas in difference equations such behaviour depends both on the *sign* and on the *absolute value* of a parameter.

These disparities depend formally on the different behaviour of the functions $e^{\lambda t}$ and λ^t.

12.2 Particular solution of the non–homogeneous equation

The problem of finding a particular solution of the non-homogeneous equation will now be tackled, applying the general method of undetermined coefficientsto to the most common functions.

12.2.1 $g(t)$ is a constant

In this case, equation (12.1) becomes

$$a_0 y' + a_1 y = a, \tag{12.14}$$

where a is a given constant. As a particular solution try $\bar{y}(t) = \mu$, where μ is an undetermined constant. Substituting in (12.14) we have $0 + a_1\mu = a$, from which

$$\mu = a/a_1. \tag{12.15}$$

The method obviously fails if $a_1 = 0$; in this case let us try $\bar{y}(t) = \mu t$. Substituting in (12.14) we have $a_0\mu = a$, so that

$$\mu = a/a_0. \tag{12.16}$$

It must be stressed that the above treatment illustrates the following general prescription, which is a necessary complement to the general principle of : *if the function that we try as a particular solution does not work, let us try the same function multiplied by t*[1].

Let us note that, when $a_1 = 0$, the general solution of the homogeneous part of (12.14) is $y(t) = A$, since the characteristic equation yields $\lambda = 0$. Thus the general solution of our non-homogeneous equation is $y(t) = (a/a_0)t + A$. We could obtain the same result directly by integrating the equation $a_0 y' = a$ (see the beginning of this chapter). It is, however important to understand how the solution is obtained even in this particular case by applying general principles.

[1]In second- or higher-order equations it may be necessary , as we shall see , to multiply by t^2, t^3, etcetera. More generally, the prescription is to try the same function multiplied by a suitable polynomial in t.

12.2.2 $g(t)$ is an exponential function

When $g(t) = Be^{dt}$, where B and d are given constants[2], as a particular solution try Ce^{dt}, where C is an undetermined constant. Substitution in (12.1) yields

$$a_0 dCe^{dt} + a_1 Ce^{dt} = Be^{dt},$$

from which

$$[(a_0 d + a_1)C - B]e^{dt} = 0. \tag{12.17}$$

Equation (12.17) will be satisfied for any t if, and only if,

$$(a_0 d + a_1)C - B = 0.$$

which gives

$$C = \frac{B}{a_0 d + a_1}. \tag{12.18}$$

If $a_0 d + a_1 = 0$ (incidentally, this means that the root of the characteristic equation of the homogeneous part of the equation equals d) the method fails. Let us then try tCe^{dt}. Substituting in (12.1) we have

$$a_0(Ce^{dt} + tdCe^{dt}) + a_1 tCe^{dt} = Be^{dt},$$

from which

$$(a_0 d + a_1)tCe^{dt} + (a_0 C - B)e^{dt} = 0,$$

i.e., since $a_0 d + a_1 = 0$,

$$(a_0 C - B)e^{dt} = 0, \tag{12.19}$$

which yields

$$C = B/a_0. \tag{12.20}$$

12.2.3 $g(t)$ is a polynomial function of degree m

As an example, consider $g(t) = c_0 + c_1 t$, where c_0 and c_1 are given constants. Try $\bar{y}(t) = \alpha + \beta t$ as a particular solution, α and β being undetermined constants. Substitution in (12.1) yields

$$a_0\beta + a_1(\alpha + \beta t) = c_0 + c_1 t,$$

which gives

$$(a_1\beta - c_1)t + (a_1\alpha + a_0\beta - c_0) = 0, \tag{12.21}$$

Eq. (12.21) will be satisfied for any t if, and only if,

$$\begin{aligned} a_1\beta - c_1 &= 0, \\ a_0\beta + a_1\alpha - c_0 &= 0. \end{aligned} \tag{12.22}$$

[2]If $g(t) = \alpha^{\beta t}, \alpha > 0$, put $\beta \log_e \alpha = d$, and proceed as before.

The solution of system (12.21) yields

$$\alpha = (a_1c_0 - a_0c_1)/a_1^2, \qquad \beta = c_1/a_1, \tag{12.23}$$

When $a_1 = 0$, we try $\bar{y}(t) = \alpha t + \beta t^2$. Substituting in (12.1) we obtain (the details are left as an exercise)

$$\begin{aligned} \alpha &= c_0/a_0, \\ \beta &= c_1/2a_0. \end{aligned} \tag{12.24}$$

12.2.4 $g(t)$ is a trigonometric function of the sine-cosine type

In this case $g(t) = B_1 \cos\omega t + B_2 \sin\omega t$, where B_1, B_2, ω are given constants. As a particular solution try $\alpha \cos\omega t + \beta \sin\omega t$, where α and β are undetermined constants[3]. Substituting in (12.1) yields

$$a_0(-\alpha\omega \sin\omega t + \beta\omega \cos\omega t) + a_1(\alpha \cos\omega t + \beta \sin\omega t) = B_1 \cos\omega t + B_2 \sin\omega t,$$

from which

$$(a_0\omega\beta + a_1\alpha - B_1)\cos\omega t + (a_1\beta - a_0\omega\alpha - B_2)\sin\omega t = 0. \tag{12.25}$$

Equation (12.25) will be satisfied for any t if, and only if,

$$\begin{aligned} a_0\omega\beta + a_1\alpha - B_1 &= 0, \\ a_1\beta - a_0\omega\alpha - B_2 &= 0. \end{aligned} \tag{12.26}$$

Equations (12.26) are a system of two linear equations, whose solution will give the values of α and β. By now the student should be able to examine the case in which this method fails.

12.2.5 $g(t)$ is a combination of the previous functions

When $g(t)$ is a combination of two or more of the functions treated above in the various cases, as a particular solution we can try the same combination of functions with undetermined coefficients. This is then substituted in Eq. (12.1), and the resulting expression—after collecting the terms containing t separately from the others if the case–usually gives the equations (in terms of the undetermined coefficients) that must hold for Eq. (12.1) to be identically satisfied. The solution of these equations gives the sought-for values of the undetermined coefficients.

Although the cases exemplified cover the usual situations, we may have to cope with the case in which $g(t)$ does not belong to any of the categories examined above.

[3]Note that the trial function must be $\alpha \cos\omega t + \beta \sin\omega t$ also when $B_1 = 0$ or $B_2 = 0$.

12.2.6 $g(t)$ is a generic function of time. The method of variation of parameters

When $g(t)$ is a generic function of time, there are various methods that can be applied in the place of the method of undetermined coefficients. Lagrange's classical method of variation of parameters, Heaveside's operational calculus, Laplace transforms are the best known alternative methods. We agree with Samuelson (1947, pp. 397-8) when he writes, in relation to ordinary differential equations with constant coefficients, that 'a careful examination of the Heaveside-Cauchy operational calculus, of Bromwich-Wagner contour integrals, of Laplace transforms, etc., will show that where these differ from the classical methods the advantages are in favor of the latter'. Indeed, we have found that Lagrange's method of variation of parameters is particularly useful since it is in principle applicable to any integrable $g(t)$. Hence it can also be used in those cases in which we do not actually know the functional form of $g(t)$.

This is a general method of solving a differential equation by considering the arbitrary constants that appear in the known solution of a simpler equation, as variable (i.e., as functions of t), and determining them so that the more general equation is identically satisfied.

In our case the equation to solve is the non-homogeneous equation, and the simpler equation with known solution is the homogeneous equation. Hence we start from Eq. (12.8) and posit

$$y(t) = A(t)e^{-bt}, \tag{12.27}$$

where $A(t)$ is an undetermined function, assumed to be differentiable. Differentiating Eq. (12.27) we obtain

$$y'(t) = A'(t)e^{-bt} - bA(t)e^{-bt}. \tag{12.28}$$

Let us now rewrite Eq. (12.1) as

$$y' + by = \frac{1}{a_0}g(t) \tag{12.29}$$

and substitute Eqs. (12.27) and (12.28) into Eq. (12.29). This yields

$$A'(t)e^{-bt} = \frac{1}{a_0}g(t),$$

that is

$$A'(t) = \frac{1}{a_0}g(t)e^{bt}. \tag{12.30}$$

Integrating both sides of Eq. (12.30) we obtain

$$A(t) = \frac{1}{a_0}\int g(t)e^{bt}\mathrm{d}t + B, \tag{12.31}$$

where B is an arbitrary constant. Hence by substitution into Eq. (12.27) we obtain

$$y(t) = Be^{-bt} + \frac{1}{a_0}e^{-bt}\int g(t)e^{bt}\mathrm{d}t \tag{12.32}$$

which is the general solution of Eq. (12.1). Note that the method of variation of parameters has directly given the solution of the non-homogeneous equation starting from the solution of the corresponding homogeneous equation. In Eq. (12.32), the second term on the right-hand side can be interpreted as the particular solution, i.e.

$$\bar{y}(t) = \frac{1}{a_0}e^{-bt}\int g(t)e^{bt}\mathrm{d}t \tag{12.33}$$

is the general formula for finding a particular solution for an arbitrarily given $g(t)$. Thus we have reduced the problem of finding a particular solution to a problem in integral calculus.

Formula (12.33) can, of course, be applied to the standard forms of $g(t)$ treated in the previous paragraphs. Consider, for example, the case in which $g(t) = a$, a constant. Then Eq. (12.33) gives

$$\bar{y}(t) = \frac{1}{a_0}e^{-bt}a\int e^{bt}\mathrm{d}t = \frac{a}{a_0}e^{-bt}\frac{1}{b}e^{bt} = \frac{a}{a_0 b} = \frac{a}{a_1},$$

which coincides with the result previously found in Eq. (12.15).

12.3 General solution of the non-homogeneous equation

After finding a particular solution of the non-homogeneous equation, we can write its general solution as

$$y\,(t) = Ae^{-bt} + \bar{y}(t). \tag{12.34}$$

Let us now determine the arbitrary constant A. It is important to note that this constant must be determined, given an additional condition, on the basis of the general solution of the equation concerned. This means that, if the equation is non-homogeneous, formula (12.13) does not hold, and we must find a new one. The method is the same: given that $y(t) = y^*$ for $t = t^*$, we substitute in (12.34) and obtain

$$y^* = Ae^{-bt^*} + \bar{y}(t^*),$$

which gives

$$A = e^{bt^*}[y^* - \bar{y}(t^*)]. \tag{12.35}$$

As we have already said above, in economics the initial value of y is usually assumed to be known, at least in principle (whence the name *initial condition*), so that $A = y_0 - \bar{y}_0$, the initial deviation between $y(t)$ and $\bar{y}(t)$.

Let us recall (see Chap. 11, Sect. 11.3) that in economic applications the particular solution $\bar{y}(t)$ can usually be interpreted as the *equilibrium value* (a stationary or a moving equilibrium according to whether $\bar{y}(t)$ is a constant or a function of t) of the variable y. Given this interpretation the general solution of the homogeneous part of the equation can be interpreted as giving the time path of the *deviations* from equilibrium since $y(t) - \bar{y}(t) = Ae^{-bt}$.

12.4 Continuously distributed lags and partial adjustment equations

We have already examined distributed lag equations in discrete time (see Chap. 3, Sect. 3.4). The results obtained in discrete time can be extended to continuous time. A *continuously distributed lag* equation has the form

$$y(t) = \int_0^\infty [f(\tau)x(t-\tau)]\mathrm{d}\tau, \quad \int_0^\infty f(\tau)\mathrm{d}\tau = 1, \quad \lim_{\tau\to\infty} f(\tau) = 0, \tag{12.36}$$

where $f(\tau)$ is the time form of the weighting function; for simplicity we have assumed that the integral of the weights is unity instead of an arbitrary positive constant. In this latter case we would write

$$y(t) = b\int_0^\infty [f(\tau)x(t-\tau)]\mathrm{d}\tau, \quad \int_0^\infty f(\tau)\mathrm{d}\tau = 1/b.$$

The continuous equivalent to the discrete-time geometric lag distribution is the *exponential lag distribution*, where $f(\tau) = \alpha e^{-\alpha\tau}$, $\alpha > 0$, so that

$$y(t) = \int_0^\infty \alpha e^{-\alpha\tau}x(t-\tau)\mathrm{d}\tau. \tag{12.37}$$

We now show that Eq. (12.37) is equivalent to the *partial adjustment equation* in continuous time

$$y'(t) = \alpha[x(t) - y(t)], \tag{12.38}$$

where in economic applications $x(t)$ can be interpreted as the desired or potential value of $y(t)$, and α is the speed of adjustment of the actual to the desired value. Let us perform a change of variable from τ to $s = t - \tau$ in the integral (12.37) so that, by elementary rules of integral calculus[4], we have

[4]Remember that $\int_a^b \psi(\tau)d\tau = \int_c^d \psi[f(s)]f'(s)ds$, where $\tau = f(s)$, $a = \psi(c)$, $b = \psi(d)$. In our case, $\tau = t - s$, and $s = t$ when $\tau = 0$ and $s = -\infty$ when $\tau = \infty$. Further remember that the property according to which the definite integral changes sign when the limits of integration are reversed also holds when one of the limits is infinite, provided that the integral converges. For these properties see, e.g. Gillespie (1959), Courant (1961), or any calculus textbook.

$$y(t) = -\int_t^{-\infty} \alpha e^{-\alpha(t-s)} x(s)\mathrm{d}s = \int_{-\infty}^t \alpha e^{-\alpha(t-s)} x(s)\mathrm{d}s. \tag{12.39}$$

Since t is a constant with respect to the integral, $\alpha e^{-\alpha t}$ can be considered as a multiplicative constant; therefore

$$y(t) = \alpha e^{-\alpha t} \int_{-\infty}^t e^{\alpha s} x(s)\mathrm{d}s,$$

whence

$$y(t)e^{\alpha t} = \alpha \int_{-\infty}^t e^{\alpha s} x(s)\mathrm{d}s. \tag{12.40}$$

Differentiating Eq. (12.40) with respect to time we have

$$\alpha y(t)e^{\alpha t} + [y'(t)]e^{\alpha t} = \alpha \frac{\mathrm{d}}{\mathrm{d}t}\left\{\int_{-\infty}^t e^{\alpha s} x(s)\mathrm{d}s\right\}.$$

Performing the differentiation of the integral[5] with respect to a parameter that appears in the upper limit, we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\left\{\int_{-\infty}^t e^{\alpha s} x(s)\mathrm{d}s\right\} = e^{\alpha t} x(t),$$

and so

$$\alpha y(t)e^{\alpha t} + [y'(t)]e^{\alpha t} = \alpha e^{\alpha t} x(t),$$

whence, eliminating $e^{\alpha t}$ and rearranging terms,

$$y'(t) = \alpha[x(t) - y(t)],$$

which coincides with the partial adjustment equation defined in Eq. (12.38).

The coefficient of adjustment α can be interpreted as the reciprocal of the mean time-lag, which in turn has an interesting economic interpretation as the time required for about 63% of the discrepancy between $y(t)$ and $x(t)$ to be eliminated by changes in $y(t)$ following a change in $x(t)$. Let us begin with the *mean time-lag*, that in continuous time is defined as

$$\overline{\tau} = \int_0^\infty \tau f(\tau)\mathrm{d}\tau, \tag{12.41}$$

[5]This a particular case of the general formula for the differentiation of an integral with respect to a parameter that occurs in the limits of integration as well as in the integrand, i.e.

$$\frac{\mathrm{d}}{\mathrm{d}t}\left\{\int_{\phi_1(t)}^{\phi_2(t)} f(t,y)dy\right\} = \int_{\phi_1(t)}^{\phi_2(t)} \frac{\partial f(t,y)}{\partial t}dy - \phi_1'(t)\left[f\left(t,\phi_1(t)\right)\right] + \phi_2'(t)\left[f\left(t,\phi_2(t)\right)\right]$$

(see any advanced calculus textbook or Courant, 1961, Vol. II, p. 220). In our particular case, the parameter (t) only appears in the upper limit.

i.e., as the weighted average of the times, where the weights are given by the lag function $f(\tau)$.

With an exponential weighting function we have

$$\overline{\tau} = \alpha \int_0^\infty \tau e^{-\alpha\tau}\mathrm{d}\tau = \frac{1}{\alpha}, \tag{12.42}$$

since (see exercise 4) $\int_0^\infty \tau e^{-\alpha\tau}\mathrm{d}\tau = 1/\alpha^2$. Therefore, when $\alpha \to \infty$ the mean time-lag tends to zero, namely $y(t)$ tends to adjust to $x(t)$ immediately.

Let us now come to the interpretation of $1/\alpha$. Consider the expression for $y(t)$ given in Eq. (12.39) and assume that θ units of time ago a sustained change in x, say Δx, took place, where Δx is a given constant. Then the change in y can be shown to be (see exercise 5)

$$\Delta y(t) = \int_{t-\theta}^t \alpha e^{-\alpha(t-s)} \Delta x(t-\theta)\mathrm{d}s = \Delta x(t-\theta)\int_{t-\theta}^t \alpha e^{-\alpha(t-s)}\mathrm{d}s, \tag{12.43}$$

from which, by returning to the original variable τ (i.e., performing a change of variable from s to $\tau = t - s$), we get

$$\Delta y(t) = \Delta x(t-\theta)\int_0^\theta \alpha e^{-\alpha\tau}\mathrm{d}\tau = \Delta x(t-\theta)\left[-e^{-\alpha\tau}\right]_0^\theta = \Delta x(t-\theta)(1-e^{-\alpha\theta}). \tag{12.44}$$

Finally, by letting the so far undetermined θ take on the value $1/\alpha$ (i.e., the mean time-lag) and carrying out the computations, we obtain

$$\Delta y(t) \simeq 0.632\Delta x(t-\theta).$$

Therefore, as stated above; $\theta = 1/\alpha$ is the time required for about 63% of the discrepancy between $y(t)$ and $x(t)$ to be eliminated by changes in $y(t)$ following a change in $x(t)$. This interpretation is always formally valid, but is particularly interesting when $x(t)$ is the desired or 'partial equilibrium' value of $y(t)$.

From Eq. (12.44) it follows that, in the case of infinite speed of adjustment, i.e. for $\alpha \to \infty$, the mean time-lag tends to zero, and we also have $\Delta y(t) = \Delta x(t-\theta)$ however small θ is, which means an instantaneous adjustment of $y(t)$ to $x(t)$. More formally, given an arbitrarily small $\theta > 0$, and an arbitrarily small $\epsilon > 0$, it is always possible to find a value of α sufficiently great so as to make $e^{-\alpha\theta} < \epsilon$ and hence to make $\Delta y(t)$ as close as we wish to $\Delta x(t-\theta)$.

12.5 Exercises

12.5.1 Example

Let us solve the following non-homogeneous equation

$$y' + y = te^t, \quad y(0) = 1. \tag{12.45}$$

The characteristic equation of the homogeneous part is $\lambda + 1 = 0$, hence the general solution of the corresponding homogeneous equation is

$$Ae^{-t}. \tag{12.46}$$

Since the given function of time is a combination (product) of a linear and an exponential function, to find a particular solution of the non-homogeneous equation let us try

$$\overline{y}(t) = (\alpha + \beta t)e^t. \tag{12.47}$$

Substitution in Eq.(12.45), account being taken that $\overline{y}'(t) = \beta e^t + (\alpha+\beta t)e^t$, yields

$$2(\alpha + \beta t)e^t + \beta e^t = te^t,$$

from which, rearranging terms,

$$(2\beta - 1)te^t + (2\alpha + \beta)e^t = 0. \tag{12.48}$$

Equation (12.48) will be identically satisfied if, and only if,

$$\begin{aligned} 2\beta - 1 &= 0, \\ 2\alpha + \beta &= 0, \end{aligned}$$

from which $\beta = 1/2, \alpha = -1/4$. Hence the general solution of Eq. (12.45) is

$$y(t) = Ae^{-t} + \left(\frac{1}{2}t - \frac{1}{4}\right)e^t. \tag{12.49}$$

The same result can be obtained by the method of variation of parameters. The use of Eq. (12.33) immediately gives

$$\overline{y}(t) = e^{-t}\int te^te^t\mathrm{d}t = e^{-t}\int te^{2t}\mathrm{d}t. \tag{12.50}$$

Now, $\int te^{2t}\mathrm{d}t = \left(\frac{1}{2}t - \frac{1}{4}\right)e^{2t}$ (see any calculus textbook), hence Eq. (12.50) gives $\overline{y}(t) = \left(\frac{1}{2}t - \frac{1}{4}\right)e^t$. The method of variation of parameters is direct and efficient, but requires the knowledge of some integral calculus, while the less direct method of undetermined coefficients does not require such a knowledge. Hence in the standard cases it may be simpler to use the method of undetermined coefficients.

To determine A it is enough to observe that $e^{-t} = e^t = 1$ for $t = 0$, hence from Eq. (12.49) we get $y(0) = A - \frac{1}{4}$. The initial condition $y(0) = 1$ then yields $A = \frac{5}{4}$.

12.5.2 Other exercises

1. Solve the following differential equations (in all cases, $y(0) = 1$):

(i)	$y' + y = 0$
(ii)	$y' + 3y = 0$
(iii)	$2y' - 12y = 0$
(iv)	$y' - y = 0$
(v)	$3y' + 6y = 0$

2. In the differential equations under exercise 1, respectively introduce the following given functions of time:

(i)	$g(t) = 2 + 2t$
(ii)	$g(t) = 2$
(iii)	$g(t) = 20\sin 2t$
(iv)	$g(t) = -t$
(v)	$g(t) = e^{-2t}$

and determine the general solution of the resulting non-homogeneous equations.

3. In the exponential lag distribution, check that $\int_0^\infty \alpha e^{-\alpha\tau} = 1$.

4. Show that $\int_0^\infty \tau e^{-\alpha\tau}\mathrm{d}\tau = 1/\alpha^2$ (Hint: a primitive of the integrand is $-e^{-\alpha\tau}(\alpha\tau+1)/\alpha^2$. Also, application of l'Hôpital's rule shows that $(\alpha\tau+1)/e^{\alpha\tau}$ tends to zero as $\tau \to \infty$, hence the infinite integral can be evaluated as a standard definite integral).

5. Prove Eq. (12.43). (Hint: define a variable X which is equal to x for times farther away than θ units ago, and equal to $x + \Delta x$ for all times from $t - \theta$ up to t, and let $Y(t)$ be the new value of y following the change in x. Since the function $X(s)$ is continuous in the interval $(-\infty, t)$ except for a finite discontinuity at $s = t - \theta$, it is integrable over this interval, so that

$$\begin{aligned} Y(t) &= \int_{-\infty}^{t} \alpha e^{-\alpha(t-s)}X(s)\mathrm{d}s = \int_{-\infty}^{t-\theta} \alpha e^{-\alpha(t-s)}x(s)\mathrm{d}s \\ &+ \int_{t-\theta}^{t} \alpha e^{-\alpha(t-s)}[x(s) + \Delta x(s)]\mathrm{d}s \end{aligned}$$

exists. Then observe that $\Delta y(t) = Y(t) - y(t)$, where $y(t) = \int_{-\infty}^{t} \alpha e^{-\alpha(t-s)}x(s)\mathrm{d}s$ is the value that would obtain in the case of no change in x).

12.6 References

Allen, R.G.D., 1959, *Mathematical Economics,* Chap. 5, Sects. 5.3, 5.8.

Baumol, W.J., 1970, *Economic Dynamics*, Chap. 14, Sect. 2.
Braun, M., 1986, *Differential Equations and their Applications*, Chap. 1, Sects. 1.2, 1.2.
Courant, R., 1961, *Differential and Integral Calculus*, Vol. II, Chap. IV, Sect. 1.
Gillespie, R.P., 1959, *Integration*, Chaps. I, II, V, VI.
Samuelson, P.A., 1947, *Foundations of Economic Analysis*, Appendix B, §§14-29.

Chapter 13

First-order Differential Equations in Economic Models

13.1 Stability of supply and demand equilibrium

We know from microeconomics that in a perfectly competitive market, equilibrium is determined by the point at which the supply function and the demand function are equal. Here we limit ourselves to the simple case in which demand and supply of a commodity are assumed to depend solely on the price of that commodity, *ceteris paribus* (i.e., we are making a partial equilibrium analysis). A general equilibrium analysis requires more complex mathematical tools and will be examined further on (see Chaps. 19 and 23).

Let us now examine what happens when the system is not in equilibrium, that is when price and quantity do not coincide with the respective equilibrium values. This may be the case either because the system, previously in equilibrium, has been displaced from it by accidental causes, or because the system has never been in equilibrium. It is obvious that these are two aspects of the same problem.

The question we are asking is the following: will the system tend to move towards its equilibrium or not? The problem that we have posed is the problem of the *stability of equilibrium*. It is important to distinguish two concepts of stability: *static stability* and *dynamic stability*[1]. Static stability only tells us whether the economic forces that act on the system tend to make it move towards the equilibrium point, but does not tell us anything about the actual time path of the system nor, therefore, whether the system converges over time to the equilibrium point. It is true that the economic forces tend to 'push' the system towards its equilibrium, but this does not

[1] Many other distinctions are made with regard to the concept of stability (perfect and imperfect, asymptotic, local and global etc.). Some of them will be treated later on, in this section. See also Chap. 19, Sect. 19.1, and Chap. 21.

exclude, for example, the case in which the equilibrium point is 'overtaken' again and again in opposite directions giving rise to oscillations which may in principle be undamped. Therefore, the study of static stability is not sufficient, and it is necessary to study dynamic stability: the latter, being based on functional equations, is able to solve the problem left unsolved by the former. Therefore, the 'true' concept of stability is the dynamic one, which is the one that we have implicitly adopted in the preceding chapters when we have examined the stability of equilibrium in the various models.

In order to study the stability of equilibrium—be it static or dynamic—it is necessary to make *assumptions about the behaviour* of the relevant variables out of equilibrium. Since, in principle, it is possible to make several equally plausible such assumptions, there follows the *relativity of stability conditions,* since in some cases a given equilibrium point may be stable or unstable according to the different assumptions made.

Let us now examine the problem from which we started, i.e. the stability of supply and demand equilibrium. The main behaviour assumptions are the Walrasian assumption and the Marshallian assumption. According to the *Walrasian* assumption, price tends to increase (decrease) if excess demand is positive (negative). According to the *Marshallian* assumption, quantity tends to increase (decrease) if excess demand price is positive (negative). Figure 13.1 serves to clarify the distinction between 'excess demand' and 'excess demand price'.

Excess demand is the difference between the quantity that buyers are willing to buy at any given price and the quantity that sellers are willing to supply at the same price.

In Fig. 13.1(a), when price is p' excess demand is positive and is measured by the segment $\overline{A'B'}(=\overline{a'b'})$; at price p'' excess demand is negative and is measured (in absolute value) by the segment $\overline{A''B''}(=\overline{a''b''})$.

Excess demand price is the difference between the price that buyers are willing to pay for any given quantity and the price that is required to call forth the supply of the same quantity. In Fig. 13.1(b) when the quantity is q' excess demand price is positive and is measured by the segment $\overline{A'B'}(=\overline{a'b'})$; at quantity q'' excess demand price is negative and is measured (in absolute value) by the segment $\overline{A''B''}(=\overline{a''b''})$.

The idea underlying the Walrasian assumption is that, when there is a positive (negative) excess demand, unsatisfied buyers (sellers) bid the price up (down). The idea underlying the Marshallian assumption is that, when there is a positive (negative) excess demand price, producers realize that they can profitably increase (decrease) the quantity supplied.

Let us now examine the stability conditions from the *static stability* point of view. If we adopt the Walrasian assumption, equilibrium is stable if a price increase (decrease)—caused by positive (negative) excess demand—

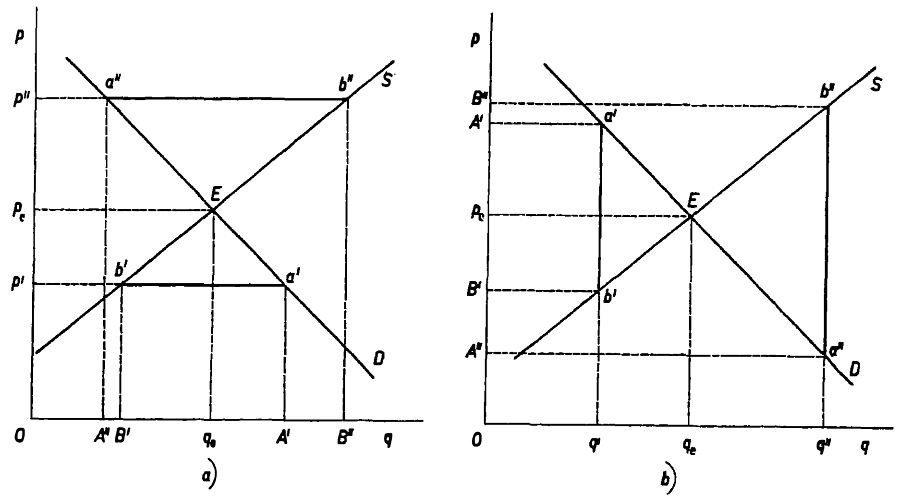


Figure 13.1: a) excess demand, b) excess demand price

diminishes (increases) excess demand, i.e. if

$$\frac{\mathrm{d}E(p)}{\mathrm{d}p}=\frac{\mathrm{d}[D(p)-S(p)]}{\mathrm{d}p}<0, \tag{13.1}$$

that is

$$\frac{\mathrm{d}D}{\mathrm{d}p}-\frac{\mathrm{d}S}{\mathrm{d}p}<0. \tag{13.2}$$

If we adopt the Marshallian assumption, equilibrium is stable if an increase(decrease) in quantity—caused by a positive (negative) excess demand price—reduces (increases) the excess demand price, i.e. if

$$\frac{\mathrm{d}E^{-1}}{\mathrm{d}q}=\frac{\mathrm{d}[p_d(q)-p_s(q)]}{\mathrm{d}q}<0, \tag{13.3}$$

that is

$$\frac{\mathrm{d}p_d}{\mathrm{d}q}-\frac{\mathrm{d}p_s}{\mathrm{d}q}<0. \tag{13.4}$$

Since $p_d=p_d(q)$ is the inverse of the function $D=D(p)$, and $p_s=p_s(q)$ is the inverse of the function $S=S(p)$, it follows from a well known theorem of elementary calculus that

$$\frac{\mathrm{d}p_d}{\mathrm{d}q}=\left(\frac{\mathrm{d}D}{\mathrm{d}q}\right)^{-1},$$

so that

$$\frac{dp_s}{dq} = \left(\frac{dS}{dq}\right)^{-1}.$$

Therefore, Eq. (13.4) can be written as

$$\left(\frac{dD}{dp}\right)^{-1} - \left(\frac{dS}{dp}\right)^{-1} < 0. \tag{13.5}$$

Before passing to compare conditions (13.2) and (13.5), let us examine the *dynamic stability* conditions. The method used to study the dynamic stability is the following: the behaviour assumption is formalized as a functional equation, which is then solved to determine the time path of the relevant variables.

The dynamic formalization of the Walrasian assumption is the following:

$$p' = f[D(p) - S(p)], \tag{13.6}$$

where

$$\text{sgn } f[...] = \text{sgn } [...], \qquad f[0] = 0, \qquad f'[0] > 0.$$

The notation sgn $f[...] =$ sgn $[...]$ means that f is a *sign-preserving function,* i.e., the dependent variable has the same sign as the independent variable (which in this case is excess demand): therefore, if excess demand is positive (negative) the time derivative of p is positive (negative), i.e. p is increasing (decreasing), and this is what the Walrasian assumption says. The condition $f[0] = 0$ means that the time derivative of p is zero when excess demand is zero, and this is obvious since zero excess demand means that the system is in equilibrium, i.e. the price does not vary. The condition $f'[0] > 0$ means that the function is increasing at the point where it passes from negative to positive values.

Equation (13.6) is a first-order differential equation and to solve it we need to know the form of the functions f, D, S. There are two methods. Either we give such functions an arbitrary form (usually linear, for simplicity's sake) or we expand them in Taylor series at the equilibrium point and neglect all terms of order higher than the first (i.e., we make a linear approximation of the functions; thus the results that we obtain are valid only in a neighbourhood of the equilibrium point and, consequently, the stability conditions are *local*). In the first case, we have

$$p' = c[a + bp - (a_1 + b_1p)] = c(b - b_1)p + c(a - a_1).$$

Since $p_e = (a - a_1)/(b_1 - b)$, it follows that $c(a - a_1) = -c(b - b_1)p_e$ and so $p' = c(b - b_1)(p - p_e)$. In the second case, we must linearise f, D, and S. Linearising f, at the point where $D = S$, we have $p' = c[D(p) - S(p)]$, where $c \equiv f'[0]$. Linearising D and S at the point where $p = p_e$, we obtain

$$p' = c[b(p - p_e) + D_e - b_1(p - p_e) - S_e],$$

where $b \equiv (dD/dp)^e, b_1 \equiv (dS/dp)^e$ (the notation $(...)^e$ means that the derivative is taken where $p = p_e$). Since $D_e = S_e$, it follows that

$$p' = c(b - b_1)(p - p_e).$$

Thus in both cases we obtain the first-order linear differential equation with constant coefficients:

$$p' = c(b - b_1)(p - p_e), \tag{13.7}$$

i.e., considering the deviations from equilibrium $\overline{p}(t) = p(t) - p_e$ and observing that $p' = \overline{p}'$ since p_e is constant,

$$\overline{p}' = c(b - b_1)\overline{p}. \tag{13.8}$$

Alternatively we can rewrite Eq. (13.7) as a non-homogeneous equation, namely

$$p' - c(b - b_1)p = -c(b - b_1)p_e. \tag{13.9}$$

Applying the methods explained in Chap. 12, the solution of (13.9) is found to be

$$p(t) = Ae^{c(b-b_1)t} + p_e, \tag{13.10}$$

i.e.,

$$\overline{p}(t) = Ae^{c(b-b_1)t}, \tag{13.11}$$

which coincides with the solution of the homogeneous equation (13.8). The arbitrary constant A is equal to the initial deviation $p(0) - p_e$. The equilibrium point is stable if $p(t)$ tends to p_e as t increases, i.e. if the deviations $\overline{p}(t)$ tend to zero. This requires the term $A\exp[c(b - b_1)t]$ to converge to zero, which in turn is equivalent to the quantity $c(b - b_1)$ being negative. Since $c > 0$, we arrive at the following stability condition

$$b - b_1 < 0. \tag{13.12}$$

Note that inequality (13.12) is the same as inequality (13.2), so that, in this case, the static and the dynamic stability conditions are the same. Let us anticipate that a similar conclusion holds in the case of the Marshallian assumption. Now, we must emphasize that the coincidence between the static and dynamic stability conditions deduced from a given behaviour assumption is not a general rule (see Chap. 19, Sect. 19.1). Given the greater importance of dynamic stability conditions over static stability conditions, whenever a 'disagreement' occurs between them the former must be preferred. More drastically, the study of static stability may be neglected and attention concentrated on dynamic stability only. However, both concepts are given here to demonstrate the conclusion just stated (which will be strengthened by the material contained in Chap. 19, Sect 19.1).

The dynamic formalization of the Marshallian assumptions is the following:

$$q' = g[p_d(q) - p_s(q)], \tag{13.13}$$

where

$$\text{sgn } g[...] = \text{sgn } [...], \qquad g[0] = 0, \qquad g'[0] > 0.$$

The interpretation of Eq. (13.13) is similar to that of Eq. (13.6): it is enough to substitute 'quantity' in the place of 'price' and 'excess demand price' in place of 'excess demand'.

From Eq. (13.13) we obtain—with a similar procedure as before (i.e. either assuming linear functions or making a linear approximation at the equilibrium point)—the following differential equation:

$$q' = k\left(\frac{1}{b} - \frac{1}{b_1}\right)(q - q_e), \tag{13.14}$$

where b and b_1 have the same meaning as before, k is a positive constant and q_e is the equilibrium quantity. The solution of Eq. (13.14) is

$$q(t) = A\exp\left[k\left(\frac{1}{b} - \frac{1}{b_1}\right)t\right] + q_e, \tag{13.15}$$

or

$$\overline{q(t)} = q(t) - q_e = A\exp\left[k\left(\frac{1}{b} - \frac{1}{b_1}\right)t\right]. \tag{13.16}$$

The stability condition is

$$\frac{1}{b} - \frac{1}{b_1} < 0, \tag{13.17}$$

which coincides with (13.5).

It is interesting to compare the Marshallian and Walrasian stability conditions. Conditions (13.17) and (13.12) give the same result both in the 'normal' case and in the 'extreme abnormal' case, whereas they give opposite results in the 'simple abnormal' cases. By 'normal' case we mean the case in which the demand function is a decreasing function and the supply function is an increasing function (i.e. $b < 0$ and so $1/b < 0$; and $b_1 > 0$ and so $1/b_1 > 0$). It is easy to see that in this case both (13.12) and (13.17) are satisfied, so that equilibrium is stable according to both the Walrasian and the Marshallian behaviour assumption.

By 'extreme abnormal' case we mean the case in which demand is an increasing function and supply is a decreasing function: neither (13.12) nor (13.17) are now satisfied, so that equilibrium is unstable according to both behaviour assumptions.

By 'simple abnormal' cases we mean the cases in which one of the two function is abnormal whereas the other is normal. Let us rewrite (13.17) as

$$\frac{b_1 - b}{bb_1} < 0. \tag{13.18}$$

Now, in both cases under consideration the product bb_1 is a positive magnitude (since we have either $b > 0, b_1 > 0$, or $b < 0, b_1 < 0$) so that (13.18) is equivalent to

$$b_1 - b < 0, \tag{13.19}$$

and it is obvious that, if (13.19) holds, then (13.12) cannot hold, and vice versa.

Therefore, in the 'simple abnormal' cases, if equilibrium is stable according to one behaviour assumption, then it is unstable according to the other, and vice versa. This illustrates the already mentioned principle of *relativity of stability conditions.* We may now wonder whether there is any point at all in studying stability, since the results that we obtain may no longer be valid if a different behaviour assumption is adopted. The answer is that, firstly, in the normal cases different behaviour assumptions usually give the same results. Secondly, although the possible behaviour assumptions are in principle many, those that may be deemed *plausible* in the study of the stability of a given equilibrium point are usually a very small number and, moreover, the simple observation of facts often enables us to see that only one is the most *realistic* (i.e., best suited to describe the behaviour of the economic agents in the case under examination). Let us note, finally, that *relativity of results is a principle valid in all economic theory,* in which results depend on the assumptions, so that different assumptions may (though not necessarily must) give rise to different results.

13.2 The neoclassical growth model

After the enormous amount of work (stimulated by the seminal papers of Solow and Swan) carried out in the late 1950's and in the 1960's on formal growth models (for a complete mathematical summing up of that literature see the books by Burmeister and Dobell, 1970, and Wan, 1971), the interest dwindled away. The late 1980's saw a revival in the intellectual interest in the problem of growth, which is still going on in the 1990's. Formal growth theory is now back in fashion (see, for example, Barro and Sala-i-Martin, 1995; Silverberg and Soete, 1994), and the point of departure remains Solow's 1956 model.

Let us consider whether a situation of growth in income with continuous full employment of labour is possible and, if so, whether this situation is stable.

With regard to the first question (existence), it is obvious that—on the assumption of no technical progress—income and the labour force must increase at the same (proportionate) rate of growth. This is called a *steady-state* (or *balanced*) growth situation, namely a situation in which all the relevant variables grow at the same rate, so that their *ratios* are constant. As a point of departure, take the fundamental relation

$$\frac{\Delta Y}{Y} = \frac{s}{k}, \tag{13.20}$$

where $\Delta Y/Y$ (i.e. Y'/Y in continuous terms) is the rate of growth in income and s, k have the usual meaning of average (and marginal) propensity to save and of average (and marginal) capital/output ratio. The relation is easily proved:

$$\frac{s}{k} = \frac{S}{Y}\left(\frac{\Delta K}{\Delta Y}\right)^{-1} = \frac{S}{Y}\frac{\Delta Y}{\Delta K},$$

and since $\Delta K = I$ (ignoring depreciation or working with net magnitude) and in equilibrium $I = S$, it follows that S and $\Delta K = I$ cancel out, so that $s/k = \Delta Y/Y$. In continuous terms $\Delta Y/Y = Y'/Y$, as said above, and $\Delta K = K'$.

On the assumption that population grows at a constant (proportionate) rate, say n, and that the labour force is a constant fraction of population so that it also grows at the same rate n , the condition for the existence of full employment growth is

$$\frac{s}{k} = n. \tag{13.21}$$

In general, if s, k and n are given and independent constants, the above equality would be purely accidental, and we would not expect it to occur. This is, in fact, the conclusion to be drawn from Harrod's model (see Chap. 4, Sect. 4.3, exercise 1). Instead, if k is assumed to be variable, it is in principle possible to satisfy the equality, provided that k can take on the value s/n.

13.2.1 Existence of a growth equilibrium

The neoclassical aggregate model of growth (due to Solow, 1956, and Swan, 1956) in its simplest version assumes that the production function is homogeneous of degree one (with the usual properties of positive but decreasing marginal productivities) and that this function admits of an *unlimited* substitutability between capital and labour. By 'unlimited' substitutability we mean that, to produce any given output, any amount—from zero (excluded or included) to infinity—of capital can be (efficiently) used, obviously using with it the appropriate amount of labour. This assumption is necessary to guarantee that the capital/output ratio can take on any (non negative) value; if it were not so, it might happen that the ratio which is needed to equate s/k to n cannot be attained given the state of technology expressed by the production function.

In this way it is always possible to equate k to s/n. From the static point of view, this means to find the capital/labour ratio which gives rise to an output/labour ratio such that the capital/output ratio is the desired one. Formally, given the production function

$$Y = f(K, L), \tag{13.22}$$

we can use the property of first-degree homogeneity to write it in the 'intensive' form

$$Y/L = f(r, 1), \tag{13.23}$$

where $r \equiv K/L$. It is as well to inform the reader of two notational conventions. The first is that we are using Solow's (1956) symbology; in other presentations the symbol k is often used in the place of r to denote the capital/labour ratio, while we have used k to indicate the capital/output ratio, consistently with the notation already used in Part I. The second is that many treatments use the dot notation instead of the prime notation to denote time derivatives. We have kept with the prime notation for the reasons explained in Chap. 11, Sect. 11.1.

The unlimited substitutability assumptions means that the following conditions

$$\begin{aligned} f(0,1) &= 0, & \lim_{r\to\infty} f(r,1) &= \infty, \\ f'(0,1) &= \infty, & \lim_{r\to\infty} f'(r,1) &= 0, \end{aligned} \tag{13.24}$$

are satisfied. Conditions (13.24) are usually called the Inada conditions (after Inada, 1963), although this denomination is not correct, because they had been previously introduced by Uzawa (1963, p. 108), as Inada himself (1963, p.121) acknowledges.

Now,

$$k = \frac{K}{Y} = \frac{K}{Lf(r,1)} = \frac{r}{f(r,1)}, \tag{13.25}$$

so that the required amount of capital per unit of labour is obtained by solving the equation

$$\frac{s}{n} = \frac{r}{f(r,1)},$$

which can be written as

$$\frac{s}{n} f(r,1) = r, \tag{13.26}$$

or as

$$sf(r,1) = nr. \tag{13.27}$$

A simple graphical interpretation of Eq. (13.27) is given by Fig. 13.2.

If the production function is 'well-behaved' (by which we mean that conditions (13.24) are satisfied), an intersection, and only one, always exists. Point A gives the required amount of capital per unit of labour. It is easy

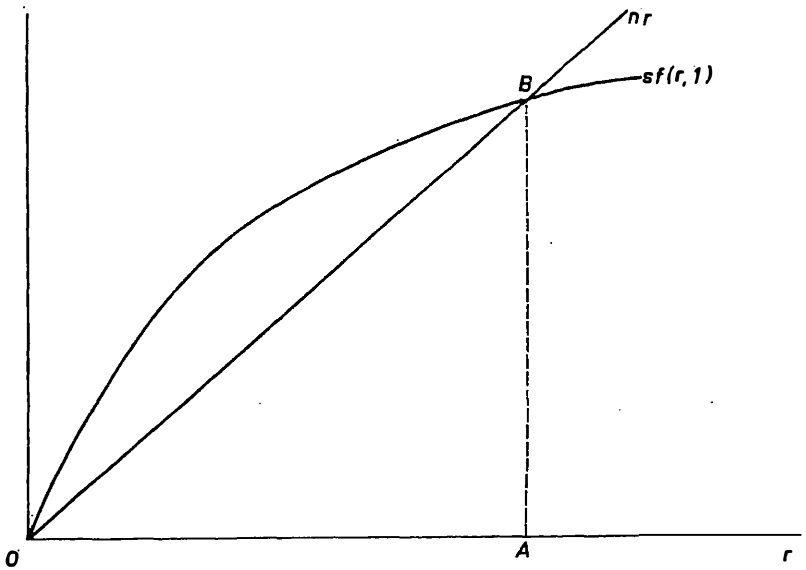


Figure 13.2: The neoclassical aggregate growth model

that this amount is indeed the one which gives rise to the capital/output ratio equal to s/n. By definition,

$$\overline{OA}/\overline{AB} = \frac{K}{L}\left(s\frac{Y}{L}\right)^{-1} = \frac{K}{sY}.$$

Geometrically,

$$\overline{OA}/\overline{AB} = 1/n$$

and equating the two results, it follows that $K/Y = s/n$.

13.2.2 Stability of growth equilibrium

We can now pass to the second question (stability). The basic equation of the model, already introduced, though not in this order, are

$$Y = f(K, L) = Lf(r, 1), \tag{13.28}$$

$$S = sY, \tag{13.29}$$

$$K' = I, \tag{13.30}$$

$$K' = sY, \tag{13.31}$$

$$L = L_0 e^{nt}. \tag{13.32}$$

Equation (13.31) implies that $I = S$ (*ex ante*) and Eq. (13.32) implies that the labour force is fully employed. Consider now the definition of r,

$$r = K/L,$$

which gives

$$K = rL_0 e^{nt}. \tag{13.33}$$

Since Eq.(13.33) is an *identity*, also the derivatives of both members are equal (with equations this would *not* in general be true). Hence we can differentiate both members with respect to time and obtain

$$K' = r' L_0 e^{nt} + rnL_0 e^{nt},$$

from which, given Eqs. (13.31), (13.28) and (13.32), we obtain

$$sL_0 e^{nt} f(r, 1) = r' L_0 e^{nt} + rnL_0 e^{nt}.$$

The term $L_0 e^{nt}$ cancels out, so that we have

$$r' = sf(r, 1) - nr, \tag{13.34}$$

which is the fundamental dynamic equation of the model.

It can easily be seen that, if the production function is 'well behaved', the point of equilibrium is stable. First of all, note that if we impose the condition that r should be in equilibrium, i.e. that $r' = 0$, then from Eq. (13.34) we obtain Eq. (13.27) already discussed above.

Consider now Fig. 13.2: to the left (right) of point A, the function $sf(r, 1)$ lies above (below) the function nr, i.e. respectively $sf(r, 1) - nr \gtrless 0$. Given this, from Eq. (13.34) it follows that to the left (right) of point A, r' is positive (negative), i.e. r increases (decreases) according to whether it is smaller (greater) than its equilibrium value. This proves stability. Similar reasoning would prove the instability of equilibrium in the case in which the function $sf(r, 1)$ were such as to cut the function nr from below.

To examine specific cases, we only have to substitute a specific function in the place of the generic function $f(r, 1)$. For example, given the well known Cobb-Douglas production function, $Y = K^{\alpha}L^{1-\alpha}, 0 < \alpha < 1$, we have $f(r, 1) = r^{\alpha}$, so that Eq. (13.34) becomes

$$r' = sr^{\alpha} - nr. \tag{13.35}$$

Eq. (13.35), although of the first order, is not linear, since r appears also to the α-th power. However, it can be brought to a linear equation by means of a simple transformation[2]. Define a new variable, k, connected to r by the relation

$$k = r^{1-\alpha}. \tag{13.36}$$

[2]This transformation of variables is a general method for solving equations of type (13.35)—they are called Bernoulli equations—and will be explained in detail in Chap. 24.

Note that this mathematical transformation—which is the standard method for solving differential equations of the type under consideration—also has an economic interpretation. In fact, given the Cobb-Douglas production function, it is easy to check that

$$\frac{K}{Y} = \left(\frac{K}{L}\right)^{1-\alpha} = r^{1-\alpha} = k, \tag{13.37}$$

so that the transformation of variables simply amounts to pass from the capital/labour ratio to the capital/output ratio, which after all is the unknown whose determination we started with.

Differentiating Eq. (13.36) with respect to time we have

$$k' = (1-\alpha)r^{-\alpha}r',$$

i.e.

$$\frac{1}{1-\alpha}k' = r^{-\alpha}r'. \tag{13.38}$$

Now, multiply both members of (13.35) by $r^{-\alpha}$:

$$r^{-\alpha}r' = s - nr^{1-\alpha},$$

and use Eqs. (13.36) and (13.38):

$$\frac{1}{1-\alpha}k' = s - nk,$$

i.e.

$$k' + n(1-\alpha)k = s(1-\alpha), \tag{13.39}$$

which is linear in k (and with constant coefficients).

The general solution of the homogeneous part of (13.39) is

$$Ae^{-n(1-\alpha)t},$$

and a particular solution of (13.39), obtained by putting $k' = 0$, is

$$\overline{k}(t) = s/n,$$

so that the general solution of (13.39) is

$$k(t) = Ae^{-n(1-\alpha)t} + s/n. \tag{13.40}$$

Assuming that k has a known value, say k_0, for $t = 0$, from (13.40) we obtain

$$k_0 = A + s/n,$$

therefore

$$A = k_0 - s/n,$$

and so

$$k(t) = (k_0 - s/n)e^{-n(1-\alpha)t} + s/n. \tag{13.41}$$

Since n and $(1-\alpha)$ are positive magnitude, the term $(k_0-s/n)\exp[-n(1-\alpha)t]$ tends to zero as t increases, so that $k(t)$ tends to its equilibrium value, s/n. This is the result we were looking for. The gap between k and its steady-state value vanishes at the constant rate

$$\beta = n(1-\alpha), \tag{13.42}$$

where β is called the coefficient of convergence.

To complete the formal treatment we might wish to go back to the original variable r. From Eq. (13.36) we have the inverse transformation

$$r = k^{1/(1-\alpha)}, \tag{13.43}$$

and by applying it to Eq. (13.41) we get

$$r(t) = \{(r_0^{1-\alpha} - s/n)\exp[-n(1-\alpha)t] + s/n\}^{1/(1-\alpha)}, \tag{13.44}$$

where we have used Eq. (13.36) to set $k_0 = r_0^{1-\alpha}$. Hence $r(t)$ tends to its equilibrium value, $(s/n)^{1/(1-\alpha)}$.

Output Y tends of course to its steady state growth rate n. In fact, from the production function we have $Y = Lr^\alpha$, from which $Y'/Y = n + \alpha r'/r$. Since Eq. (13.35) gives $r'/r = sr^{\alpha-1} - n$, using Eq. (13.44) it is easy to see that r'/r converges to zero, hence Y'/Y converges to n.

13.2.3 Refinements

13.2.3.1 Depreciation and technical progress

The introduction of depreciation and technical progress can be performed without changing the basic conclusions of the model.

As regards depreciation, the usual assumption is that it is of the 'radioactive' type, namely a fraction δ of the capital stock has to be replaced per unit of time. Then Eq. (13.30) becomes

$$K' = I - \delta K,$$

after which the procedure used above yields the following modified fundamental dynamic equation

$$r' = sf(r,1) - (n+\delta)r. \tag{13.45}$$

that replaces Eq. (13.34). The analysis can then proceed as before, with $(n+\delta)$ in the place of n.

The introduction of technical progress in this model can be made in a very simple manner if we assume that it is of the 'disembodied' type, that is, something like manna that falls from heaven on *all* capital goods, old and new. Disembodied technical progress explicates its beneficial influence on output even if the capital stock and the labour force remain the same both quantitatively and qualitatively. It is as if technical progress were simply a way of improving the organisation and operation of inputs without reference to the nature of inputs themselves. Such kind of technical progress is simply expressed by introducing a multiplicative factor in the production function, say $A(t) = e^{\mu t}$, where μ is an exogenously given rate.

It is then easy to see that the rate of growth to which the system tends is no longer n, but (with a Cobb-Douglas production function) $n^* = n + \mu/(1-\alpha)$. In fact, from $Y = e^{\mu t}K^{\alpha}L^{1-\alpha}$ we obtain $Y = K^{\alpha}L^{*1-\alpha}$, where $L^* = L_0 e^{n^* t}$. Then the analysis can proceed as before. However, with a generic production function this is no longer true: as we shall see below, the only technical progress compatible with a steady-state solution is the so-called 'labour augmenting' technical progress.

Labour-augmenting technical progress increases the productivity of labour, so that at any time t the same physical unit of labour produces more than at time $t - \mathrm{d}t$. In this case we only have to measure labour in 'efficiency units' rather than in natural units. For example, if technical progress occurs at a constant proportional rate τ, we can write the equation of the labour force as $L = L_0 e^{\nu t}$, where $\nu = n + \tau$, namely the rate of growth of the labour force *plus* the rate of labour-augmenting technical progress (so that $L = L_0 e^{\nu t}$ is labour measured in efficiency units, also called *effective labour*), and then proceed as before in the analysis of the model. The fundamental dynamic equation, account being taken of depreciation, becomes

$$r' = sf(r,1) - (n + \tau + \delta)r, \tag{13.46}$$

and the steady-state growth rate will be $n + \tau + \delta$, that—in the case of a Cobb-Douglas production function—will be approached according to the convergence coefficient

$$\beta = (1-\alpha)(n + \tau + \delta), \tag{13.47}$$

which replaces the one defined earlier in Eq. (13.42).

Let us now come to technical progress in general, and write the production function as

$$Y(t) = A(t)F[B(t)K(t), C(t)L(t)], \tag{13.48}$$

where $A(t)$ is output-augmenting (also called Hicks-neutral), $B(t)$ capital-augmenting (also called Solow-neutral), and $C(t)$ labour-augmenting (also called Harrod-neutral) technical progress. It is then easy to see that, letting $y \equiv Y/L$,

$$y = f(r,1), \quad y = A(t)f(r,1), \quad y = f(B(t)r, 1) \tag{13.49}$$

respectively for labour-augmenting (with labour measured in efficiency units), output-augmenting and capital-augmenting technical progress. It follows that when r is constant, y is constant only in the case of labour-augmenting technical progress. Hence there can be a steady-state (both r and y constant) only when technical progress is labour-augmenting.

The advantage of the Cobb-Douglas functional form is that both output-augmenting and capital-augmenting technical progress can be expressed as labour-augmenting simply by shifting terms, as we have shown above in the output-augmenting case.

More complicated types of technical progress can be considered, for example the case of technical progress embodied in new capital goods, which gives rise to 'vintage' models (Solow, 1960). Yet these do not depart from the assumption that technical progress is *exogenous* to the model. In the growth literature of the 1950's and 1960's there were however a few attempts at endogenizing technical progress. Among these Kaldor's 'technical progress function' (Kaldor, 1957, Kaldor and Mirrlees, 1961), where productivity increases with the rate of investment I/K, and Arrow's 'learning by doing' (Arrow, 1962), according to which labour requirements per unit of output on new machines fall over time as experience permits a better design of machines, are worth mentioning.

13.2.3.2 Golden rule

An interesting conclusion that has been drawn from the neoclassical model is the so-called '*golden rule*' of capital accumulation (Phelps, 1966), which is the answer to the following question: given that the system is on its equilibrium path and assuming that the only goal of society is to maximize consumption per unit of labour, how is the goal to be reached? The analysis is very simple. On the equilibrium path, all variables grow at the same proportionate rate n, so that

$$K' = nK.$$

Consumption equals output minus investment, that is

$$C = f(K, L) - nK,$$

and, given that f is homogeneous of degree one

$$C/L = f(r,1) - nr.$$

Now, C/L is at a maximum when

$$\frac{\mathrm{d}[f(r,1) - nr]}{\mathrm{d}r} = \frac{\mathrm{d}f(r,1)}{\mathrm{d}r} - n = 0. \tag{13.50}$$

The second derivative $\mathrm{d}^2 f(r,1)/\mathrm{d}r^2$, is negative since f is assumed to be a well-behaved function, so that the stationary point given by (13.50) is indeed

Hence, by adding and subtracting 1 in the expression in square brackets and rearranging terms,

$$\frac{r'}{r_e} = (n+\tau+\delta)\left[\frac{f(r,1)-f(r_e,1)}{f(r_e,1)} - \left(\frac{r-r_e}{r_e}\right)\right]. \tag{13.56}$$

Since we are considering a linear approximation, we expand $f(r,1)$ in Taylor's series at the equilibrium point r_e and neglect all higher order terms, which gives

$$f(r,1)-f(r_e,1) \simeq f'(r_e,1)(r-r_e) = r_e f'(r_e,1)\left(\frac{r-r_e}{r_e}\right). \tag{13.57}$$

Evaluating r'/r_e in the neighbourhood of the equilibrium point, so that $r'/r_e = (d\ln r/dt)_{r=r_e}$, and substituting Eq. (13.57) into Eq. (13.56), we obtain

$$\begin{aligned}\left(\frac{d\ln r}{dt}\right)_{r=r_e} &= (n+\tau+\delta)\left[\frac{r_e f'(r_e,1)}{f(r_e,1)}\left(\frac{r-r_e}{r_e}\right) - \left(\frac{r-r_e}{r_e}\right)\right] \\ &= (n+\tau+\delta)\left[\frac{r_e f'(r_e,1)-f(r_e,1)}{f(r_e,1)}\left(\frac{r-r_e}{r_e}\right)\right].\end{aligned} \tag{13.58}$$

Since $\ln(1+z) \simeq z$ for small z, by setting $z=(r-r_e)/r_e$ and observing that $1+z = 1+(r-r_e)/r_e = r/r_e$, we finally obtain

$$\left(\frac{d\ln r}{dt}\right)_{r=r_e} = -\beta \ln\left(\frac{r}{r_e}\right), \tag{13.59}$$

where

$$\beta \equiv \frac{f(r_e,1)-r_e f'(r_e,1)}{f(r_e,1)}(n+\tau+\delta) \tag{13.60}$$

has the same interpretation as the convergence coefficient defined in Eq. (13.47). Note that, since factors are rewarded according to their marginal productivity, the fraction on the r.h.s. of Eq. (13.60) is one minus the share of capital in total output, namely the share of labour.

It can easily be checked that, when the production function is Cobb-Douglas, where $f(r,1)=r^\alpha$, Eq. (13.60) yields Eq. (13.47).

Equation (13.59) is a differential equation in $\ln(r/r_e)$, that has the solution

$$\ln\frac{r}{r_e} = \ln\left(\frac{r_0}{r_e}\right)e^{-\beta t} \tag{13.61}$$

or

$$\ln r = (1-e^{-\beta t})\ln r_e + e^{-\beta t}\ln r_0. \tag{13.62}$$

Equation (13.62) can be used to calculate the adjustment time according to formula (13.51). It should however be stressed that these calculations are valid only for small deviations from the equilibrium.

13.2.4.2 β-convergence, σ-convergence, and all that

The renewed theoretical interest in formal growth theory focuses on two related aspects: the *convergence hypothesis* and *endogenous growth,* respectively examined in this and the following paragraph.

By convergence hypothesis the new growth theorists mean the hypothesis that poor countries (i.e., countries with a low amount of capital per head and output per head) grow faster than rich countries, *other things being equal.* Hence, in the long run, growth rates of poor and rich countries should tend to converge. The negative relationship between the growth rate of income per head and the initial level of income per head has been labelled β-convergence by Barro and Sala-i-Martin (1992), to distinguish it from σ-convergence. This latter means that the dispersion (as measured, for example, by the variance) of per capita income across groups of economies tends to fall over time. It might seem that β-convergence implies σ-convergence, and this would be true in a deterministic context. But in a stochastic context, which is the one that we find in real life, the presence of random shocks on the growth rates of the various countries tends to increase the variance. Hence β-convergence does not imply σ-convergence.

Let us now go back to β-convergence and stress the importance of the *ceteris paribus* clause, that serves to dissipate a possible misunderstanding.

It is a general rule that, when we perform any analysis in *comparative dynamics* (which is the dynamic equivalent to the well known comparative statics analysis: see Chap. 20, Sect. 20.6), what we are really making is a comparison between the dynamic behaviour of two economies that are equal in all respects except for the one under consideration or, which is the same thing, we are comparing the dynamic behaviour of the same economy with different values of the item being considered. This is exactly what we are doing when we examine the convergence hypothesis, where the element of difference is the distance from the steady state.

Thus, the neoclassical model does not make the loose, and possibly wrong, statement that poor countries tend to grow faster than rich countries (this has been called the *absolute convergence* hypothesis). It makes the precise (and certainly correct) comparative dynamic statement that the growth rate of an economy is directly related to the distance *of that economy* from its steady state. Hence the poorer an economy is, the faster *this same economy* tends to grow. To extend this statement to inter-country comparisons one must postulate that all countries are equal in all respects except for the distance from the (same) steady state. In fact, only if poor and rich countries converge to the *same* steady state can we say that poor countries, that are more distant from the steady state, grow faster than rich countries. The concept of convergence conditional on the steady state (which is the precise concept of convergence) has been called *conditional convergence,* and this we are going to examine.

Let us consider output per head $y \equiv Y/L$. From the production function we have

$$y = K^{\alpha}L^{-\alpha} = r^{\alpha}. \tag{13.63}$$

Differentiating with respect to time we obtain $y' = \alpha r^{\alpha-1}r'$, and so

$$y'/y = \alpha r^{-1}r',$$

from which, by using Eq. (13.35),

$$y'/y = \alpha s r^{\alpha-1} - \alpha n, \tag{13.64}$$

where r is given by Eq. (13.44). Since r converges to its steady state value $(s/n)^{1/(1-\alpha)}$, it is easy to see that y'/y tends to zero as $t \to \infty$, which we already knew from our previous result that steady-state output grows at the same rate as the labour force.

To show that y'/y is directly related to the distance of output per head from its steady state value, we first observe that we are considering situations where the initial value of output per head is below its steady state value. Hence, given the monotonic movement of the variables, we can define the distance as the positive magnitude

$$D(t) = y_e - y(t), \quad y(t) = y_e - D(t). \tag{13.65}$$

From Eq. (13.63) we have $r = y^{1/\alpha}$, and by substituting this into Eq. (13.64), account being taken of Eq. (13.65), we obtain

$$y'/y = \alpha s\,[y_e - D(t)]^{(\alpha-1)/\alpha} - \alpha n. \tag{13.66}$$

It is now easy to see that

$$\frac{\partial(y'/y)}{\partial D(t)} = -(\alpha-1)s\,[y_e - D(t)]^{-1/\alpha} > 0, \tag{13.67}$$

which proves (conditional) convergence.

As we have seen, an equivalent way of stating conditional convergence is that the per capita growth rate is inversely related to the starting level of output per head, given of course the same parameters of the model. This derives from the observation that, at any time t, the lower y_0, the greater the distance. In fact, by using Eq. (13.44) we get

$$D(t) = y_e - r^{\alpha} = (s/n)^{\alpha/(1-\alpha)} - \{(y_0^{(1-\alpha)/\alpha} - s/n)\exp[-n(1-\alpha)t] + s/n\}^{\alpha/(1-\alpha)}. \tag{13.68}$$

It is then easy to see that, given the values of the parameters and given $y_0^{(1-\alpha)/\alpha} - s/n < 0$ (since we are starting from below the steady state), the distance at any time t is an inverse function of y_0. More formally,

$$\frac{\partial D(t)}{\partial y_0} = -\left\{(y_0^{(1-\alpha)/\alpha} - s/n)\exp[-n(1-\alpha)t] + s/n\right\}^{\frac{2\alpha-1}{1-\alpha}} y_0^{\frac{1-2\alpha}{\alpha}} e^{-n(1-\alpha)t} < 0. \tag{13.69}$$

Therefore

$$\frac{\partial(y'/y)}{\partial y_0} = \frac{\partial(y'/y)}{\partial D(t)}\frac{\partial D(t)}{\partial y_0} < 0. \tag{13.70}$$

A third way of stating conditional convergence is that economies with lower capital per person tend to grow faster in per capita terms, other things equal, namely y'/y is inversely related to r. This can be shown by computing, from Eq. (13.64), the partial derivative

$$\frac{\partial(y'/y)}{\partial r} = (\alpha-1)\alpha s r^{\alpha-2} < 0. \tag{13.71}$$

Let us note, finally, that the same results can be obtained with the more general model (see Sect. 13.2.3.1) in which depreciation and technical progress are present.

13.2.4.3 Endogenous growth

Growth in the basic neoclassical model is exogenous: the steady state path, in fact, depends on exogenous factors such as the rate of growth of the labour force and technical progress. Both are exogenous: the labour force grows according to exogenous demographic factors, and technical progress is no more than an exogenous time trend.

The theory of endogenous growth stresses the *endogenous* determination of technical progress, which actually means an endogenous determination of the main source of growth (hence the name of endogenous growth theory). The basic ideas were already present in neoclassical growth theory (see above, Sect. 13.2.3.1), but in endogenous growth theory (and in current growth theory in general: see Chap. 19, Sect 19.2, for a neoclassical model in which human capital plays a crucial role) they are at the centre of the stage.

Another point emphasized by endogenous growth theory is the absence of decreasing returns to capital. This implies that there is normally *no* convergence, and in this sense endogenous growth is closely (but negatively) related to the convergence hypothesis. It is however possible to reconcile endogenous growth and convergence, as we shall see in Chap. 22, Sect. 22.2.2.

For a general treatment of endogenous growth see Barro and Sala-i-Martin (1995), Romer (1994), and Solow (1992).

13.3 Exercises

1. Suppose that demand and supply are isoelastic functions, $D = ap^b, S = a_1p^{b_1}, a > 0, a_1 > 0$. Furthermore, assume that the adjustment mechanism is described either by the differential equation $d\log p/dt = c\log[D(p)/S(p)], c > 0$, or by the differential equation $d\log q/dt = k\log[p_d(t)/p_s(t)], k > 0$. Express the stability conditions in terms of elasticities.

2. The following model is the continuous-time equivalent to the cobweb model examined in discrete time in Chap. 4, Sect. 4.1:

$$\begin{aligned} D(t+\mathrm{d}t) &= a + bp(t+\mathrm{d}t), \\ S(t+\mathrm{d}t) &= a_1 + b_1 p(t), \\ D(t+\mathrm{d}t) &= S(t+\mathrm{d}t), \end{aligned}$$

where it is assumed that suppliers expect that in the next instant ($t+\mathrm{d}t$) the price will be equal to the current value. Examine the stability of the static equilibrium point taking $\mathrm{d}t = 1$ (Hint: by elementary calculus, $p(t+\mathrm{d}t) = p(t)+\mathrm{d}p(t)/\mathrm{d}t = p(t) + p'$).

3. In the model considered in the previous exercise, let $S(t+\mathrm{d}t) = a_1+b_1\hat{p}(t)$, where $\hat{p}(t)$ is the price that, at time t, producers expect will hold in the next instant. Then assume $\hat{p}(t) = p(t) + c\,[\mathrm{d}p(t)/\mathrm{d}t)]$, where c is a coefficient of expectations that can be either positive (extrapolative expectations) or negative (regressive expectations)—the case $c = 0$ has already been examined in exercise 2. Study the stability in the various cases.

4. In the model $D(t) = a+bp(t), S(t) = a_1+b_1\hat{p}(t)$, suppose that expectations are formed adaptively, i.e. $\hat{p}'(t) = \beta\,[p(t) - \hat{p}(t)]\,, \beta > 0$. Show that, at the equilibrium point, $p = \hat{p} = p_e$, and examine the conditions under which the paths of $p(t)$ and $\hat{p}(t)$ converge to p_e.

5. In a macroeconomic long-run model (Domar, 1946), investment has the *dual* nature of being a component of aggregate demand and of increasing the productive capacity of the economy. Suppose that we want to keep productive capacity fully employed. Hence, there must be an increase in aggregate demand to match the increase in (potential) output brought about by the increase in productive capacity. Assume that output (which equals income) actually increases by its potential value. Now, given that the propensity to consume is smaller than one, the increase in output is only partly absorbed by an increase in consumption, so that an increase in investment (in its aspect of component of aggregate demand) is called for to absorb the remaining part. But the increase in investment further increases productive capacity, hence investment must further increase, and so on and so forth. We are required to find the growth path of investment that maintains full capacity utilization, and the corresponding growth paths of the other variables.
For this purpose, call σ the (constant) 'potential social average productivity of investment', so that—assuming no depreciation—$P' = \sigma I$ is the increase in potential output brought about by investment (σ can also be interpreted as the reciprocal of the capital/output ratio, $\sigma = 1/k$). Let Y' be the increase in output, equal to the increase in aggregate demand, $C' + I'$. Full employment of productive capacity requires $Y' = P'$. Assume then the simple consumption function $C = (1 - s)Y$. (Note: the final formula of Domar's model coincides with that found in Harrod's model—see Chap. 4, exercise 1—and so their models are often referred to as the Harrod-Domar growth model, although they started from different problems).

6. In the neoclassical aggregate growth model assume that the production function is $Y = K^{1/3}L^{2/3}$, and $n = 0.025$, $s = 0.10$, $r_0 = 54.812$. Calculate the time required for 90% of the initial deviation of $r(t), k(t), y(t)$ from their respective equilibrium values to be eliminated.

7. With the same production function of exercise 6, introduce labour augmenting technical progress at the rate $\tau = 0.02$, depreciation at the rate $\delta = 0.05$, and assume $n = 0.01$. Calculate the time required for (a) one-half, and (b) seventy-five per cent of the initial deviation to be eliminated.

8. Suppose that the supply of labour is a function of the real wage rate and time, $L(t) = L_0 e^{nt}(w/p)^h, h > 0$. For simplicity assume that the price level is constant.

(a) Show that the fundamental dynamic equation of the neoclassical growth model is now $r' = sf(r,1) - nr - hr(w'/w)$.

(b) In a competitive economy the real wage rate is determined by the marginal productivity condition, $\partial f/\partial L = w/p$. Assume a Cobb-Douglas production function, derive the final form of the fundamental dynamic equation, and compare the results with those of the basic model.

9. Show that the basic neoclassical growth model can also be reduced to the differential equation $K' = sf(K, L_0e^{nt})$, and work out its solution assuming a Cobb-Douglas production function.

13.4 References

Arrow, K.J., 1962, The Economic Implications of Learning by Doing.
Barro, R.J. and X. Sala-i-Martin, 1992, Convergence.
Barro, R.J. and X. Sala-i-Martin, 1995, *Economic Growth.*
Beckmann, M.J. and J.P. Wallace, 1967, Marshallian versus Walrasian Stability.
Burmeister, E. and A.R. Dobell, 1970, *Mathematical Theories of Economic Growth.*
Domar, E.D., 1946, Capital Expansion, Rate of Growth, and Employment.
Inada, K.-I., 1963, On a Two-Sector Model of Economic Growth: Comments and a Generalization.

Kaldor, N., 1957, A Model of Economic Growth.
Kaldor, N. and J.A. Mirrlees, 1961, A New Model of Economic Growth.
Marshall, A., 1920, *Principles of Economics,* Book V, Chap. III, § 6.
Mankiw, N.G., D. Romer and D.N. Weil, 1992, A Contribution to the Empirics of Economic Growth.
Phelps, E., 1966, *Golden Rules of Economic Growth.*
Romer, P.M., 1994, The Origins of Endogenous Growth.
Samuelson, P.A., 1947, *Foundations of Economic Analysis,* Chap. IX, pp. 260-264.
Sato, R., 1963, Fiscal Policy in a Neo-Classical Growth Model: An Analysis of Time Required for Equilibrating Adjustment.
Sato, R., 1980, Adjustment Time and Economic Growth Revisited.
Silverberg, G. and L. Soete (eds.), 1994, *The Economics of Growth and Technical Change.*
Solow, R.M., 1956, A Contribution to the Theory of Economic Growth.
Solow, R.M., 1960, Investment and Technical Progress.
Solow, R.M., 1992, *Siena Lectures on Endogenous Growth Theory.*
Swan, T., 1956, Economic Growth and Capital Accumulation.
Uzawa, H., 1963, On a Two-sector Model of Economic Growth, II.
Walras, L., 1954 (1926), *Elements of Pure Economics,* lesson 6, §§ 60, 61; lesson 7, §§ 66, 67.
Wan, H.Y., Jr., 1971, *Economic Growth.*

Chapter 14

Second-order Differential Equations

The general form of second order differential equation is

$$a_0 y'' + a_1 y' + a_2 y = g(t), \tag{14.1}$$

where a_0, a_1, a_2 are given constants and $g(t)$ is a known function. Of course, the coefficient a_0 must be different from zero since the equation is of the second order; a_1 or a_2 or both may be zero.

14.1 Solution of the homogeneous equation

Let us begin by studying the homogeneous equation

$$a_0 y'' + a_1 y' + a_2 y = 0. \tag{14.2}$$

Since $a_0 \neq 0$, we can divide through by a_0 and write the equation as

$$y'' + b_1 y' + b_2 y = 0, \tag{14.3}$$

where

$$b_1 \equiv a_1/a_0, \quad b_2 \equiv a_2/a_0. \tag{14.4}$$

What might be the form of the unknown function $y(t)$ that satisfies the equation is not so clear as in the first-order equations, at least at first sight. However, analogy suggests that this form might be similar to that found in first-order equations. Let us then try a function of the type $e^{\lambda t}$, where λ is a constant to be determined in terms of the coefficients of the equation. Substituting in (14.3) we have

$$\lambda^2 e^{\lambda t} + b_1 \lambda e^{\lambda t} + b_2 e^{\lambda t} = e^{\lambda t}(\lambda^2 + b_1 \lambda + b_2) = 0. \tag{14.5}$$

If $e^{\lambda t}$ has to be a solution, Eq. (14.5) must be satisfied for any value of t, and this is possible if, and only if,

$$\lambda^2 + b_1\lambda + b_2 = 0. \tag{14.6}$$

Eq. (14.6) is called the *characteristic* (or *auxiliary*) equation of (14.3). Note that here too we have reduced the solution of a functional equation to the solution of an algebraic equation.

The two roots of Eq. (14.6) are given by the well-known formula

$$\lambda_1, \lambda_2 = \frac{1}{2}[-b_1 \pm (b_1^2 - 4b_2)^{1/2}]. \tag{14.7}$$

The nature of the solution of Eq. (14.3) depends on the nature of these roots, which we are going to examine in some detail, according to the sign of the discriminant $\Delta \equiv b_1^2 - 4b_2$. Three cases are possible.

14.1.1 Positive discriminant ($\Delta > 0$)

The roots λ_1, λ_2 are real and distinct. Then both exp $(\lambda_1 t)$ and exp $(\lambda_2 t)$ satisfy Eq. (14.3) so that, according to general principles, their linear combination with two arbitrary constants

$$y(t) = A_1 \exp(\lambda_1 t) + A_2 \exp(\lambda_2 t) \tag{14.8}$$

is the general solution of Eq. (14.3). The time path of $y(t)$ for $t \to +\infty$ is monotonic. The stability of this movement depends on the *signs* of the roots[1]. Descartes' theorem[2] enables us to reach the following general proposition given that $\Delta > 0$:

(1) If b_1 and b_2 are both positive, the roots are both negative, since we have two continuations in the signs of the coefficients of Eq. (14.6). From

[1]Note again that, contrary to the result holding for difference equation, here the absolute value of the roots is immaterial as far as stability is concerned. This is a general property, already mentioned in Chap. 12, and depends on the fact that in the solution the λ's appear in one case in the form λ^t, and in the other in the form $e^{\lambda t}$.

[2]The sign of the roots can be ascertained by **Descartes' theorem**: *In any algebraic equation, complete or incomplete, the number of positive roots cannot exceed the number of changes of signs of the coefficients, and in any complete equation the number of negative roots cannot exceed the number of continuations in the signs of the coefficients.* In our case, given of course that $\Delta > 0$, there are the following possibilities:

+	+	+	two negative roots,
+	+	−	one negative and one positive root (the negative root being greater in absolute value),
+	−	+	two positive roots,
+	−	−	one negative and one positive root (the positive root being greater in absolute value),
+	0	−	one positive and one negative root (with the same absolute value).

(14.8) it follows that $y(t)$ tends to zero as t increases, i.e. the movement is convergent.

(2) If b_1 and b_2 are both negative or if $b_1 > 0$, $b_2 < 0$, the signs of the coefficients show one variation and one continuation, so that one root is positive and the other is negative. In (14.8), the term containing the negative root converges to, and the term containing the positive root diverges from, zero as t increases. The overall movement is obviously divergent.

(3) If $b_1 < 0$, $b_2 > 0$ the signs of the coefficients show two variations, so that both roots are positive. The movement is clearly divergent, towards $\pm\infty$ according to whether the sign of the arbitrary constant (to be determined by means of appropriate condition, see below) multiplying the term containing the greater root is positive or negative.

Two particular cases may arise, when either b_1 or b_2 are zero (both cannot be zero here since $\Delta > 0$). When $b_1 = 0, b_2$ must be negative since $\Delta > 0$, and from (14.7) it is easy to see that the two roots are equal in absolute value and opposite in sign, so that a similar conclusion as in (2) holds. When $b_1 \neq 0, b_2 = 0$, from (14.7) it is easy to see that one root is zero and the other is equal to $-b_1$. The solution is

$$y(t) = A_1 + A_2 \exp(-b_1 t),$$

and so the movement will diverge if $b_1 < 0$ and converge if $b_1 > 0$: note, however, that in the second case it will not converge to zero but to the constant A_1.

14.1.2 Null discriminant ($\Delta = 0$)

The roots λ_1, λ_2 are real and equal : $\lambda_1 = \lambda_2 = -\frac{1}{2}b_1$, call it λ^*. Now $\exp(\lambda^* t)$ is one solution of Eq. (14.3), and we must find a second one to be able to write down the general solution, according to general principles. Let us prove that t exp($\lambda^* t$) is another function which satisfies Eq. (14.3). Substituting in (14.3) we have

$$[2\lambda^* \exp(\lambda^* t) + t\lambda^{*2} \exp(\lambda^* t)] + b_1[\exp(\lambda^* t) + t\lambda^* \exp(\lambda^* t)] + b_2 t \exp(\lambda^* t) = 0,$$

from which

$$(2\lambda^* + b_1)\exp(\lambda^* t) + (\lambda^{*2} + b_1\lambda^* + b_2)t\exp(\lambda^* t) = 0. \tag{14.9}$$

If t exp$(\lambda^* t)$ is a solution, then Eq. (14.9) must be satisfied for any t so that, since $\exp(\lambda^* t) \neq 0$, the expressions $2\lambda^* + b_1$ and $\lambda^{*2} + b_1\lambda^* + b_2$ must be both zero. The former is zero because $\lambda^* = -\frac{1}{2}b_1$, and the latter is zero because λ^* is a root of the characteristic equation. Thus $t\exp(\lambda^* t)$ is a solution of Eq.(14.3), and the general solution is

$$y(t) = A_1 \exp(\lambda^* t) + A_2 t \exp(\lambda^* t) = (A_1 + A_2 t)\exp(\lambda^* t), \tag{14.10}$$

where A_1 and A_2 are arbitrary constants. Note that (14.10) also holds in the particular case $\lambda^* = 0$, which gives $y(t) = A_1 + A_2 t$. Since $\lambda^* = 0$ implies $b_1 = b_2 = 0$, the differential equation is $y'' = 0$, from which, integrating twice, we actually obtain $y(t) = A_1 + A_2 t$ without any need to apply the method so far expounded. However, the fact that we can obtain the solution also from (14.10), shows that the method also works in particular cases.

The function $A_1 + A_2 t$ obviously diverges, so that the overall movement will be divergent if $\exp(\lambda^* t)$ also diverges ($\lambda^* > 0$, i.e. $b_1 < 0$). What happens when $\lambda^* < 0$? In this case $\exp(\lambda^* t)$ tends to zero and, since it tends to zero more rapidly than $A_1 + A_2 t$ tends to infinity, the overall movement will be convergent [3].

14.1.3 Negative discriminant ($\Delta < 0$)

The roots are two complex conjugate numbers, say $\alpha \pm \mathrm{i}\theta$, where $\mathrm{i} = \sqrt{-1}$ is the imaginary unit and

$$\alpha = -\frac{1}{2}b_1, \qquad \theta = \frac{1}{2}(\left|b_1^2 - 4b_2\right|)^{1/2}. \tag{14.11}$$

Then we have

$$\begin{aligned} y(t) &= A' \exp[(\alpha + \mathrm{i}\theta)t] + A'' \exp[(\alpha - \mathrm{i}\theta)t] \\ &= A' \exp[(\alpha t + \mathrm{i}\theta t)] + A'' \exp[(\alpha t - \mathrm{i}\theta t)], \end{aligned} \tag{14.12}$$

where A' and A'' are two arbitrary constants that we may take as two arbitrary complex conjugate numbers.

Let us now recall some elementary results in the theory of functions of a complex variable, that will enable us to write the solution in a more suitable form.

In the real domain, the exponential function e^x (x a real number) can be defined as the sum of the power series

$$1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$

[3]More formally,

$$\begin{aligned} &\lim_{t\to+\infty}[A_1 \exp(\lambda^* t) + A_2 t \exp(\lambda^* t)] \\ &= A_1 \lim_{t\to+\infty} \exp(\lambda^* t) + A_2 \lim_{t\to+\infty} t \exp(\lambda^* t) \\ &= 0 + A_2 \lim_{t\to+\infty} t \exp(\lambda^* t) = A_2 \lim_{t\to+\infty} \frac{t}{\exp(-\lambda^* t)}. \end{aligned}$$

Applying L'Hôpital's theorem,

$$\lim_{t\to+\infty} \frac{t}{\exp(-\lambda^* t)} = \lim_{t\to+\infty} [-\lambda^* \exp(-\lambda^* t)]^{-1} = 0.$$

In the same way we can define e^z (z a complex number) as the sum of the series of complex terms:

$$e^z = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \dots \tag{14.13}$$

Since the series converges for all values of z , it defines a function which is analytic (i.e. one valued and differentiable) in the whole z-plane. It can also be proved by multiplication of series, that the real number e with a complex exponent obeys the formal laws of indices of elementary algebra (so that, for example, $e^{a\pm \mathrm{i}x} = e^a e^{\pm \mathrm{i}x}$). When z is pure imaginary, say $z = \pm \mathrm{i}x$, applying (14.13) we have

$$e^{\pm \mathrm{i}x} = 1 + (\pm \mathrm{i}x) + \frac{(\pm \mathrm{i}x)^2}{2!} + \dots \tag{14.14}$$

If we recall that $\mathrm{i}^2 = -1$, whence $\mathrm{i}^3 = -\mathrm{i}$, $\mathrm{i}^4 = 1$, $\mathrm{i}^5 = \mathrm{i}$, etc. from (14.14) we have, on separating the real and the imaginary terms of the series,

$$\begin{aligned} e^{\pm \mathrm{i}x} = &(1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + - \dots) \\ &\pm \mathrm{i}(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + - \dots), \end{aligned} \tag{14.15}$$

so that

$$e^{\pm \mathrm{i}x} = \cos x \pm \mathrm{i} \sin x, \tag{14.16}$$

since the two power series on the right-hand side of (14.15) define, respectively, the cosine and the sine of the real variable x. Thus we have the *exponential form of* complex numbers

$$e^{a\pm \mathrm{i}x} = e^a(\cos x \pm \mathrm{i} \sin x).$$

Let us now go back to Eq. (14.12), and rewrite it as

$$y(t) = e^{\alpha t}(A' e^{\mathrm{i}\theta t} + A'' e^{-\mathrm{i}\theta t}). \tag{14.17}$$

Putting $x = \theta t$ we can use the transformation

$$e^{\pm \mathrm{i}\theta t} = \cos \theta t \pm \mathrm{i} \sin \theta t \tag{14.18}$$

to obtain

$$\begin{aligned} y(t) &= e^{\alpha t}[A'(\cos \theta t + \mathrm{i} \sin \theta t) + A''(\cos \theta t - \mathrm{i} \sin \theta t)] \\ &= e^{\alpha t}[(A' + A'') \cos \theta t + (A' - A'')\mathrm{i} \sin \theta t]. \end{aligned} \tag{14.19}$$

Now, since A' and A'' are complex conjugate numbers, say $a \pm \mathrm{i}b$ (where a and b are arbitrary real numbers), it follows that $A' + A'' = 2a$, a real number (call it A_1) and that $(A' - A'')$i$=(2ib)$i$=-2b$, a real number too (call it A_2). The solution becomes

$$y(t) = e^{\alpha t}[A_1 \cos\theta t + A_2 \sin\theta t]. \tag{14.20}$$

An alternative way of writing the solution is

$$y(t) = Ae^{\alpha t}\cos(\theta t - \epsilon), \tag{14.21}$$

where the new arbitrary constants A, ϵ are expressed in terms of A_1 and A_2 by the transformation

$$\begin{aligned} A\cos\epsilon &= A_1, \\ A\sin\epsilon &= A_2. \end{aligned} \tag{14.22}$$

Substituting (14.22) into (14.20) and using the fact that

$$\cos\epsilon\cos\theta t + \sin\epsilon\sin\theta t = \cos(\theta t - \epsilon),$$

we obtain (14.21). This second form is perhaps easier to interpret since it involves only one trigonometric function instead than two; the form (14.20) is more suitable for the determination of the values of the two arbitrary constants.

In any case the resulting movement is a trigonometric oscillation, whose period is $2\pi/\theta$ and whose amplitude is increasing, constant, or decreasing, if, respectively, $\alpha \gtreqless 0$, i.e. if the real part of the complex roots is positive, zero, or negative. Thus, since $\alpha = -\frac{1}{2}b_1$, the stability condition (damped oscillations) is that $b_1 > 0$.

14.1.4 Stability conditions

We have now completed the study of all the cases that may arise in a homogeneous equation of the second order. From the point of view of stability (i.e. convergence towards zero of the general solution) the general condition is that the real roots are negative and that the complex roots have a negative real part. In fact, since we are interested in the convergence towards zero, we must also exclude the case of a zero real root (convergence to a constant, see at the end of Sect. 14.1.1), and the case of complex roots with a zero real part (constant-amplitude oscillation).

If we want to check the stability without solving the characteristic equation, it follows from the results obtained in the various cases that we only have to check that both b_1 and b_2 are positive. In other words, the necessary and sufficient conditions for the roots of (14.6) to be negative if real, and have negative real parts if complex, are

$$\begin{aligned} b_1 &> 0, \\ b_2 &> 0. \end{aligned} \tag{14.23}$$

The following table provides a classification of all possible cases and subcases.

	b_1	b_2	Δ	**Kind of movement for $t \to \infty$**
			+	Monotonic, convergent to zero
+	+	+	0	Monotonic, convergent to zero
			−	Oscillating, damped
			+	Monotonic, divergent
+	−	+	0	Monotonic, divergent
			−	Oscillating, divergent
+	+	−	+	Monotonic, divergent
+	−	−	+	Monotonic, divergent
+	0	+	−	Oscillating, constant amplitude
+	0	−	+	Monotonic, divergent
+	+	0	+	Monotonic, convergent to an arbitrary constant
+	−	0	+	Monotonic, divergent
+	0	0	0	Monotonic (linear and not exponential as in all previous cases), divergent

14.2 Particular solution of the non-homogeneous equation

Let us now examine the non-homogeneous equation. According to general principles we only have to find a particular solution of this equation and add it to the general solution of the corresponding homogeneous equation. We shall exemplify the general method of undetermined coefficients (see Chap. 11) in the case in which $g(t)$ is constant, say F. In other cases the student may proceed along the same lines as in the examples expounded with reference to first-order equations (see Chap. 12).

As a particular solution, try $\bar{y}(t) = B$, where B is an undetermined constant. Substituting in (14.1) we have $a_2 B = F$, so that

$$\bar{y}(t) = F/a_2 \tag{14.24}$$

is a particular solution. If $a_2 = 0$, try $\bar{y}(t) = Bt$. Substituting in (14.1) we have $a_1 B = F$, so that

$$\bar{y}(t) = \frac{F}{a_1}t \tag{14.25}$$

is a particular solution. If also $a_1 = 0$, try $\bar{y}(t) = Bt^2$ and substitute in (14.1), obtaining $2a_0 B = F$, so that

$$\bar{y}(t) = \frac{F}{2a_0}t^2 \tag{14.26}$$

is a particular solution (remember that $a_0 \neq 0$).

The particular solution of the non-homogeneous equation may usually be interpreted, as we shall see in the economic applications, as the (stationary or moving) equilibrium of the variable $y(t)$.

14.2.1 Variation of parameters

The principles of this method have already been illustrated in Chap. 12, Sect. 12.2.6. Let us recall that this is a general method of solving a differential equation by considering the arbitrary constants that appear in the known solution of a simpler equation, as variable (i.e., as functions of t), and determining them so that the more general equation is identically satisfied.

In our case the simpler equation with known solution is the homogeneous equation. Hence we start from Eq. (14.1) and posit

$$y(t) = A_1(t)y_1(t) + A_2(t)y_2(t), \tag{14.27}$$

where $y_1(t), y_2(t)$ are two linearly independent solutions of the homogeneous equation, i.e. the functions appearing in the solutions (14.8), (14.10), (14.20), as the case may be; the arbitrary constants have been set as functions of time. Thus $A_1(t), A_2(t)$ are undetermined functions, assumed to be differentiable. This procedure might appear a dumb thing to do, since we are replacing the problem of finding one unknown function (the particular solution) with the problem of finding *two* unknown functions $A_1(t), A_2(t)$. The fact is that, as we shall see, this method reduces the problem of finding $A_1(t), A_2(t)$ to a simple problem of solving two first-order equations.

Differentiating Eq. (14.27) and rearranging terms we obtain

$$y' = [A_1y_1' + A_2y_2'] + [A_1'y_1 + A_2'y_2], \tag{14.28}$$

from which we see that the differential equation (14.1) will not contain the second-order derivatives of $A_1(t), A_2(t)$ if

$$A_1'y_1 + A_2'y_2 = 0. \tag{14.29}$$

Since the differential equation (14.1) imposes only one condition on the two unknown functions $A_1(t), A_2(t)$, we are free to impose an additional condition on them, and this we will do so as to make the solution of the problem as simple as possible. The obvious simplifying additional condition is (14.29). Thus we have

$$\begin{aligned} y' &= A_1y_1' + A_2y_2', \\ y'' &= A_1'y_1' + A_1y_1'' + A_2'y_2' + A_2y_2''. \end{aligned} \tag{14.30}$$

If we now substitute Eqs. (14.27) and (14.30) into the differential equation (14.1) we obtain

$$a_0(A_1'y_1' + A_1y_1'' + A_2'y_2' + A_2y_2'') + a_1(A_1y_1' + A_2y_2') + a_2(A_1y_1 + A_2y_2) = g(t).$$

Rearranging terms and taking account that both $y_1(t)$ and $y_2(t)$ are a solution to the homogeneous equation (14.2), so that $a_0y_i'' + a_1y_i' + a_2y_i = 0, i = 1, 2$, we get

$$\begin{aligned} &[A_1(a_0y_1'' + a_1y_1' + a_2y_1) + A_2(a_0y_2'' + a_1y_2' + a_2y_2)] + a_0(A_1'y_1' + A_2'y_2') \\ &= a_0(A_1'y_1' + A_2'y_2') = g(t). \end{aligned} \tag{14.31}$$

Thus $y(t) = A_1(t)y_1(t)+A_2(t)y_2(t)$ will be a solution to the non-homogeneous equation (14.1) if $A_1(t), A_2(t)$ satisfy the two equations (14.29) and (14.31):

$$\begin{aligned} A_1'(t)y_1(t) + A_2'(t)y_2(t) &= 0, \\ A_1'(t)y_1'(t) + A_2'(t)y_2'(t) &= a_0^{-1}g(t). \end{aligned} \tag{14.32}$$

We can use system (14.32) to express A_1', A_2' in terms of the remaining functions. Note that the determinant of this linear system is

$$W(t) = \begin{vmatrix} y_1(t) & y_2(t) \\ y_1'(t) & y_2'(t) \end{vmatrix} = y_1(t)y_2'(t) - y_2(t)y_1'(t),$$

namely the Wronskian, that according to general principles (see Chap. 11, Theorem 11.5) is non zero. Therefore we have

$$\begin{aligned} A_1'(t) &= \frac{-a_0^{-1}g(t)y_2(t)}{W(t)}, \\ A_2'(t) &= \frac{a_0^{-1}g(t)y_1(t)}{W(t)}, \end{aligned} \tag{14.33}$$

from which we obtain $A_1(t), A_2(t)$ by integrating the right-hand sides. Hence we have

$$\begin{aligned} A_1(t) &= \int \frac{-a_0^{-1}g(t)y_2(t)}{W(t)}\mathrm{d}t + B_1, \\ A_2(t) &= \int \frac{a_0^{-1}g(t)y_1(t)}{W(t)}\mathrm{d}t + B_2, \end{aligned} \tag{14.34}$$

where B_1, B_2 are arbitrary constants of integration. Note that when B_1, B_2 are allowed to be non-zero, we directly obtain the general solution of the non-homogeneous equation. In fact, by substituting Eq. (14.34) back into Eq. (14.27) we obtain

$$y(t) = [B_1y_1(t)+B_2y_2(t)]+\left\{y_1(t)\int \frac{-a_0^{-1}g(t)y_2(t)}{W(t)}\mathrm{d}t + y_2(t)\int \frac{a_0^{-1}g(t)y_1(t)}{W(t)}\mathrm{d}t\right\}, \tag{14.35}$$

which is the general solution of Eq. (14.2). Hence if we are only interested in the particular solution itself we take B_1, B_2 to be zero, and the particular solution we are looking for is

$$\bar{y}(t) = y_1(t)\int \frac{-a_0^{-1}g(t)y_2(t)}{W(t)}\mathrm{d}t + y_2(t)\int \frac{a_0^{-1}g(t)y_1(t)}{W(t)}\mathrm{d}t. \tag{14.36}$$

The method of variation of parameters thus reduces the problem of finding a particular solution to a problem in integral calculus.

Formula (14.36) can, of course, be applied to the standard forms of $g(t)$ treated with the method of undetermined coefficients. Consider, for example, the case in which $g(t) = F$, a constant, and consider for simplicity the case in which we have distinct real roots of the characteristic equation, so that $y_1(t) = e^{\lambda_1 t}, y_2(t) = e^{\lambda_2 t}$. Then the Wronskian is $W(t) = e^{(\lambda_1+\lambda_2)t}(\lambda_2 - \lambda_1)$, and Eq. (14.36) gives

$$\begin{aligned}\bar{y}(t) &= \frac{F}{a_0}\left\{e^{\lambda_1 t}\int \frac{-e^{-\lambda_1 t}}{(\lambda_2-\lambda_1)}\mathrm{d}t + e^{\lambda_2 t}\int \frac{e^{-\lambda_2 t}}{(\lambda_2-\lambda_1)}\mathrm{d}t\right\}\\ &= \frac{F}{a_0}\left\{e^{\lambda_1 t}\frac{\lambda_1^{-1}e^{-\lambda_1 t}}{(\lambda_2-\lambda_1)} + e^{\lambda_2 t}\frac{-\lambda_2^{-1}e^{-\lambda_2 t}}{(\lambda_2-\lambda_1)}\right\}\\ &= \frac{F}{a_0}\frac{\lambda_1^{-1}-\lambda_2^{-1}}{\lambda_2-\lambda_1} = \frac{F}{a_0}\frac{\lambda_2-\lambda_1}{\lambda_1\lambda_2(\lambda_2-\lambda_1)} = \frac{F}{a_0(\lambda_1\lambda_2)}.\end{aligned} \tag{14.37}$$

Now, since λ_1, λ_2 are the roots of the characteristic equation (14.6), the relations between the roots and the coefficients give us

$$\lambda_1\lambda_2 = b_2 = a_2/a_0, \tag{14.38}$$

and by substituting (14.38) into (14.37) we obtain $\bar{y}(t) = F/a_2$. It is clear that in the standard cases the method of undetermined coefficients is much quicker and simpler than the method of variation of parameters.

14.3 General solution of the non homogeneous equation

After finding a particular solution $\bar{y}(t)$ of the non-homogeneous equation, we can write its general solution as

$$y(t) = f(t; A_1, A_2) + \bar{y}(t), \tag{14.39}$$

where $f(t; A_1, A_2)$ is the general solution of the homogeneous equation.

Let us recall (see Chap. 11, Sect. 11.2.2) that in economic applications the particular solution $\bar{y}(t)$ can usually be interpreted as the *equilibrium value* (a stationary or a moving equilibrium according to whether $\bar{y}(t)$ is a

constant or a function of t) of the variable y. Given this interpretation the general solution of the homogeneous part of the equation can be interpreted as giving the time path of the *deviations* from equilibrium since $y(t) - \bar{y}(t) = f(t; A_1, A_2)$.

14.4 Determination of the arbitrary constants

We must, finally, treat the determination of the arbitrary constants. Of course, two additional conditions are needed. They usually take the form that for a known value of t, say t^*, the value of $y(t)$, say y^*, and of its first derivative, say y'^*, are known (in economics t^* is generally taken to be zero, whence the name *initial conditions*). Substituting t^* and y^* in the general solution and substituting t^* and y'^* in its first derivative, we obtain two linear equation in the two unknowns A_1, A_2. Formally, given that $f(t; A_1, A_2) = A_1y_1(t) + A_2y_2(t)$, where $y_1(t), y_2(t)$ are two distinct solutions of the homogeneous equation, we have

$$\begin{aligned}A_1y_1^* + A_2y_2^* &= y^* - \bar{y}^*,\\ A_1y_1'^* + A_2y_2'^* &= y'^* - \bar{y}'^*,\end{aligned} \tag{14.40}$$

where of course $\bar{y}^*, \bar{y}'^*$ are absent if the differential equation is homogeneous. The solution of this system yields the values of the two arbitrary constant that satisfy the two additional conditions. Note that the determinant of the linear system (14.40) coincides with the Wronskian evaluated at time $t = t^*$. Since the Wronskian is different from zero for any t when the solutions are distinct (see Theorem 11.5), system (14.40) can always be solved.

14.5 Exercises

14.5.1 Examples

1. Let us solve the following second-order differential equation

$$y'' - 6y' + 9y = e^{3t}/t^2$$

with the initial conditions $y(1) = 0, y'(1) = 1.5e^3$.

The characteristic equation of the corresponding homogeneous equation is

$$\lambda^2 - 6\lambda + 9 = 0,$$

that has the double real root $\lambda^* = 3$. Hence the general solution of the homogeneous equation—see Eq. (14.10)—is

$$y(t) = A_1e^{3t} + A_2te^{3t}.$$

To find a particular solution of the non-homogeneous equation we apply the method of variation of parameters. The Wronskian is

$$W(t) = \begin{vmatrix} e^{3t} & te^{3t} \\ 3e^{3t} & (1+3t)e^{3t} \end{vmatrix} = e^{6t},$$

hence equation (14.36) gives us

$$\begin{aligned} \overline{y}(t) &= e^{3t}\int -t^{-1}\mathrm{d}t + te^{3t}\int t^{-2}\mathrm{d}t = e^{3t}(-\ln t) + te^{3t}(-t^{-1}) \\ &= -e^{3t}\ln t - e^{3t}. \end{aligned}$$

The general solution of the non-homogeneous equation thus is

$$y(t) = A_1 e^{3t} + A_2 te^{3t} - e^{3t}\ln t - e^{3t}.$$

To determine the arbitrary constants we have the conditions

$$A_1e^3 + A_2e^3 - e^3 \quad = y(1) \quad = 0,$$

$$3A_1e^3 + 4A_2e^3 - 2e^3 \quad = y'(1) \quad = 1.5e^3,$$

from which, by dividing through by e^3 and solving, we get $A_1 = A_2 = 0.5$.

2. Let us solve the following second-order differential equation

$$Y'' + 3Y' + 6.25Y = -25. \tag{14.41}$$

Two initial conditions, $Y(0) = 0$ and $Y'(0) = -4$, are given.

A particular solution may be obtained assuming Y as constant. Therefore substituting in (14.41) $Y'' = 0$ and $Y' = 0$ we get:

$$\bar{Y} = -4.$$

The characteristic root of the homogeneous part of (14.41) is

$$\lambda^2 + 3\lambda + 6.25 = 0,$$

whose roots are

$$\lambda_1, \lambda_2 = -1.5 \pm 2i.$$

Thus the general solution of (14.41) is

$$Y(t) = e^{-1.5t}(A_1\cos 2t + A_2\sin 2t) - 4.$$

Using the initial conditions we can determine the value of the two arbitrary constants. We have

$$Y(0) = 0 = e^0(A_1\cos 0 + A_2\sin 0) - 4;$$

therefore

$$0 = A_1 - 4,$$

and so

$$A_1 = 4.$$

Moreover

$$\begin{aligned} Y' &= -1.5A_1e^{-1.5t}\cos 2t - 2tA_1e^{-1.5t}\sin 2t \\ &\quad -1.5A_2e^{-1.5t}\sin 2t + 2A_2e^{-1.5t}\cos 2t, \end{aligned}$$

and so, for $t = 0$, since $\cos 0 = 1, \sin 0 = 0$, we have

$$Y'(0) = -4 = -1.5A_1 + 2A_2.$$

Since $A_1 = 4$, we obtain

$$A_2 = 1.$$

14.5.2 Other exercises

(a) Find the solution of the following differential equations:

$$\begin{array}{lllll} (i) & y''(t) & +y'(t) & -y & = -10; \\ (ii) & y''(t) & +y'(t) & & = -5; \\ (iii) & y''(t) & +12y'(t) & +18y & = 54; \\ (iv) & -0.2y''(t) & -0.4y'(t) & -10y & = -200 - 10\cos 4\pi t. \end{array}$$

(b) Find the solution of the following differential equations, and determine the arbitrary constants:

$$\begin{array}{llllcll} (i) & y''(t) & & -3y & = & 0 & y(0)=0, & y'(0)=-3; \\ (ii) & y''(t) & +2y'(t) & +y & = & t^2 & y(0)=0, & y'(0)=1; \\ (iii) & 2y''(t) & -3y'(t) & +y & = & (t^2+1)e^t & y(0)=5, & y'(0)=14; \\ (iv) & 3y''(t) & +4y'(t) & +y & = & (\sin t)e^{-t} & y(0)=1, & y'(0)=0. \end{array}$$

14.6 References

Allen, R.D.G., 1959, *Mathematical Economics,* Chap. 4, Sects. 4.1-4.5; Chap. 5, Sect. 5.4.

Baumol, W.J., 1970, *Economic Dynamics*, Chap. 14, Sects. 14.3-14.6, 14.8.

Braun, M., 1986, *Differential Equations and Their Applications,* Chap. 2.

Courant, R., 1937, *Differential and Integral Calculus,* Vol. I, Chap. VIII, Sect. 7; Chap. XI, Sects. 2, 3.

Chapter 15

Second-order Differential Equations in Economic Models

While second- and higher-order difference equations are frequently met in economics, examples of the use of 'direct' second- or higher-order differential equations are scarce. Let us first explain what we mean by 'direct'. As we shall see in Chap. 18, a simultaneous system of differential equations can always be transformed into a single differential equation of a certain order. Thus a second- or higher-order differential equation can result from the transformation of a system. By 'direct' we mean differential equations that arise directly, and not as a transformation of a system, in economic dynamics.

Now, in the context of period analysis, it is normally the case that lags can be plausibly poured over the model: for example, in any macroeconomic model built in discrete time one can introduce investment and consumption functions with distributed lags. On the contrary, in the context of continuous-time analysis it is more difficult to give an economic meaning to derivatives of order higher than the first, and hence to introduce them 'directly' into an economic model.

15.1 The second-order accelerator

We have already met the accelerator in the context of Samuelson's multiplier-accelerator model in discrete time (see Sect. 6.1). A more general investment function relates net fixed investment to the gap between the desired and the actual capital stock, i.e.

$$K' = I = \alpha(K^* - K), \tag{15.1}$$

where $\alpha > 0$ denotes the speed of adjustment, and for simplicity we have assumed no depreciation. Equation (15.1) is a partial adjustment equation of the type illustrated in Sect. 12.4. and is called the *capital stock adjustment* *principle*. This equation leaves the desired capital stock undetermined, and can give rise to different investment functions according to the specification of K^*. The desired capital stock is presumably related to output (we neglect financial factors), which should be *expected* output, but to keep the matter simple it is often assumed that $K^* = kY$, namely K^* is proportional to current output according to the capital/output ratio k, that can be taken as fixed in the short run. Thus we have

$$K' = I = \alpha(kY - K). \tag{15.2}$$

Equation (15.2) is called the *flexible accelerator*, and has been introduced by Goodwin (1951), although in a non-linear context. The simple accelerator can be shown to be a particular case of the flexible accelerator when the adjustment speed tends to infinity. In fact, in the case of $\alpha \to \infty$, we have instantaneous adjustment (see Sect. 12.4), i.e. $K = kY$. Hence the capital stock varies simultaneously with output, i.e., differentiating with respect to time we get $I = kY'$, that is the continuous-time formulation of the simple accelerator.

The second-order accelerator represents a further development of the capital stock adjustment principle and of the flexible accelerator, and arises when I in Eq. (15.1) is interpreted not as the actual, but as the desired level of investment, I^*. Actual investment is then carried out according to a partial adjustment equation of I toward I^*, namely

$$I' = \beta(I^* - I) = \beta[\alpha(K^* - K) - I]. \tag{15.3}$$

The idea underlying the second-order accelerator is that actual investment expenditure is the result of a two-stage decision process: in the first stage entrepreneurs determine the gap between the actual and the desired capital stock, that they would like to fill with a certain adjustment speed measured by α; this gives rise to desired investment. In the second stage they carry out investment, that however is not instantaneously equal to its desired level due to production lags and other frictions, hence the partial adjustment equation of I toward I^*. The second-order accelerator was developed independently by Gandolfo and Hillinger in the 1970's, and can be given a rigorous microeconomic foundation (for this aspect see Hillinger et al., 1992, where previous studies are also examined).

Since $I' = K''$, by rearranging terms in Eq. (15.3) we have

$$K'' + \beta K' + \alpha\beta K = \alpha\beta K^*. \tag{15.4}$$

Let us neglect for the moment the determinants of K^*, and consider it as a generic function of time. Then Eq. (15.4) is a second-order non-homogeneous equation in K. The characteristic equation of the homogeneous part is

$$\lambda^2 + \beta\lambda + \alpha\beta = 0.$$

Since the succession of signs of the coefficients is $+ + +$, the movement will be stable. As regards the nature of the roots, let us consider the discriminant $\Delta = \beta^2 - 4\alpha\beta$, which will be positive, zero, or negative according as

$$\beta \gtreqless 4\alpha. \tag{15.5}$$

Hence the model is capable of generating stable oscillations endogenously, i.e. its ability to describe investment cycles is independent of the specification of K^*. Hillinger refers to the gradual, rather than immediate, adjustment of I to I^* as the 'inertia' of investment, and observes that 'as in physical systems, inertia potentially leads to overshooting and cyclical behaviour because $K^* = K$ does not imply $I = 0$, so that a positive or negative rate of net investment continues and K again departs from its equilibrium level' (p. 168).

A particular solution to Eq. (15.4) can be found by applying the general formula of variation of parameters, Eq. (14.36) in Chap. 14. This gives

$$\overline{K}(t) = K_1(t) \int \frac{K^*(t)K_2(t)}{W(t)} dt + K_2(t) \int \frac{K^*(t)K_1(t)}{W(t)} dt, \tag{15.6}$$

where $K_1(t), K_2(t)$ are the two components of the solution of the homogeneous equation, and $W(t)$ is the Wronskian of that equation.

This treatment is valid only in so far as K^* is independent of Y. When K^* is related to Y, for example through the naive assumption $K^* = kY$ that we have already used before, the structure of the homogeneous part changes unless we are willing to assume that Y is wholly exogenous. In fact, in the context of a simple model of income determination, Y is related to I through the multiplier. If we build a very simple macroeconomic model by adding a standard consumption function without any lag (the introduction of a partial adjustment equation in the consumption function would increase the order of the final equation), $C = a + bY, 0 < b < 1$, we obtain

$$Y = C + I = \frac{1}{1-b}(I + a) = \frac{1}{1-b}(K' + a), \tag{15.7}$$

which is the standard multiplier equation. With $K^* = kY$, Eq. (15.4) becomes

$$K'' + \beta K' + \alpha\beta K = \alpha\beta k \frac{K'}{1-b} + \alpha\beta k \frac{a}{1-b},$$

i.e.

$$K'' + \beta\left(1 - \frac{\alpha k}{1-b}\right) K' + \alpha\beta K = \alpha\beta k \frac{a}{1-b}. \tag{15.8}$$

A particular solution is easily found by letting $K = \overline{K}$= constant, whence

$$\overline{K} = k\frac{a}{1-b} = kY_e, \tag{15.9}$$

which is the capital stock corresponding to the static equilibrium value of output.

The characteristic equation of the homogeneous part is

$$\lambda^2 + \beta\frac{1-b-\alpha k}{1-b}\lambda + \alpha\beta = 0. \tag{15.10}$$

The model is no longer certainly stable: the crucial stability condition is $1-b-\alpha k > 0$, namely

$$\alpha < \frac{s}{k}, \tag{15.11}$$

where $s \equiv 1-b$. Thus the coefficient of adjustment of K to K^* must be smaller than the critical value s/k. Empirical research carried out by Gandolfo et al. (1990) shows α to be very low (the corresponding mean time-lag is of the order of a few years) and such that the stability condition would certainly be satisfied.

Oscillations require $\Delta < 0$, i.e.

$$\beta < \left(\frac{1-b}{1-b-\alpha k}\right)^2 4\alpha. \tag{15.12}$$

If the stability condition is satisfied, the fraction on the right-hand side turns out to be greater than one, hence inequality (15.12) is more likely to be satisfied than the analogous inequality $\beta < 4\alpha$ that we found in the case of an exogenous K^*.

It should be noted, in conclusion, that a feature of the second-order accelerator is that cycles can be generated solely in the investment sector, which implies a minor role for consumption in causing business cycles. This result contrasts with the standard multiplier-accelerator model of the business cycle (see Chap. 6, Sect. 6.1), that gives an equally important role to consumption.

15.2 Exercises

1. Suppose that K^* is related to expected output, and that expectations are of the extrapolative type, from which

$$K^* = k(Y + \gamma Y'), \quad \gamma > 0.$$

Examine the behaviour of the solution to the second-order accelerator model. Consider then the case of regressive expectations ($\gamma < 0$) and compare the results in the two cases.

2. In the foreign exchange market non-speculators (commercial traders, etc.) are permanently present, whose excess demand depends on the current exchange rate and on seasonal factors, represented by a periodic oscillation:

$$E_n(t) = a_0 + a_1 SR(t) + B\cos\omega t, \quad a_0 > 0, a_1 < 0, 0 < B < a_0.$$

where $SR(t)$ denotes the current spot exchange rate (number of units of domestic currency per unit of foreign currency). Hence in the absence of other agents the exchange rate would follow the periodic path

$$SR(t) = \frac{B}{-a_1}\cos\omega t - \frac{a_0}{a_1},$$

which is determined by the equilibrium condition $E_n(t) = 0$. Let us now introduce speculators, who demand and supply foreign exchange in the expectation of a change in the exchange rate. Their excess demand for foreign exchange is given by

$$E_s(t) = m[ER(t) - SR(t)], \quad m > 0,$$

where $ER(t)$ denotes the expected spot exchange rate. The market equilibrium condition is now $E_n(t) + E_s(t) = 0$. To determine the path of the exchange rate we need to know how expectations are formed. Let us assume the following (admittedly ad hoc) process of expectation formation:

$$ER(t) = SR(t) + b_1 SR'(t) + b_2 SR''(t),$$

i.e., speculators base their expectations on the current rate of exchange, on the direction in which it is moving, and on the acceleration of its movement. The signs of b_1, b_2 are left unspecified. In fact, the exercise consists of the following problems:

(2.a) determine the signs of b_1, b_2 so that the model is stable;

(2.b) show that the particular solution involves a periodic oscillation having the same frequency as the basic seasonal factors;

(2.c) is it possible to say that the particular solution has a smaller amplitude than the one involved in the path of the exchange rate when speculators are absent?

3. Suppose that the monetary authorities are also operating in the foreign exchange market with the aim of stabilizing the exchange rate at the constant value a_0/a_1. Assume that the authorities' excess demand can be represented by the following function:

$$E_G(t) = f_1\left[SR(t) - \left(\frac{a_0}{-a_1}\right)\right] + f_2 SR'(t),$$

where f_1, f_2 are policy parameters.

(3.a) Prove that the stability conditions are always satisfied for a suitable choice of f_1, f_2;

(3.b) Examine the influence of f_1, f_2 on the particular solution;

(3.c) Find an explicit expression for the change in the authorities' foreign exchange reserves over the interval $0-T$, where T is a given positive number (Hint: the authorities excess demand for foreign exchange gives the instantaneous value of the change in reserves. Hence $\int_0^T E_G(t)dt$ is the change over the period considered. Given the solution for $SR(t)$, etcetera).

15.3 References

Gandolfo, G. and P.C. Padoan, 1990, The Italian Continuous Time Model: Theory and Empirical Results, p. 97 and p. 107.
Goodwin, R.M., 1951, The Nonlinear Accelerator and the Persistence of Business Cycles.
Hillinger, C. (ed.), 1992, *Cyclical Growth in Market and Planned Economies,* Chap. 8.

Chapter 16

Higher-order Differential Equations

The general form of an n-th order differential equation is

$$a_0 y^{(n)} + a_1 y^{(n-1)} + ... + a_{n-1}y' + a_n y = g(t), \tag{16.1}$$

where $y^{(n)}, y^{(n-1)}$, etc., indicate the derivatives of the order $n, n-1$, etc.; the a's are given constants and $g(t)$ is a known function. Some a's may be zero, but of course a_0 must be different from zero if the equation is of order n.

16.1 Solution of the homogeneous equation

Let us begin, as usual, with the homogeneous equation

$$a_0 y^{(n)} + a_1 y^{(n-1)} + ... + a_{n-1}y' + a_n y = 0. \tag{16.2}$$

To solve this equation we adopt the same method that worked successfully in the case of first- and second-order equations, namely we assume a solution of the type $e^{\lambda t}$ and substitute it into Eq. (16.2). Thus we obtain

$$\begin{aligned} & a_0\lambda^n e^{\lambda t} + a_1\lambda^{n-1}e^{\lambda t} + ... + a_{n-1}\lambda e^{\lambda t} + a_n e^{\lambda t} \\ = \; & e^{\lambda t}\left(a_0\lambda^n + a_1\lambda^{n-1} + ... + a_{n-1}\lambda + a_n\right) = 0. \end{aligned} \tag{16.3}$$

If $e^{\lambda t}$ is a solution of the differential equation (16.2), Eq. (16.3) must be satisfied identically. Since $e^{\lambda t} \neq 0$, it is necessary and sufficient that

$$a_0\lambda^n + a_1\lambda^{n-1} + ... + a_{n-1}\lambda + a_n = 0. \tag{16.4}$$

Equation (16.4) is called the *characteristic* (or *auxiliary*) *equation* of the differential equation (16.2), and the polynomial on the left-hand side is called the *characteristic polynomial.* Thus we have reduced the solution of a differential equation to the solution of an algebraic equation. The solution of Eq. (16.4) yields exactly n roots (the *characteristic roots*), which may be real and/or complex, simple and/or repeated.

In the case of n distinct real roots, we have n functions $\exp(\lambda_i t)$, each being a distinct solution to the homogeneous differential equation (16.2). According to general principles—see Theorem 11.4—we can combine them linearly and obtain the general solution

$$y(t) = A_1 e^{\lambda_1 t} + A_2 e^{\lambda_2 t} + ... + A_n e^{\lambda_n t}, \tag{16.5}$$

where $A_1, A_2, ..., A_n$ are arbitrary constants.

If λ^* is a repeated real root of multiplicity $m \leq n$, then also $t\exp(\lambda^* t)$, $t^2\exp(\lambda^* t), ..., t^{m-1}\exp(\lambda^* t)$ are solutions of the homogeneous equation. In general, the solution of the homogeneous equation in the case of repeated real roots is

$$y(t) = \sum_{j=1}^{k} P_j(t) e^{\lambda_j^* t}, \tag{16.6}$$

where λ_j^* denotes the distinct roots of the characteristic equation, each counted with its own multiplicity, and $P_j(t)$ are polynomials of the type

$$P_j(t) = A_{1_j} + A_{2_j} t + ... + A_{m_j} t^{m_j - 1}, \tag{16.7}$$

where the A's are arbitrary constants and m_j is the multiplicity of j-th root.

In the case of complex roots, that always occur in conjugate pairs, each pair will give rise to a trigonometric oscillation of the kind

$$e^{\alpha t}(A_1 \cos\theta t + A_2 \sin\theta t), \tag{16.8}$$

exactly as in the case of second-order equations (see Sect. 14.1.3). A complication that could not arise in the case of second-order equations is that one pair (or more) of complex roots may be repeated. Then in the solution we shall have terms of the kind

$$e^{\alpha t}[(A_{11_j} + A_{12_j} t + ... + A_{1m_j} t^{m_j-1}) \cos\theta t$$

$$+(A_{21_j} + A_{22_j} t + ... + A_{2m_j} t^{m_j-1})] \sin\theta t, \tag{16.9}$$

where m_j is the number of times that the j-th pair of complex roots is repeated, and the A's are $2m_j$ arbitrary constants.

Since in the same equation complex (simple or repeated) roots may occur together with real (simple or repeated) roots, a great variety of movements is possible.

16.2 Solution of the non-homogeneous equation

A particular solution of the non-homogeneous equation can usually be found by applying the method of undetermined coefficients (see Chap. 12, Sect. 12.2). We shall exemplify in the case in which $g(t) = G$, a constant. As a particular solution, try $\overline{y} = B$, an undetermined constant. Since all the derivatives of a constant are zero, substitution in Eq. (16.1) yields $a_n B = G$, and so $B = G/a_n$.

If $a_n = 0$, try $\overline{y} = Bt$. Substituting in Eq. (16.1) and solving for B we obtain $B = G/a_{n-1}$. If also $a_{n-1} = 0$, try $\overline{y} = Bt^2$, whence $B = G/2a_{n-2}$, and so on. In the extreme case in which all coefficients but a_0 are zero, a particular solution will be $\overline{y} = Bt^n$, where $B = G/a_0 n!$.

Note that in this last case the characteristic equation of the homogeneous part, which is $a_0 y^{(n)} = 0$, will be $a_0\lambda^n = 0$, which has a root $\lambda^* = 0$ repeated n times. By applying Eqs. (16.6) and (16.7), account being taken that $e^0 = 1$, we obtain the general solution of the homogeneous equation as $A_1 + A_2 t + ... + A_n t^{n-1}$, and so the general solution of the non-homogeneous equation $a_0 y^{(n)} = G$ is $y(t) = A_1 + A_2 t + ... + A_n t^{n-1} + (G/a_0 n!)t^n$. We could have obtained this result directly by integrating the differential equation n times.

16.2.1 Variation of parameters

The principles of this method have already been illustrated in Chap. 12, Sect. 12.2.6. Let us recall that this is a general method of solving a differential equation by considering the arbitrary constants that appear in the known solution of a simpler equation, as variable (i.e., as functions of t), and determining them so that the more general equation is identically satisfied.

In our case the simpler equation with known solution is the homogeneous equation. Hence we start from Eq. (16.1) and posit

$$y(t) = A_1(t)y_1(t) + A_2(t)y_2(t) + ... + A_n(t)y_n(t), \tag{16.10}$$

where $y_1(t), y_2(t), ..., y_n(t)$ are n linearly independent solutions of the homogeneous equation, i.e. the functions appearing in the solutions (16.5)-(16.9) as the case may be; the arbitrary constants have been set as functions of time. Thus $A_1(t), A_2(t), ..., A_n(t)$ are undetermined functions, assumed to be differentiable. This procedure might appear a dumb thing to do, since we are replacing the problem of finding one unknown function (the particular solution) with the problem of finding n unknown functions $A_1(t), A_2(t), ...,$ $A_n(t)$. The fact is that, as we shall see, this method reduces the problem of finding $A_1(t), A_2(t)$,..., $A_n(t)$ to the simpler problem of solving n separate first-order equations.

Differentiating Eq. (16.10), and rearranging terms, we obtain

$$y' = [A_1y_1' + A_2y_2' + ... + A_ny_n'] + [A_1'y_1 + A_2'y_2 + ... + A_n'y_n]. \tag{16.11}$$

To keep the finding of the functions $A_i(t)$ as simple as possible, we want to prevent their second- and higher-order derivatives from appearing in the differential equation (16.1). From Eq. (16.11) we see that the differential equation (16.1) will not contain the second-order derivatives of $A_1(t), A_2(t), ..., A_n(t)$ if

$$A_1'y_1 + A_2'y_2 + ... + A_n'y_n = 0. \tag{16.12}$$

Again differentiating Eq. (16.11), account being taken of Eq. (16.12), to obtain y'', and rearranging terms, we have

$$y'' = [A_1y_1'' + A_2y_2'' + ... + A_ny_n''] + [A_1'y_1' + A_2'y_2' + ... + A_n'y_n'], \tag{16.13}$$

from which we see that the differential equation (16.1) will not contain the second-order derivatives of $A_1(t), A_2(t), ..., A_n(t)$ if

$$A_1'y_1' + A_2'y_2' + ... + A_n'y_n' = 0. \tag{16.14}$$

Proceeding in like manner we obtain the further conditions

$$\begin{array}{rcl} A_1'y_1'' + A_2'y_2'' + ... + A_n'y_n'' & = & 0, \\ \dots\dots\dots\dots\dots\dots & ... & ... \\ A_1'y_1^{(n-2)} + A_2'y_2^{(n-2)} + ... + A_n'y_n^{(n-2)} & = & 0. \end{array} \tag{16.15}$$

Since the differential equation (16.1) imposes only one condition on the n unknown functions $A_1(t), A_2(t), ..., A_n(t)$, we are free to impose $n-1$ additional conditions on them, and this we will do so as to make the solution of the problem as simple as possible. The obvious simplifying additional conditions are Eqs. (16.12), (16.14), (16.15). Thus we obtain the following expressions for the various derivatives of the assumed solution (16.10) for $y(t)$:

$$\begin{array}{rcl} y' & = & A_1y_1' + A_2y_2' + ... + A_ny_n', \\ y'' & = & A_1y_1'' + A_2y_2'' + ... + A_ny_n'', \\ \dots\dots & ... & \dots\dots\dots\dots\dots\dots \\ y^{(n-2)} & = & A_1y_1^{(n-2)} + A_2y_2^{(n-2)} + ... + A_ny_n^{(n-2)}, \\ y^{(n-1)} & = & A_1y_1^{(n-1)} + A_2y_2^{(n-1)} + ... + A_ny_n^{(n-1)}, \\ y^{(n)} & = & A_1y_1^{(n)} + A_2y_2^{(n)} + ... + A_ny_n^{(n)} \\ & & +A_1'y_1^{(n-1)} + A_2'y_2^{(n-1)} + ... + A_n'y_n^{(n-1)}. \end{array} \tag{16.16}$$

If we now substitute Eqs. (16.10) and (16.16) into the differential equation (16.1) we obtain

$$\begin{array}{l} a_0\left(A_1y_1^{(n)} + A_2y_2^{(n)} + ... + A_ny_n^{(n)} + A_1'y_1^{(n-1)} + A_2'y_2^{(n-1)} + ... + A_n'y_n^{(n-1)}\right) \\ +a_1\left(A_1y_1^{(n-1)} + A_2y_2^{(n-1)} + ... + A_ny_n^{(n-1)}\right) \\ +\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ +a_n\left(A_1y_1 + A_2y_2 + ... + A_ny_n\right) = g(t). \end{array}$$

Rearranging terms we get

$$\begin{array}{l} A_1\left(a_0y_1^{(n)} + a_1y_1^{(n-1)} + ... + a_ny_1\right) + A_2\left(a_0y_2^{(n)} + a_1y_2^{(n-1)} + ... + a_ny_2\right) \\ +... + A_n\left(a_0y_n^{(n)} + a_1y_n^{(n-1)} + ... + a_ny_n\right) \\ +a_0\left(A_1'y_1^{(n-1)} + A_2'y_2^{(n-1)} + ... + A_n'y_n^{(n-1)}\right) = g(t), \end{array}$$

from which—taking account that $y_1(t),\ y_2(t), ..., y_n(t)$ are solutions to the homogeneous equation (16.2), so that $a_0y_i^{(n)} + a_1y_i^{(n-1)} + ... + a_ny_i = 0, i = 1, 2, ..., n$—we obtain

$$a_0\left(A_1'y_1^{(n-1)} + A_2'y_2^{(n-1)} + ... + A_n'y_n^{(n-1)}\right) = g(t). \tag{16.17}$$

Thus $y(t) = A_1(t)y_1(t) + A_2(t)y_2(t) + ... + A_n(t)y_n(t)$ will be a solution to the non-homogeneous equation (16.1) if $A_1(t), A_2(t), ..., A_n(t)$ satisfy the n equations (16.12), (16.14), (16.15) and (16.17), that is the system

$$\begin{array}{rcl} A_1'(t)y_1(t) + A_2'(t)y_2(t) + ... + A_n'(t)y_n(t) & = & 0, \\ A_1'(t)y_1'(t) + A_2'(t)y_2'(t) + ... + A_n'(t)y_n'(t) & = & 0, \\ \dots\dots\dots\dots\dots\dots\dots\dots & ... & ... \\ A_1'(t)y_1^{(n-1)}(t) + A_2'(t)y_2^{(n-1)}(t) + ... + A_n'(t)y_n^{(n-1)}(t) & = & a_0^{-1}g(t). \end{array} \tag{16.18}$$

We can use system (16.18) to express $A_1', A_2', ..., A_n'$ in terms of the remaining functions, that are all known. Note that the determinant of this linear system is

$$W(t) = \begin{vmatrix} y_1(t) & y_2(t) & ... & y_n(t) \\ y_1'(t) & y_2'(t) & ... & y_n'(t) \\ ... & ... & ... & ... \\ y_1^{(n-1)}(t) & y_2^{(n-1)}(t) & ... & y_n^{(n-1)}(t) \end{vmatrix},$$

namely the Wronskian, that according to general principles (see Chap. 11, Theorem 11.5) is non zero. Therefore, solving by Cramer's rule, we have

$$\begin{array}{rcl} A_1'(t) & = & (-1)^{n+1}a_0^{-1}g(t)\dfrac{W_1(t)}{W(t)}, \\ A_2'(t) & = & (-1)^{n+2}a_0^{-1}g(t)\dfrac{W_2(t)}{W(t)}, \\ ... & ... & \dots\dots\dots\dots \\ A_n'(t) & = & (-1)^{2n}a_0^{-1}g(t)\dfrac{W_n(t)}{W(t)}, \end{array} \tag{16.19}$$

where $W_1(t), W_2(t), ..., W_n(t)$ are the appropriate minors of the Wronskian. We finally obtain $A_1(t), A_2(t), ..., A_n(t)$ by integrating the right-hand sides. Hence we have

$$\begin{aligned} A_1(t) &= (-1)^{n+1}a_0^{-1}\int \frac{g(t)W_1(t)}{W(t)}\mathrm{d}t + B_1, \\ A_2(t) &= (-1)^{n+2}a_0^{-1}\int \frac{g(t)W_2(t)}{W(t)}\mathrm{d}t + B_2, \\ \dots &\quad \dots \dots\dots\dots\dots\dots\dots\dots \\ A_n(t) &= (-1)^{2n}a_0^{-1}\int \frac{g(t)W_n(t)}{W(t)}\mathrm{d}t + B_n, \end{aligned} \tag{16.20}$$

where $B_1, B_2, ..., B_n$ are arbitrary constants of integration. Note that when $B_1, B_2, ..., B_n$ are allowed to be non-zero, we directly obtain the general solution of the non-homogeneous equation. In fact, by substituting Eq. (16.20) back into Eq. (16.10) we obtain

$$\begin{aligned} y(t) &= [B_1y_1(t) + B_2y_2(t) + ... + B_ny_n(t)] \\ &\quad + y_1(t)(-1)^{n+1}a_0^{-1}\int \frac{g(t)W_1(t)}{W(t)}\mathrm{d}t \\ &\quad + y_2(t)(-1)^{n+2}a_0^{-1}\int \frac{g(t)W_2(t)}{W(t)}\mathrm{d}t \\ &\quad + ... + y_n(t)(-1)^{2n}a_0^{-1}\int \frac{g(t)W_n(t)}{W(t)}\mathrm{d}t, \end{aligned} \tag{16.21}$$

which is the general solution of Eq. (16.2). Hence if we are only interested in the particular solution itself we take $B_1, B_2, ..., B_n$ to be zero, and the particular solution we are looking for is

$$\begin{aligned} \overline{y}(t) &= y_1(t)(-1)^{n+1}a_0^{-1}\int \frac{g(t)W_1(t)}{W(t)}\mathrm{d}t + y_2(t)(-1)^{n+2}a_0^{-1}\int \frac{g(t)W_2(t)}{W(t)}\mathrm{d}t \\ &\quad + ... + y_n(t)(-1)^{2n}a_0^{-1}\int \frac{g(t)W_n(t)}{W(t)}\mathrm{d}t. \end{aligned} \tag{16.22}$$

The method of variation of parameters thus reduces the problem of finding a particular solution to a problem in integral calculus.

Formula (16.22) can, of course, be applied to the standard forms of $g(t)$ treated with the method of undetermined coefficients, but this latter method is much simpler and quicker in the standard cases.

16.3 Determination of the arbitrary constants

To determine the n arbitrary constants appearing in the solution, n additional conditions are needed. These usually take the form of $y(t), y'(t), ..., y^{(n-1)}(t)$

being known values for a certain $t = t_0$ which is usually taken as the initial point ($t_0 = 0$), whence the name of *initial conditions*. The problem of solving a differential equation subject to the initial conditions is also known as the *initial value problem*.

Suppose then that $y(0), y'(0), ..., y^{(n-1)}(0)$ are known, and substitute these values in the general solution

$$y(t) = A_1y_1(t) + A_2y_2(t) + ... + A_ny_n(t) + \overline{y}(t).$$

The result is

$$\sum_{j=1}^{n} A_jy_j^{(i)}(0) = y^{(i)}(0) - \overline{y}_i(0), \quad i = 0, 1, ..., n-1, \tag{16.23}$$

where the superscript (i) denotes the order of the derivative ($i = 0$ means no differentiation), and $\overline{y}(0) \equiv 0$ if the equation is homogeneous. Equations (16.23) form a linear system whose determinant is $W(0)$, namely the Wronskian evaluated at the given point of time. Since—according to general principles (see Theorem 11.5)—the Wronskian is different from zero, this system can always be solved.

16.4 Stability conditions

As we have seen, in the solution of higher-order difference equations there are no *conceptual* difficulties greater than those met in relation to second-order equations. The 'jump' in conceptual difficulty occurs when we pass from first- to second-order equations (complex roots, etc.). From the second-order on we do not think that the conceptual difficulties are greater. The greater difficulty of higher-order equations lies in the practical problem of how to find the roots of the characteristic equation. This is a problem in numerical analysis (nowadays easily solvable with the current computing equipment) which is outside the scope of this book.

This problem, in any case, is *not* of great importance for the economic theorist, who works with *qualitative* information only. In this connection it would be highly desirable to have conditions—of the kind of Eqs. (14.23)—to check the stability of the movement by means of inequalities involving the coefficients of the characteristic equation, i.e. to check whether all the roots of the characteristic equation have negative real parts without solving the characteristic equation.

Another way of stating stability is that the roots are strictly lying in the left half of complex plane, that is, they are 'stable' roots. The reason why also the roots with zero real part (this includes both zero real roots and complex roots with zero real part) are to be excluded is that we want *asymptotic* stability, that is $\lim_{t\to+\infty} y(t) = 0$, where $y(t)$ is the general solution of the homogeneous equation (for a more detailed and rigorous treatment of the notion of stability see Part III, Chap. 21). Now, a root with zero real part gives rise, in the solution, either to a constant term (zero real root, whence $e^0 = 1$) or to a constant-amplitude oscillation (pair of complex conjugate roots with zero real part). In each of these cases the time path, while not being divergent, is not stable in the sense defined above.

These conditions are fairly easy to find in the case of real roots: in fact, the condition that all the non-zero coefficients in the characteristic equation have the same sign excludes, by Descartes' theorem, the presence of positive real roots, and the condition $a_n \neq 0$ excludes the presence of zero real roots. But in principle also complex roots may occur, which makes the matter rather complicated, as the following simple case of a third-order equation makes clear. Let us then consider the characteristic equation

$$a_0\lambda^3 + a_1\lambda^2 + a_2\lambda + a_3 = 0, \quad a_0 > 0, \tag{16.24}$$

where we have taken $a_0 > 0$, which implies no loss of generality. The conditions $a_i > 0, i = 1, 2, 3$ exclude non-negative real roots. As regards the complex roots, let us consider the relations between the roots and the coefficients of an algebraic equation (see, for example, Turnbull, 1957, p. 66), that in the case of a third-degree equation are

$$\begin{aligned} a_1/a_0 &= -(\lambda_1 + \lambda_2 + \lambda_3), \\ a_2/a_0 &= \lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3, \\ a_3/a_0 &= -\lambda_1\lambda_2\lambda_3. \end{aligned} \tag{16.25}$$

Now suppose that $\lambda_2, \lambda_3 = \alpha \pm i\theta$ (the conditions for a cubic equation to have one real and two complex roots is given, for example, in Turnbull, 1957, Chap. IX, Sect. 52, but need not interest us here). The positivity of all the coefficients ensures that $\lambda_1 < 0$. From Eqs. (16.25) we have

$$\begin{aligned} -(\lambda_1 + 2\alpha) &= a_1/a_0 > 0, \\ 2\alpha\lambda_1 + \alpha^2 + \theta^2 &= a_2/a_0 > 0, \\ -\lambda_1(\alpha^2 + \theta^2) &= a_3/a_0 > 0. \end{aligned} \tag{16.26}$$

Consider now the additional inequality

$$\frac{a_1}{a_0}\frac{a_2}{a_0} - \frac{a_3}{a_0} > 0 \quad \text{or } a_1a_2 - a_0a_3 > 0, \tag{16.27}$$

whence, by substituting from Eqs. (16.26) and rearranging terms, we obtain

$$2[-(\lambda_1 + 2\alpha)\lambda_1 - (\alpha^2 + \theta^2)]\alpha > 0. \tag{16.28}$$

Since λ_1 is negative, and $-(\lambda_1 + 2\alpha) > 0$ by the first inequality in (16.26) while $(\alpha^2 + \theta^2) > 0$ by definition, the expression in square brackets in (16.28) is negative. Hence inequality (16.28) implies, and is implied by, $\alpha < 0$.

When all roots are real, by substituting from Eqs. (16.25) into (16.27) we see that the latter is certainly satisfied when the real roots are negative, hence it is implied by the conditions $a_i > 0$, which is obvious.

From this it also follows that, when $a_i > 0$ (no non-negative real root) and inequality (16.27) turns into an *equality*, then there is certainly a pair of complex roots with zero real part. In fact, $a_1a_2 - a_0a_3 = 0$ cannot hold when there are three negative real roots (in such a case it must be $a_1a_2 - a_0a_3 > 0$, as we have just noted), hence there must be a pair of complex roots with $\alpha = 0$, as can be seen from (16.28). For the same reason it can be seen that, when $a_i > 0$, if inequality (16.28) is reversed, there will be a pair of complex roots with positive real part.

Thus we have reached the conclusion that a set of necessary and sufficient conditions for all the roots of Eq. (16.24) to have negative real parts is

$$\begin{aligned} a_1 &> 0, \\ a_2 &> 0, \\ a_3 &> 0, \\ a_1a_2 - a_0a_3 &> 0. \end{aligned} \tag{16.29}$$

Let us note that either the first or the second inequality can be eliminated, since either one is implied by the remaining three.

In the case of an n-th order equation general conditions exist which are necessary and sufficient for all the roots to be stable. We shall first state the classical Routh-Hurwitz theorem and then the Liénard-Chipart stability criteria that allow a simplification in the conditions (Gantmacher, 1959; see also Barnett, 1973, 1983).

Let there be given the polynomial equation with real coefficients

$$a_0\lambda^n + a_1\lambda^{n-1} + ... + a_{n-1}\lambda + a_n = 0, \quad a_0 > 0, \tag{16.30}$$

where without loss of generality a_0 has been taken as positive. Then we can state the following conditions.

16.4.1 Necessary and sufficient stability conditions (Routh-Hurwitz)

Form the array of coefficients

$$\begin{bmatrix} a_1 & a_3 & a_5 & a_7 & \dots & 0 \\ a_0 & a_2 & a_4 & a_6 & \dots & 0 \\ 0 & a_1 & a_3 & a_5 & \dots & 0 \\ 0 & a_0 & a_2 & a_4 & \dots & 0 \\ 0 & 0 & a_1 & a_3 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & a_n \end{bmatrix} \tag{16.31}$$

Note that all the a's with a subscript greater than n or negative must be treated as zero. Only the first n rows and columns of the array (16.31) have to be considered. The Routh-Hurwitz theorem states that necessary and sufficient conditions for all the roots of Eq. (16.30) to have negative real parts is given by the simultaneous verification of the following inequalities:

$$\Delta_1 > 0, \quad \Delta_2 > 0, ..., \quad \Delta_n > 0, \tag{16.32}$$

where $\Delta_1, \Delta_2, ..., \Delta_n$ are the leading principal minors of (16.31), namely

$$\Delta_1 = a_1, \quad \Delta_2 = \begin{vmatrix} a_1 & a_3 \\ a_0 & a_2 \end{vmatrix}, \qquad \Delta_3 = \begin{vmatrix} a_1 & a_3 & a_5 \\ a_0 & a_2 & a_4 \\ 0 & a_1 & a_3 \end{vmatrix}, \quad ...,$$

$$\Delta_n = \begin{vmatrix} a_1 & a_3 & a_5 & a_7 & ... & 0 \\ a_0 & a_2 & a_4 & a_6 & ... & 0 \\ 0 & a_1 & a_3 & a_5 & ... & 0 \\ 0 & a_0 & a_2 & a_4 & ... & 0 \\ 0 & 0 & a_1 & a_3 & ... & 0 \\ ... & ... & ... & ... & ... & ... \\ 0 & 0 & 0 & 0 & ... & a_n \end{vmatrix}. \tag{16.33}$$

For example, in the case of a third-order equation the basic array is

$$\begin{bmatrix} a_1 & a_3 & 0 \\ a_0 & a_2 & 0 \\ 0 & a_1 & a_3 \end{bmatrix}$$

and only $\Delta_1, \Delta_2, \Delta_3$ have to be considered. Thus the stability conditions are

$$\begin{aligned} a_1 &> 0, \\ a_1a_2 - a_0a_3 &> 0, \\ a_3(a_1a_2 - a_0a_3) &> 0. \end{aligned} \tag{16.34}$$

The third inequality, given the second, is equivalent to $a_3 > 0$, so that we can rewrite the stability conditions as

$$\begin{aligned} a_1 &> 0, \\ a_1a_2 - a_0a_3 &> 0, \\ a_3 &> 0, \end{aligned} \tag{16.35}$$

which are the same as (16.29) with the second inequality eliminated. Note now that, since $a_0 > 0$, the three inequalities in (16.35) together imply $a_2 > 0$: the second inequality, in fact, can be satisfied only when $a_2 > 0$. If

we prefer, the first inequality in (16.35) can be replaced with $a_2 > 0$, since the inequalities

$$\begin{aligned} a_2 &> 0, \\ a_1a_2 - a_0a_3 &> 0, \\ a_3 &> 0, \end{aligned} \tag{16.36}$$

imply $a_1 > 0$. Another way of noting the same thing is to state the stability conditions in the form (16.29) and to note that either the first or the second inequality can be eliminated, since either one is implied by the remaining three.

16.4.2 Necessary and sufficient stability conditions (Liénard-Chipart)

Necessary and sufficient conditions for Eq. (16.30) to have only roots with negative real parts may be expressed in any of the four following alternative forms:

$$\begin{array}{lrl} \text{(a)} & a_n > 0, a_{n-2} > 0, ...; & \Delta_1 > 0, \Delta_3 > 0, ..., \\ \text{(b)} & a_n > 0, a_{n-2} > 0, ...; & \Delta_2 > 0, \Delta_4 > 0, ..., \\ \text{(c)} & a_n > 0, a_{n-1} > 0, a_{n-3} > 0, ...; & \Delta_1 > 0, \Delta_3 > 0, ..., \\ \text{(d)} & a_n > 0, a_{n-1} > 0, a_{n-3} > 0, ...; & \Delta_2 > 0, \Delta_4 > 0, ..., \end{array} \tag{16.37}$$

where the Δ's are as defined in Eq. (16.33).

The advantage of the Liénard-Chipart conditions is that they involve about half many determinantal inequalities as the latter, and this is an important simplification, especially when the degree of the characteristic equation is high (and so the order of the Δ's is correspondingly high).

In either form, however, the stability conditions become increasingly complicated as the order of the equation increases and, correspondingly, their economic interpretation becomes more and more difficult. Indeed, there is not much hope of extricating a clear economic meaning from the stability conditions (except in particular cases) when the order of the equation is even moderately high (say four or more).

16.5 Exercises

16.5.1 Example

Consider the following non-homogeneous equation

$$\begin{aligned} y''' + 2y'' - y' - 2y &= e^t + t^2, \\ y(0) &= -7/4, \quad y'(0) = -8/6, \quad y''(0) = 16/3. \end{aligned}$$

The characteristic equation of the homogeneous part is

$$\lambda^3 + 2\lambda^2 - \lambda - 2 = 0.$$

Application of the stability conditions (16.29) shows that they are not satisfied. This means that the movement will be unstable.

Numerically, the characteristic equation clearly has the root $\lambda = 1$. Hence we can divide by $\lambda - 1$ and obtain

$$\lambda^3 + 2\lambda^2 - \lambda - 2 = (\lambda - 1)(\lambda^2 + 3\lambda + 2) = 0,$$

so that the remaining two roots will be given by solving $\lambda^2 + 3\lambda + 2 = 0$, whence $\lambda_2, \lambda_3 = -1, -2$. Hence the general solution of the homogeneous part is

$$y(t) = A_1 e^t + A_2 e^{-t} + A_3 e^{-2t}.$$

As regards a particular solution of the non-homogeneous equation, we note that the given functions are standard types, hence we use the method of undetermined coefficients. However, since one of them coincides with one of the terms of the solution of the homogeneous equation, we immediately begin with the same function multiplied by t. Thus the function that we try as a particular solution is

$$\overline{y}(t) = c_1 t^2 + c_2 t + c_3 + c_4 t e^t.$$

We have

$$\begin{aligned} \overline{y}'(t) &= 2c_1 t + c_2 + c_4 e^t + c_4 t e^t, \\ \overline{y}''(t) &= 2c_1 + 2c_4 e^t + c_4 t e^t, \\ \overline{y}'''(t) &= 3c_4 e^t + c_4 t e^t, \end{aligned}$$

and so, substituting into the non-homogeneous equation we have

$$\begin{aligned} 3c_4 e^t + c_4 t e^t + 4c_1 + 4c_4 e^t + 2c_4 t e^t - 2c_1 t - c_2 - c_4 e^t - c_4 t e^t & \\ -2c_1 t^2 - 2c_2 t - 2c_3 - 2c_4 t e^t &= e^t + t^2. \end{aligned}$$

Collecting terms we get

$$(-2c_1 - 1)t^2 - 2(c_1 + c_2)t + (4c_1 - c_2 - 2c_3) + (6c_4 - 1)e^t = 0,$$

that will be identically satisfied if, and only if

$$\begin{aligned} -2c_1 - 1 &= 0, \\ c_1 + c_2 &= 0, \\ 4c_1 - c_2 - 2c_3 &= 0, \\ 6c_4 - 1 &= 0. \end{aligned}$$

Thus we get $c_1 = -1/2, c_2 = 1/2, c_3 = -5/4, c_4 = 1/6$. The general solution of the non-homogenous equation is then

$$y(t) = A_1 e^t + A_2 e^{-t} + A_3 e^{-2t} - \frac{1}{2}t^2 + \frac{1}{2}t - \frac{5}{4} + \frac{1}{6}te^t.$$

To determine the arbitrary constants we have the equations

$$\begin{aligned} y(0) &= A_1 + A_2 + A_3 - 5/4 &&= 7/4, \\ y'(0) &= A_1 - A_2 - 2A_3 + 4/6 &&= -8/6, \\ y''(0) &= A_1 + A_2 + 4A_3 - 2/3 &&= 16/3, \end{aligned}$$

which gives $A_1 = A_2 = A_3 = 1$.

16.5.2 Other exercises

1. Solve the following homogeneous equations:
   (i) $y''' - 2y' + 4y = 0$ (Hint: one root is -2),
   (ii) $y''' - 4y'' + 3y = 0$,
   (iii) $y^{iv} + 2y''' - 3y'' = 0$.

2. Respectively add the following given functions of time to the homogeneous equations in exercise 1, and solve the resulting non-homogeneous equations:
   (i) $g(t) = t^2 + 2t - 1$,
   (ii) $g(t) = t^4 + 3t^2 - 5t + 2$,
   (iii) $g(t) = t^2 + 3e^{2t} + 4\sin t$.

16.6 References

Barnett, S., 1973, New Reductions of Hurwitz Determinants.
Barnett, S., 1983, *Polynomials and Linear Control Systems,* Chap. 3, Sect. 3.2.1.
Gantmacher, F.R., 1959, *Applications of the Theory of Matrices,* Chap. 5, Sects. 6, 13.
Ince, E.L., 1956, *Ordinary Differential Equations,* Chap. V, Sect. 5.23; Chap. VI.
Samuelson, P.A., 1947, *Foundations of Economic Analysis,* Appendix B, §36.
Turnbull, H.W., 1957, *Theory of Equations,* pp. 66, 68, 102; Sect. 52.

Chapter 17

Higher-order Differential Equations in Economic Models

We have already explained in Chap. 15 why it is difficult to find examples of 'direct' application of higher-order equations in economic models. The model that we are going to illustrate is actually the result of the transformation of a system.

17.1 Feedback control and stabilisation policies

17.1.1 Introduction

In the past decades the application of control theory to economics has emerged as an important field of research. Control theory has been developed not only by mathematicians, but also (and primarily) by mechanical and electronic engineers with practical problems in mind. This theory has many aspects, and here we just mention the classical method of feedback control. For a complete survey see Petit (1990), where a way of overcoming the Lucas critique is also shown.

Unlike the more recent theory of optimal control (see Sect. 22.1.2), in which there is some kind of objective function to be optimised dynamically, and the various targets admit of a trade-off, feedback control is concerned with fixed targets (these can be a given point or a given dynamic path) and defines control actions that are functions of the controlled variables or, more precisely, of their deviations from the given respective target. The purpose of feedback control is to stabilise the behaviour of the controlled variables with reference to their target values. Therefore, a stabilising control is a control that either makes stable an otherwise unstable system, or makes 'more stable' a system that is already stable.

The qualification 'more stable' is rather loose, and refers generically to a better dynamic performance of the model, such as a larger interval for stability or a smaller margin for oscillations, or a path to equilibrium that is closer to the equilibrium point (or path). To be precise, if damped oscillations around the equilibrium were present, a more stable behaviour means more heavily damped oscillations or even elimination of the oscillations. If the variables were monotonically convergent to equilibrium, more stable means a faster convergence.

Although the suggestion of applying feedback policy rules to stabilise an economic system had already been set forth a couple of years before (see Petit, 1990, p. 68), it was effectively implemented by Phillips (1954) in a pioneering contribution that was not followed up by economists at that time, mainly because it was too advanced for the period. Here we shall describe this contribution.

17.1.2 Three types of stabilisation policy

The level of national output (income) is determined, as we know from elementary macroeconomics, by the level of aggregate demand. The latter is made up of a part originating from private economic agents and a part originating from the government. The government manoeuvres its expenditure in order that aggregate demand is such that national income attains a given desired level. Thus we can talk of stabilisation of aggregate demand. Phillips analyses this economic policy problem from the point of view of the principle of servomechanisms, i.e. in a feedback control context.

Let us assume that national income is initially at the desired level and that an exogenous decrease in aggregate demand occurs. The variable are measured as deviations from their desired levels, so that a negative value simply means that the actual value is smaller than the desired value. Before introducing the stabilisation policy we must know, to have a standard of comparison, the 'spontaneous' behaviour of the economic system, i.e. its behaviour when there is no government expenditure.

The basic endogenous dynamic mechanism is the multiplier[1], where it is assumed that producers react to excess demand by making adjustments in output: if aggregate demand exceeds (falls short of) current output, the latter will be increased (decreased). Obviously this mechanism operates independently of the origin of excess demand, and so is the same both without and with government expenditure. In formal terms

$$Y' = \alpha(D - Y), \qquad \alpha > 0, \tag{17.1}$$

where Y is national output, D is total aggregate demand and α is a reaction coefficient, representing the velocity of adjustment to a discrepancy between and aggregate demand and current output.

[1]Phillips also considers the case in which the basic dynamic mechanism is of the multiplier-accelerator type. See below, exercise 4.

When the government is absent, total aggregate demand coincides with aggregate private demand, which is a function of national income:

$$D = (1 - l)Y, \tag{17.2}$$

where $(1 - l)$ is the marginal propensity to spend of the private sector (i.e. the marginal propensity to consume plus the marginal propensity to invest). This propensity is assumed to be smaller than unity, whence $0 < l < 1$. Introducing the exogenous disturbance u we have

$$D = (1 - l)Y - u. \tag{17.3}$$

Substituting (17.3) into (17.1), and giving u a unit value, we have

$$Y' + \alpha l Y = -\alpha. \tag{17.4}$$

This is a first order equation, whose solution is

$$Y(t) = Ae^{-\alpha l t} - \frac{1}{l}. \tag{17.5}$$

We assumed that in the initial period national income was at the desired level, so that $Y(0) = 0$ and so $A = 1/l$, whence

$$Y(t) = -\frac{1}{l}(1 - e^{-\alpha l t}). \tag{17.6}$$

Since $\alpha l > 0, Y(t)$ tends monotonically to $-1/l$, i.e. to the value obtained applying the multiplier $1/l$ to the exogenous decrease in expenditure (-1).

We may now go on to examine the effects of a stabilisation policy. Phillips enumerates three types of stabilisation policy. They are the following:

(i) *Proportional stabilisation policy*: government expenditure is proportional and of opposite sign to the deviation between the actual and the desired value of output, i.e. $G^* = -f_p Y$, where $f_p > 0$ is the coefficient of proportionality.

(ii) *Derivative stabilisation policy:* government expenditure is proportional and of opposite sign to the variation in (that is to the derivative of) current output, i.e. $G^* = -f_d Y'$, where $f_d > 0$ is the coefficient of proportionality.

(iii) *Integral stabilisation policy*: government expenditure is proportional and opposite to the sign of the sum (in continuous term to the integral) of all the differences that have occurred, from time zero to the current moment, between the actual and the desired values of output, i.e. $G^* = -f_i \int_0^t Y \, dt$, where $f_i > 0$ is the coefficient of proportionality.

We have marked government demand with an asterisk. The reason is that the various values of such demand indicated the enumeration of the various policies are the *theoretical* or *potential* values, i.e. the values that define in theory the different policies. Now in Phillips' words (Phillips, 1954, p. 294) 'the actual policy demand will be different from the potential policy demand, owing to the time required for observing changes in error, adjusting the correcting action accordingly and for changes in the correcting action to produce their full effects.... whenever such a difference exists the actual policy demand will be changing in a direction which tends to eliminate the difference and at a rate proportional to the difference'. Thus using the symbol G to indicate *actual* government demand, we have the partial adjustment equation (see Chap. 12, Sect. 12.4)

$$G' = \beta(G^* - G), \qquad \beta > 0, \tag{17.7}$$

where β is a reaction coefficient, indicating the speed of response to a discrepancy between potential and actual government expenditure.

When government demand is present, equation (17.3) becomes

$$D = (1-l)Y + G - u. \tag{17.8}$$

The stabilisation model consists of equations (17.1), (17.7), (17.8), and of one or more of the relations defining G^*. We now manipulate the model to reduce it to a single equation.

Substituting Eq. (17.8) into Eq. (17.1) and rearranging terms we have

$$Y' + \alpha l Y + \alpha u = \alpha G, \tag{17.9}$$

from which, differentiating with respect to time,

$$Y'' + \alpha l Y' = \alpha G'. \tag{17.10}$$

From (17.7) we have, multiplying through by α,

$$\alpha G' + \alpha\beta G = \alpha\beta G^*. \tag{17.11}$$

We can now substitute Eqs. (17.9) and (17.10) into Eq. (17.11). This yields, after rearranging terms and considering a unit decrease in aggregate demand,

$$Y'' + (\alpha l + \beta)Y' + \alpha\beta l Y - \alpha\beta G^* = -\alpha\beta, \tag{17.12}$$

which is the basic differential equation of the model. Inserting the various relation defining G^* we can determine the time path of Y and so study the effect of the single stabilisation policies or combinations of them, when two or more are used simultaneously.

Although only integral stabilisation policy gives rise to a higher-order differential equation, the study of proportional and derivative stabilisation policies (that give rise to second-order equations) is a necessary prerequisite. Thus we shall begin by examining the 'pure proportional' case and the 'mixed proportional-derivative' case.

17.1.2.1 Proportional stabilisation policy

Inserting $G^* = -f_p Y$ in Eq. (17.12) and collecting terms we have

$$Y'' + (\alpha l + \beta)Y' + \alpha\beta(l + f_p)Y = -\alpha\beta. \tag{17.13}$$

A particular solution is

$$\bar{Y} = -\frac{1}{l + f_p}. \tag{17.14}$$

The characteristic equation of the homogeneous part of (17.13) is

$$\lambda^2 + (\alpha l + \beta)\lambda + \alpha\beta(l + f_p) = 0. \tag{17.15}$$

The succession of the signs of the coefficients is $+ + +$, so that stability is ensured. To determine whether the movement will be monotonic or oscillatory we must examine the discriminant

$$\Delta = (\alpha l + \beta)^2 - 4\alpha\beta(l + f_p), \tag{17.16}$$

and it easy to see that the greater f_p is, the more likely it is that $\Delta < 0$. More precisely,

$$\Delta \gtreqless 0 \quad \text{according to} \quad f_p \lesseqgtr \frac{(\alpha l - \beta)^2}{4\alpha\beta}. \tag{17.17}$$

Let us recall that the limit value of output in the absence of any stabilisation policy is $-1/l$; since $f_p > 0$, the new limit value $-1/(l + f_p)$ is smaller in absolute value. This means that the decrease in income determined by an exogenous decrease in aggregate demand is smaller than the decrease occurring without the stabilisation policy. The policy under consideration, however, cannot completely eliminate this decrease (an infinite value of f_p is not possible). Let us also note that the greater f_p is, the more effective is the stabilisation policy (i.e. the smaller is the absolute value of $-1/(l + f_p)$).

In Fig. 17.1—taken from Phillips—curve (a) is the time path of income in the absence of stabilisation, curve (b) holds for f_p =0.5 and curve (c) for $f_p = 2$. In all cases $\alpha = 4, l = 0.25$; the value of β is 2 and the initial conditions are $Y(0) = 0$ for curve (a) and $Y(0) = 0, Y'(0) = -4$ for curves (b) and (c).

In conclusion the drawbacks of the purely proportional stabilisation policy are two:

(a) it fails to eliminate the reduction in income completely;

(b) it tends to provoke oscillations (which could not occur in the absence of the policy under consideration). Although damped, these oscillations occur when f_p is too great, and this contrasts with the desirability of having as great an f_p as possible, in order to minimize the reduction in income.

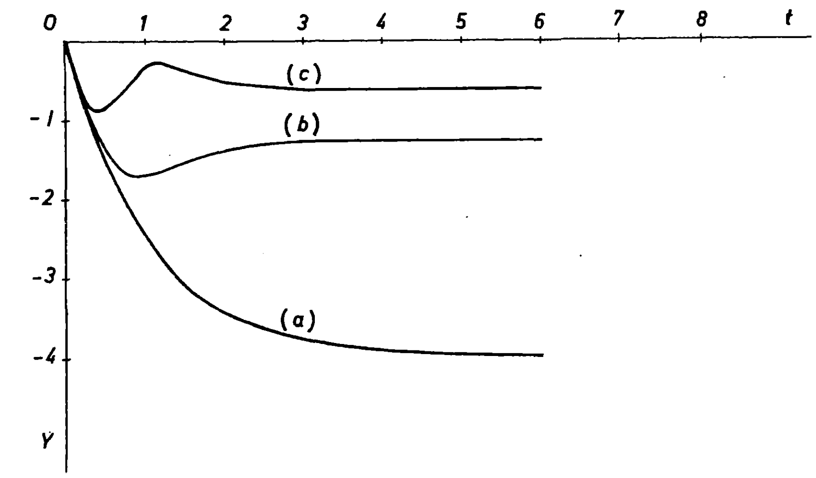


Figure 17.1: Proportional stabilisation policy in the Phillips model

17.1.2.2 Mixed proportional-derivative stabilisation policy

In this case the two policies are used simultaneously, so that in Eq. (17.12) we must substitute $G^* = -f_p Y - f_d Y'$, obtaining

$$Y'' + (\alpha l + \beta + \alpha\beta f_d)Y' + \alpha\beta(l + f_p)Y = -\alpha\beta. \tag{17.18}$$

A particular solution is

$$\bar{Y} = -\frac{1}{l + f_p}. \tag{17.19}$$

The characteristic equation of the homogeneous part of (17.18) is

$$\lambda^2 + (\alpha l + \beta + \alpha\beta f_d)\lambda + \alpha\beta(l + f_p) = 0. \tag{17.20}$$

Here too the succession of signs of the coefficients is + + +, so that as t increases $Y(t)$ will converge to the particular solution (17.19). The oscillatory or monotonic behaviour of the time path depends on the sign of

$$\Delta = (\alpha l + \beta + \alpha\beta f_d)^2 - 4\alpha\beta(l + f_p). \tag{17.21}$$

We now have all the elements to compare the mixed policy with the pure proportional policy. The first thing to note is that (17.19) is the same as (17.14): the addition of the derivative policy has no effect as far as the reduction in income is concerned. The effect of this addition can be seen by examining Δ: the greater f_d, the greater the positive term in Δ and the less

likely it is that $\Delta < 0$. Thus f_d offsets the bias of f_p toward oscillations. Moreover, even if oscillations should occur, they would be more heavily damped than in case of a purely proportional policy. In fact, as we know, it is the real part of the complex roots that gives the damping: the greater the absolute value of this real part, the heavier is the damping.

Now, the real part of the complex roots, when they occur, is $-\frac{1}{2}(\alpha l + \beta)$ in the case of the purely proportional policy and $-\frac{1}{2}(\alpha l + \beta + \alpha\beta f_d)$ in the case of the mixed proportional-derivative policy. The latter expression is obviously greater in absolute value than the former. We shall presently see that the general property of the derivative stabilisation policy is to offset the oscillatory bias of other policies.

17.1.2.3 Integral stabilisation policy

Substituting $G^* = -f_i \int_0^t Y\, dt$ in Eq. (17.12) we get

$$Y'' + (\alpha l + \beta)Y' + \alpha\beta l Y + \alpha\beta f_i \int_0^t Y dt = -\alpha\beta. \tag{17.22}$$

To eliminate the integral we differentiate with respect to time, thus obtaining

$$Y''' + (\alpha l + \beta)Y'' + \alpha\beta l Y' + \alpha\beta f_i Y = 0. \tag{17.23}$$

The first thing to observe is that (17.23) is a homogeneous equation, so that the movement of Y is referred to zero (the target). Therefore, if we find that this movement is stable, it follows that the integral stabilisation policy succeeds where the other policies fail, namely in completely eliminating the decrease in income due to an exogenous decrease in aggregate demand.

The characteristic equation of (17.23) is

$$\lambda^3 + (\alpha l + \beta)\lambda^2 + \alpha\beta l\lambda + \alpha\beta f_i = 0. \tag{17.24}$$

Since all the coefficients are positive, no positive real root can exist. Thus instability, if any, can only take the form of undamped oscillations. If we apply the stability conditions (see Eq. (16.29) in Chap. 16) we find that the crucial stability condition is

$$(\alpha l + \beta)\alpha\beta l - \alpha\beta f_i > 0, \quad \text{or } f_i < (\alpha l + \beta)l. \tag{17.25}$$

It follows that integral stabilisation policy is successful provided that the policy parameter f_i is smaller than the critical value $(\alpha l + \beta)l$. In the contrary case this policy destabilises output, causing undamped oscillations.

Thus an integral stabilisation policy on the one hand has the advantage of completely eliminating the reduction in output caused by an exogenous disturbance, a result that could not be achieved with the other policies. On the other hand, it gives rise to the risk of an unstable movement, a risk that was not present with the other policies. These are interesting results, and it should be noted that we could not have reached them without the use of the stability conditions, which gave rise to inequality (17.25).

17.2 Exercises

1. Consider the joint use of integral and proportional stabilisation policy, and show that the addition of the proportional policy reduces the risk of explosive oscillations with respect to the pure integral policy.

2. Consider the use of the derivative policy together with the integral policy, and show that the danger of explosive oscillations is reduced with respect to the pure integral policy.

3. Show that the simultaneous adoption of all three policies is the best solution.

4. Suppose that, in Phillips' model, the basic dynamic mechanism is of the multiplier-accelerator type. Desired investment is proportional to the rate of change in output, $I^* = kY'$, and entrepreneurs react to a discrepancy between desired and actual investment according to the partial adjustment equation $I' = \eta(I^* - I)$, where η is the speed of adjustment. Total private demand is now $D = (1-l)Y + I - u$.

(4.a) Show that in the case of a pure proportional or mixed proportional-derivative policy the resulting differential equation is of the third order, and analyse the stability of the model.

(4.b) Show that the introduction of an integral policy gives rise to a fourth-order equation. Apply the stability conditions and discuss the results.

5. Consider the following continuous time reformulation of Metzler's inventory cycle model (see Chap. 8, Sect. 8.1 for the original discrete-time model). We know from elementary macroeconomics that a discrepancy between ex ante saving and investment gives rise to an equivalent variation in inventories, $Q'(t) = S(t) - I(t)$, where $Q(t)$ is the stock of inventories. Assume that $S(t) = sY(t), I(t) = I_0e^{gt}$. The inventory change induces firms to change the level of current output, assumed to be initially in equilibrium, so as to bring inventories back to their desired or equilibrium level $\hat{Q}(t)$, namely $Y'(t) = \gamma[\hat{Q}(t) - Q(t)]$. Suppose that $\hat{Q}(t)$ is proportional to *expected* output, $\hat{Q}(t) = k\hat{Y}(t)$. Finally assume that expectations are formed according to the (admittedly ad hoc) mechanism $\hat{Y}(t) = Y(t) + a_1Y'(t) + a_2Y''(t)$.

Solve the resulting non-homogeneous equation and examine the stability of the equilibrium path. (Hint: to eliminate the variable Q, differentiate $Y'(t) = \gamma[\hat{Q}(t) - Q(t)]$ with respect to time and observe that $\hat{Q}'(t) = k\hat{Y}'(t)$).

17.3 References

Petit, M.L., 1990, *Control Theory and Dynamic Games in Economic Policy Analysis.*

Phillips, A.W., 1954, Stabilisation Policy in a Closed Economy.

Phillips, A.W., 1957, Stabilisation Policy and the Time-Form of Lagged Responses.

Chapter 18

Simultaneous Systems of Differential Equations

A simultaneous system is made up of two (or more) differential equations in which two (or more) unknown functions are involved. For the system to be solvable, it must have as many equations as unknowns, provided that the equations are independent and consistent.

18.1 First-order 2×2 systems in normal form

The simplest type of system is the following first-order system in 'normal' form

$$\begin{aligned} y'(t) &= a_{11}y(t) + a_{12}z(t) + g_1(t), \\ z'(t) &= a_{21}y(t) + a_{22}z(t) + g_2(t), \end{aligned} \tag{18.1}$$

where the coefficients a_{ij} are given constants and $g_1(t), g_2(t)$ are known functions. The system is first-order because only first-order derivatives appear, and is called 'normal' because each equation involves the derivative of only one unknown function in turn. It must be noted that the *order* of a system is equal to the *degree* of the characteristic equation (see below). Therefore, the expression 'first-order system' must be understood only in the sense defined above, and not as indicating the order of the system (actually the order of system (18.1) is 2).

System (18.1) is a non-homogeneous system. It can be proved by direct substitutions (which are left as an exercise) that, as in single equations, the general solution of (18.1) is obtained adding a particular solution to the general solution of the corresponding homogeneous system. So let us begin with the latter.

18.1.1 General solution of the homogeneous system: first method

Let us consider the homogenous system corresponding to (18.1),

$$\begin{aligned} y'(t) &= a_{11}y(t) + a_{12}z(t), \\ z'(t) &= a_{21}y(t) + a_{22}z(t). \end{aligned} \tag{18.2}$$

A method of solving (18.2) consists of reducing it to a single equation in which only one unknown function appears; this reduction is always possible by suitable transformations. From the first equation of system (18.2) we obtain, provided that $a_{12} \neq 0$,

$$z = \frac{1}{a_{12}}y' - \frac{a_{11}}{a_{12}}y, \tag{18.3}$$

whence, differentiating with respect to time,

$$z' = \frac{1}{a_{12}}y'' - \frac{a_{11}}{a_{12}}y'. \tag{18.4}$$

If $a_{12} = 0$, we use the second equation to isolate $y(t)$, etcetera. The steps are the same as those used above for $z(t)$. Note that, when $a_{12} = 0$, a_{21} must be different from zero, since, if also $a_{21} = 0$, we would no longer have a simultaneous system, but two separate equations with no interdependency. Another method when either a_{12} or a_{21} is zero is to solve separately the equation where only one unknown function appears and to substitute the result in the other equation.

Substituting (18.3) and (18.4) in the second equation of system (18.2), we have

$$\frac{1}{a_{12}}y'' - \frac{a_{11}}{a_{12}}y' = a_{21}y + \frac{a_{22}}{a_{12}}y' - \frac{a_{11}a_{22}}{a_{12}}y. \tag{18.5}$$

Multiplying both members by a_{12} and rearranging terms, we obtain

$$y'' - (a_{11} + a_{22})y' - (a_{12}a_{21} - a_{11}a_{22})y = 0. \tag{18.6}$$

Thus we have a second-order differential equation in which only one unknown function, y, appears. We solve it in the usual way, and to obtain z we only have to substitute the solution in Eq. (18.3). Suppose for example that the roots of the characteristic equation of (18.6) are real and distinct. Then the general solution of (18.6) is, as we know,

$$y(t) = A_1e^{\lambda_1 t} + A_2e^{\lambda_2 t}, \tag{18.7}$$

where A_1 and A_2 are arbitrary constants. Substituting (18.7) in (18.3) we have

$$z(t) = \frac{\lambda_1 A_1}{a_{12}}e^{\lambda_1 t} + \frac{\lambda_2 A_2}{a_{12}}e^{\lambda_2 t} - \frac{a_{11}A_1}{a_{12}}e^{\lambda_1 t} - \frac{a_{11}A_2}{a_{12}}e^{\lambda_2 t}.$$

Collecting terms we obtain

$$z(t) = \frac{\lambda_1 - a_{11}}{a_{12}}A_1e^{\lambda_1 t} + \frac{\lambda_2 - a_{11}}{a_{12}}A_2e^{\lambda_2 t}. \tag{18.8}$$

Eqs.(18.7) and (18.8) are the general solution of system (18.2); of course, if the roots of the characteristic equation (18.6) are real and equal or are complex, we shall proceed as explained in Chap. 14 to obtain $y(t)$, and then substitute in (18.3) to obtain $z(t)$. If such roots are real and equal $(\lambda_1 = \lambda_2 = \lambda^*)$, the solution for $y(t)$ is, as we know,

$$y(t) = (A_1 + A_2t)e^{\lambda^* t}, \tag{18.9}$$

from which

$$y' = A_2e^{\lambda^* t} + \lambda^*(A_1 + A_2t)e^{\lambda^* t}. \tag{18.10}$$

Substituting Eqs. (18.9) and (18.10) in (18.3), we have

$$\begin{aligned} z(t) &= \frac{1}{a_{12}}[A_2 + \lambda^*(A_1 + A_2t)]e^{\lambda^* t} - \frac{a_{11}}{a_{12}}(A_1 + A_2t)e^{\lambda^* t} \\ &= e^{\lambda^* t}\left[\frac{\lambda^*(A_1 + A_2t) + A_2}{a_{12}} - \frac{a_{11}}{a_{12}}(A_1 + A_2t)\right]; \end{aligned}$$

therefore, by rearranging terms,

$$z(t) = \left[\frac{(\lambda^* - a_{11})A_1 + A_2}{a_{12}} + \frac{\lambda^* - a_{11}}{a_{12}}A_2t\right]e^{\lambda^* t}. \tag{18.11}$$

Using the fact that $\lambda^* = \frac{1}{2}(a_{11} + a_{22})$, Eq. (18.11) may also be written as

$$z(t) = \left[\frac{(a_{22} - a_{11})A_1 + 2A_2}{2a_{12}} + \frac{a_{22} - a_{11}}{2a_{12}}A_2t\right]e^{\lambda^* t}. \tag{18.12}$$

If, finally, the roots of the characteristic equation of (18.6) are complex, say $\lambda_1, \lambda_2 = \alpha \pm i\theta$, then oscillations will arise, as we know from the treatment of second-order equations, namely

$$y(t) = e^{\alpha t}(A_1 \cos\theta t + A_2 \sin\theta t), \tag{18.13}$$

from which

$$\begin{aligned} y'(t) &= \alpha e^{\alpha t}(A_1 \cos\theta t + A_2 \sin\theta t) + e^{\alpha t}(-\theta A_1 \sin\theta t + \theta A_2 \cos\theta t) \\ &= e^{\alpha t}[(\alpha A_1 + \theta A_2)\cos\theta t + (\alpha A_2 - \theta A_1)\sin\theta t]. \end{aligned} \tag{18.14}$$

Substituting Eqs. (18.13) and (18.14) in (18.3) we have

$$\begin{aligned} z(t) &= \frac{e^{\alpha t}\left[(\alpha A_1 + \theta A_2)\cos\theta t + (\alpha A_2 - \theta A_1)\sin\theta t\right]}{a_{12}} \\ &\quad + \frac{-a_{11}e^{\alpha t}(A_1\cos\theta t + A_2\sin\theta t)}{a_{12}}, \end{aligned}$$

whence, rearranging terms,

$$z(t) = e^{\alpha t}\left[\frac{(\alpha - a_{11})A_1 + \theta A_2}{a_{12}}\cos\theta t + \frac{(\alpha - a_{11})A_2 - \theta A_1}{a_{12}}\sin\theta t\right]. \tag{18.15}$$

The method that we have illustrated is fairly simple and yields the required solution. There is, however, another method, which at first sight might look more complicated, but has the advantage of being more direct, in the sense that it simultaneously gives both unknown functions $y(t)$ and $z(t)$ without any need to reduce the system to a single equation in one unknown function. Such a method, moreover, can easily be generalized.

18.1.2 General solution of the homogeneous system: second (or direct) method

The alternative method consists in directly trying as a solution—by analogy with single equations—the functions $y(t) = \alpha_1 e^{\lambda t}$, $z(t) = \alpha_2 e^{\lambda t}$, where α_1, α_2 are constants not both zero. Substituting in (18.2) we have

$$\begin{aligned} \alpha_1\lambda e^{\lambda t} &= a_{11}\alpha_1 e^{\lambda t} + a_{12}\alpha_2 e^{\lambda t}, \\ \alpha_2\lambda e^{\lambda t} &= a_{21}\alpha_1 e^{\lambda t} + a_{22}\alpha_2 e^{\lambda t}, \end{aligned}$$

from which

$$\begin{aligned} e^{\lambda t}\left[(a_{11} - \lambda)\alpha_1 + a_{12}\alpha_2\right] &= 0, \\ e^{\lambda t}\left[a_{21}\alpha_1 + (a_{22} - \lambda)\alpha_2\right] &= 0. \end{aligned} \tag{18.16}$$

The functions that we have tried will be the solution of system (18.2) if, and only if, system (18.16) is satisfied for any t, i.e.(apart from the trivial case $\lambda = 0$) if, and only if,

$$\begin{aligned} (a_{11} - \lambda)\alpha_1 + a_{12}\alpha_2 &= 0, \\ a_{21}\alpha_1 + (a_{22} - \lambda)\alpha_2 &= 0. \end{aligned} \tag{18.17}$$

System (18.17) has the trivial solution $\alpha_1 = \alpha_2 = 0$, but we have excluded it from the beginning for obvious reasons. From elementary algebra we know that the necessary and sufficient condition for a linear and homogeneous

system to have non-trivial solutions, in addition to the trivial one, is that the determinant of the system be zero. In our case, then, it must be

$$\begin{vmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{vmatrix} = 0. \tag{18.18}$$

Expanding the determinant we have

$$\begin{aligned} &(a_{11} - \lambda)(a_{22} - \lambda) - a_{12}a_{21} \\ &= \lambda^2 - (a_{11} + a_{22})\lambda + (a_{11}a_{22} - a_{12}a_{21}) = 0. \end{aligned} \tag{18.19}$$

The determinantal equation (18.18) and its expanded form (18.19) are called the characteristic equation of the system of differential equations (18.2). Let us note that such an equation is the same as the characteristic equation of (18.6) above, and this is correct since the λ values must be the same whichever method is followed to solve the system of differential equations.

18.1.2.1 Unequal real roots

From the solution of (18.19) we obtain two values of λ; let us assume for the moment that they are real and distinct. Thus the determinant of the system (18.17) equals zero for $\lambda = \lambda_1$ and for $\lambda = \lambda_2$; correspondingly, we shall have two solutions of that system. Let us call $[\alpha_1^{(1)}, \alpha_2^{(1)}]$ the solution that we obtain putting $\lambda = \lambda_1$ in (18.17), and $[\alpha_1^{(2)}, \alpha_2^{(2)}]$ the solution that we have putting $\lambda = \lambda_2$. For $\lambda = \lambda_1$ we have

$$\begin{aligned} (a_{11} - \lambda)\alpha_1^{(1)} + a_{12}\alpha_2^{(1)} &= 0, \\ a_{21}\alpha_1^{(1)} + (a_{22} - \lambda)\alpha_2^{(1)} &= 0. \end{aligned} \tag{18.20}$$

From elementary algebra we know that, since the determinant of the system is zero, we can fix the value of one of the unknowns arbitrarily and then determine the value of the other (in other words, only the ratio between the two unknowns is determined). Such arbitrariness does not give any trouble in the solution of the system (18.2), since it combines in a multiplicative way with the arbitrary constants A_1, A_2 which appear in the general solution. Thus we choose $\alpha_1^{(1)} = 1$ (and similarly we shall choose $\alpha_1^{(2)} = 1$) so that the solution of system (18.2) that we shall obtain can be immediately compared, without any further manipulations, with (18.7) and (18.8) above. Setting $\alpha_1^{(1)} = 1$, from the first equation of (18.20) we have

$$\alpha_2^{(1)} = \frac{\lambda_1 - a_{11}}{a_{12}}.$$

Note that, from the second equation, $\alpha_2^{(1)} = a_{21}/(\lambda_1 - a_{22})$. The two values, however, are equal since λ_1 is a root of the characteristic equation, i.e.

$$(a_{11} - \lambda_1)(a_{22} - \lambda_1) - a_{12}a_{21} = 0,$$

so that

$$\frac{\lambda_1 - a_{11}}{a_{12}} = \frac{a_{21}}{\lambda_1 - a_{22}}.$$

In a similar way for $\lambda = \lambda_2$ we fix $\alpha_1^{(2)} = 1$ and obtain

$$\alpha_2^{(2)} = \frac{\lambda_2 - a_{11}}{a_{12}} = \frac{a_{21}}{\lambda_2 - a_{22}}.$$

Thus we have reached the result that $y(t) = \alpha_1^{(1)} e^{\lambda_1 t}$, $z(t) = \alpha_2^{(1)} e^{\lambda_1 t}$ is a solution of system (18.2), and that $y(t) = \alpha_1^{(2)} e^{\lambda_2 t}$, $z(t) = \alpha_2^{(2)} e^{\lambda_2 t}$ is another solution. We can then combine them linearly with two arbitrary constants A_1, A_2 and obtain the general solution of system(18.2):

$$y(t) = A_1 \alpha_1^{(1)} e^{\lambda_1 t} + A_2 \alpha_1^{(2)} e^{\lambda_2 t}, \tag{18.21}$$

$$z(t) = A_1 \alpha_2^{(1)} e^{\lambda_1 t} + A_2 \alpha_2^{(2)} e^{\lambda_2 t}, \tag{18.22}$$

i.e., with the values of $\alpha_i^{(j)}$, $i, j, = 1, 2$, found above,

$$y\,(t) = A_1 e^{\lambda_1 t} + A_2 e^{\lambda_2 t}, \tag{18.23}$$

$$z(t) = A_1 \frac{\lambda_1 - a_{11}}{a_{12}} e^{\lambda_1 t} + A_2 \frac{\lambda_2 - a_{11}}{a_{12}} e^{\lambda_2 t}, \tag{18.24}$$

which are the same as Eqs. (18.7) and (18.8). The student may check by direct substitution that (18.23) and (18.24) indeed satisfy system (18.2). The number of arbitrary constants appearing in the general solution of this system is two, since the system is reducible to a second-order equation, as we have seen above. In general, the number of arbitrary constants is equal to the *order* of the system, that is to the *degree* of its characteristic equation.

The solution of the system may also be written in the equivalent form, which sometimes appears in the literature,

$$\begin{aligned} y(t) &= A_1 e^{\lambda_1 t} + A_2 e^{\lambda_2 t}, \\ z(t) &= A_1' e^{\lambda_1 t} + A_2' e^{\lambda_2 t}, \end{aligned} \tag{18.25}$$

where the arbitrary constants A_1, A_2, A_1', A_2' are connected by the relations

$$\frac{A_1'}{A_1} = \frac{\lambda_1 - a_{11}}{a_{12}} = \frac{a_{21}}{\lambda_1 - a_{22}}, \qquad \frac{A_2'}{A_2} = \frac{\lambda_2 - a_{11}}{a_{12}} = \frac{a_{21}}{\lambda_2 - a_{22}}. \tag{18.26}$$

Since the ratios $A_1'/A_1, A_2'/A_2$ are uniquely determined in terms of the λ's and of the coefficients of the system, the independent arbitrary constants are actually only two. If we compare Eqs. (18.25) with Eqs. (18.21)-(18.22) we see that A_1'/A_1 must be equal to $A_1\alpha_1^{(1)}/A_1\alpha_2^{(1)}$, etcetera. This is indeed true, since $\alpha_1^{(1)}/\alpha_2^{(1)} = (\lambda_1 - a_{11})/a_{12} = a_{21}/(\lambda_1 - a_{22})$, etcetera.

18.1.2.2 Equal real roots

If the characteristic equation has two real and equal roots $\lambda_1 = \lambda_2 = \lambda^*$, let us try

$$\begin{aligned} y(t) &= (A_1 + A_2 t) e^{\lambda^* t}, \\ z(t) &= (A_1' + A_2' t) e^{\lambda^* t}, \end{aligned} \tag{18.27}$$

where the arbitrary constants A_1, A_2, A_1', A_2' are related in some way. Differentiating we obtain

$$\begin{aligned} y'(t) &= (\lambda^* A_1 + A_2 + \lambda^* A_2 t) e^{\lambda^* t}, \\ z'(t) &= (\lambda^* A_1' + A_2' + \lambda^* A_2' t) e^{\lambda^* t}. \end{aligned} \tag{18.28}$$

Substituting Eqs. (18.27) and (18.28) in system (18.2) we have

$$\begin{aligned} (\lambda^* A_1 + A_2 + \lambda^* A_2 t) e^{\lambda^* t} &= a_{11}(A_1 + A_2 t) e^{\lambda^* t} + a_{12}(A_1' + A_2' t) e^{\lambda^* t}, \\ (\lambda^* A_1' + A_2' + \lambda^* A_2' t) e^{\lambda^* t} &= a_{21}(A_1 + A_2 t) e^{\lambda^* t} + a_{22}(A_1' + A_2' t) e^{\lambda^* t}. \end{aligned} \tag{18.29}$$

Dividing through by $e^{\lambda^* t} \neq 0$ and rearranging terms we obtain

$$\begin{aligned} &\left[(\lambda^* - a_{11}) A_2 - a_{12} A_2'\right] t + \left[(\lambda^* - a_{11}) A_1 + A_2 - a_{12} A_1'\right] = 0, \\ &\left[(\lambda^* - a_{22}) A_2' - a_{21} A_2\right] t + \left[(\lambda^* - a_{22}) A_1' + A_2' - a_{21} A_1\right] = 0. \end{aligned} \tag{18.30}$$

Eqs. (18.30) are identically satisfied if, and only if, the expressions in square brackets are all zero, i.e.

$$A_2(\lambda^* - a_{11}) - a_{12} A_2' = 0, \tag{18.31}$$

$$(\lambda^* - a_{11}) A_1 + A_2 - a_{12} A_1' = 0, \tag{18.32}$$

$$(\lambda^* - a_{22}) A_2' - a_{21} A_2 = 0, \tag{18.33}$$

$$(\lambda^* - a_{22}) A_1' + A_2' - a_{21} A_1 = 0, \tag{18.34}$$

whence

$$A_2' = \frac{\lambda^* - a_{11}}{a_{12}} A_2, \tag{18.35}$$

$$A_1' = \frac{(\lambda^* - a_{11}) A_1 + A_2}{a_{12}}, \tag{18.36}$$

$$A_2' = \frac{a_{21}}{\lambda^* - a_{22}} A_2, \tag{18.37}$$

$$A_1' = \frac{a_{21}}{\lambda^* - a_{22}} A_1 - \frac{1}{\lambda^* - a_{22}} A_2'. \tag{18.38}$$

Since λ^* is a root of the characteristic equation, we have $(\lambda^* - a_{11})/a_{12} = a_{21}(\lambda^* - a_{22})$, and so (18.35) and (18.37) coincide. Using (18.35) and the fact that $(\lambda^* - a_{11})/a_{12} = a_{21}/(\lambda^* - a_{22})$, Eq. (18.38) can be written as

$$A_1' = \frac{\lambda^* - a_{11}}{a_{12}} A_1 - \frac{\lambda^* - a_{11}}{a_{12}(\lambda^* - a_{22})} A_2, \tag{18.39}$$

which coincides with (18.36) if, and only if, $-(\lambda^* - a_{11}) = \lambda^* - a_{22}$, i.e. if and only if, $\lambda^* = \frac{1}{2}(a_{11} + a_{22})$, which is indeed true if, and only if, λ^* is a double root of the characteristic Eq. (18.19). Thus (18.27) is indeed the (general) solution of the system. Using (18.36) and (18.35) to express A_1', A_2' in terms of A_1, A_2, it can be seen that Eqs. (18.27) coincide with (18.9) and (18.11) above.

18.1.2.3 Complex roots

If the roots of the characteristic equation are complex, i.e. $\lambda_1, \lambda_2 = \alpha \pm i\theta$, then the solution can at first be written (the procedure is the same as for the case of distinct real roots) as

$$\begin{aligned} y(t) &= B_1 e^{(\alpha+i\theta)t} + B_2 e^{(\alpha-i\theta)t}, \\ z(t) &= \frac{(\alpha+i\theta)-a_{11}}{a_{12}} B_1 e^{(\alpha+i\theta)t} + \frac{(\alpha-i\theta)-a_{11}}{a_{12}} B_2 e^{(\alpha-i\theta)t}, \end{aligned} \tag{18.40}$$

where B_1, B_2 are arbitrary complex conjugate constants, say $A_1 \pm A_2 i$. Using standard transformations of complex numbers (see above, Chap. 14, Sect. 14.1.3), and noting that $(B_1 + B_2) \equiv A_1$, $(B_1 - B_2)i \equiv A_2$ are arbitrary real constants, we immediately obtain

$$y(t) = e^{\alpha t}(A_1 \cos\theta t + A_2 \sin\theta t). \tag{18.41}$$

As regards $z(t)$, we have

$$\begin{aligned} z(t) = {} & e^{\alpha t}\Big[\tfrac{(\alpha-a_{11})+i\theta}{a_{12}}(B_1\cos\theta t + B_1 i\sin\theta t) \\ & + \tfrac{(\alpha-a_{11})-i\theta}{a_{12}}(B_2\cos\theta t - B_2 i\sin\theta t)\Big] \\ = {} & e^{\alpha t}\Big[\tfrac{(\alpha-a_{11})(B_1+B_2)+i\theta(B_1-B_2)}{a_{12}}\cos\theta t + \tfrac{(\alpha-a_{11})(B_1-B_2)+i\theta(B_1+B_2)}{a_{12}} i\sin\theta t)\Big] \\ = {} & e^{\alpha t}\Big[\tfrac{(\alpha-a_{11})(B_1+B_2)+\theta(B_1-B_2)i}{a_{12}}\cos\theta t + \tfrac{(\alpha-a_{11})(B_1-B_2)i+i^2\theta(B_1+B_2)}{a_{12}}\sin\theta t)\Big] \\ = {} & e^{\alpha t}\Big[\tfrac{(\alpha-a_{11})A_1+\theta A_2}{a_{12}}\cos\theta t + \tfrac{(\alpha-a_{11})A_2-\theta A_1}{a_{12}}\sin\theta t)\Big], \end{aligned} \tag{18.42}$$

where in the last passage we have used $(B_1 + B_2) \equiv A_1$, $(B_1 - B_2)i \equiv A_2$, and the fact that $i^2 = -1$. Equations (18.41) and (18.42) coincide with (18.13) and (18.15). This completes the exposition of the 'direct' method of solution.

18.1.3 Particular solution. Determination of the arbitrary constants

We can now turn to the problem of finding a particular solution of the non-homogeneous system. The general method of undetermined coefficients can be applied here too, and we shall illustrate it in the case in which $g_1(t), g_2(t)$ are two given constants, say b_1, b_2. Thus we have

$$\begin{aligned} y' &= a_{11}y_t + a_{12}z + b_1, \\ z' &= a_{21}y_t + a_{22}z + b_2. \end{aligned} \tag{18.43}$$

As a particular solution let us try $\overline{y}_t = \mu_1, \overline{z}_t = \mu_2$, where μ_1, μ_2 are undetermined constants. Substituting in Eq. (18.43) and rearranging terms we have

$$\begin{aligned} a_{11}\mu_1 &+ a_{12}\mu_2 = -b_1, \\ a_{21}\mu_1 &+ a_{22}\mu_2 = -b_2, \end{aligned} \tag{18.44}$$

whence

$$\mu_1 = \frac{-b_1 a_{22} + b_2 a_{12}}{a_{11}a_{22} - a_{12}a_{21}}, \quad \mu_2 = \frac{-b_2 a_{11} + b_1 a_{21}}{a_{11}a_{22} - a_{12}a_{21}}. \tag{18.45}$$

The method is successful only if the determinant of system (18.44) is different from zero, i.e.

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \neq 0. \tag{18.46}$$

Note that this condition means that *zero* is *not* a root of the characteristic equation of the homogeneous differential system (18.2). When this occurs, the method fails; a way out is to try $\overline{y}(t) = \mu_{11} + \mu_{12}t$, $\overline{z}(t) = \mu_{21} + \mu_{22}t$, where μ_{ij} are undetermined constants.

Finally, the arbitrary constants that appear in the general solution can be determined as usual by means of a number of additional conditions equal to the number of arbitrary constants. These conditions give the information that, for a given value of t (usually for $t = 0$, whence the name *initial conditions*), the values of the various functions are know values (e.g., that y_0, z_0 are known). Substituting in the general solution of the system under consideration, we obtain a system of linear equations which can be solved for the values of the arbitrary constants.

18.2 First order n×n systems in normal form

In general, an $n \times n$ first-order system in normal form has the typical matrix form

$$\mathbf{y}' = \mathbf{A}\mathbf{y} + \mathbf{g}(t), \tag{18.47}$$

where $\mathbf{y} = [y_1(t), y_2(t), ..., y_n(t)]$ is the column vector of the unknown functions of time to be found,

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \qquad (18.48)$$

is the square matrix of the given coefficients, and $\mathbf{g}(t) = [g_1(t), g_2(t), ..., g_n(t)]$ is a column vector of known functions of time.

Let us first observe that an $n \times n$ system can always be reduced, by a procedure similar to that used in relation to the 2×2 system, to a single nth order equation in one unknown function. Incidentally, note that the converse is also true, i.e. a nth order equation can always be transformed into a first-order system in normal form having n equations in n unknown functions. To do this, new variables $y_1(t), y_2(t), ..., y_{n-1}(t)$ are defined such that

$$\begin{aligned} &y' = y_1, \\ &y_1' = y_2, \\ &\vdots \\ &y_{n-2}' = y_{n-1}. \end{aligned} \qquad (18.49)$$

Substituting in the given nth order equation in y, and taking into account that $y^{(n)} = y_{n-1}'$, a first-order equation in $y_{n-1}', y_{n-1}, ..., y_1, y$ is obtained, that—together with the equations defining the new variables—forms a first-order system in normal form. To wit, if we consider the nth order differential equation

$$a_0 y^{(n)} + a_1 y^{(n-1)} + a_2 y^{(n-2)} + ... + a_{n-1} y' + a_n y = g(t), \qquad (18.50)$$

we can write it, taking into account the definition of the new variables, as

$$\begin{aligned} y_{n-1}' &= -\frac{a_1}{a_0} y_{n-1} - \frac{a_2}{a_0} y_{n-2} - \dots - \frac{a_{n-1}}{a_0} y_1 - \frac{a_n}{a_0} y + \frac{g(t)}{a_0} \\ &= -\frac{a_n}{a_0} y - \frac{a_{n-1}}{a_0} y_1 - \dots - \frac{a_2}{a_0} y_{n-2} - \frac{a_1}{a_0} y_{n-1} + \frac{g(t)}{a_0}. \end{aligned} \qquad (18.51)$$

Hence the definitions (18.49) and the differential equation in the form (18.51) make up the following differential system

$$\mathbf{Y}' = \mathbf{AY} + \mathbf{G}(t), \qquad (18.52)$$

where

$$\mathbf{Y} \equiv \begin{bmatrix} y \\ y_1 \\ \vdots \\ y_{n-2} \\ y_{n-1} \end{bmatrix}, \mathbf{A} \equiv \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \\ \frac{-a_n}{a_0} & \frac{-a_{n-1}}{a_0} & \frac{-a_{n-2}}{a_0} & \dots & \frac{-a_1}{a_0} \end{bmatrix}, \mathbf{G}(t) \equiv \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \frac{g(t)}{a_0} \end{bmatrix}. \qquad (18.53)$$

System (18.52) is clearly a first-order system in the normal form (18.47).

Going back to systems, the reduction process is time-consuming when the number of equations increases, hence it is preferable to apply the direct method of solution. If we consider for example the homogeneous system

$$\begin{aligned} y' &= a_{11} y + a_{12} z + a_{13} w, \\ z' &= a_{21} y + a_{22} z + a_{23} w, \\ w' &= a_{31} y + a_{32} z + a_{33} w, \end{aligned}$$

we can immediately write its characteristic equation

$$D(\lambda) = \begin{vmatrix} a_{11} - \lambda & a_{12} & a_{13} \\ a_{21} & a_{22} - \lambda & a_{23} \\ a_{31} & a_{32} & a_{33} - \lambda \end{vmatrix} = 0. \qquad (18.54)$$

Expanding the determinant we have a third-degree algebraic equation in the unknown λ, whose solution will give three values, $\lambda_1, \lambda_2, \lambda_3$. Since the stable or unstable behaviour over time of the solution depends exclusively on the roots $\lambda_1, \lambda_2, \lambda_3$, to analyse the stability of the system we can examine only the nature of such roots, without any need to compute the coefficients $\alpha_i^{(j)}$. For this purpose we can apply to the characteristic equation the stability conditions stated in Chap. 16, Sect, 16.4, which allow us to check whether the roots of a polynomial have negative real parts without finding them explicitly.

With the help of a little matrix algebra, the direct method can easily be generalized to first-order systems of type (18.1) having any number of equations.

18.2.1 Solution of the homogeneous system

Let us begin with the homogeneous system

$$\mathbf{y}' = \mathbf{Ay}. \qquad (18.55)$$

An immediate extension of the direct method of solution leads us to consider the possible solution $\mathbf{y} = \boldsymbol{\alpha} e^{\lambda t}$, where $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, ..., \alpha_n]$ is a vector of constants not all zero. By substituting into (18.55) we get

$$\lambda \boldsymbol{\alpha} e^{\lambda t} = \mathbf{A} \boldsymbol{\alpha} e^{\lambda t}, \tag{18.56}$$

i.e.

$$e^{\lambda t}[\mathbf{A} - \lambda \mathbf{I}]\boldsymbol{\alpha} = \mathbf{0}, \tag{18.57}$$

which will be identically satisfied if, and only if,

$$[\mathbf{A} - \lambda \mathbf{I}]\boldsymbol{\alpha} = \mathbf{0}. \tag{18.58}$$

System (18.58) will have a non-trivial solution ($\boldsymbol{\alpha} \neq \mathbf{0}$) if, and only if, its determinant is zero, namely

$$|\mathbf{A} - \lambda \mathbf{I}| = 0, \tag{18.59}$$

which is the determinantal form of the characteristic equation of the matrix $\mathbf{A}$. Expansion of this determinant yields an nth order polynomial equation in λ of the type

$$(-1)^n \lambda^n + c_1(-1)^{n-1}\lambda^{n-1} + ... + c_r(-1)^r \lambda^r + ... + c_{n-1}(-1)\lambda + c_n = 0 \tag{18.60}$$

or, multiplying through by $(-1)^n$ and taking account that $(-1)^{2n} = 1$, $(-1)^{2n-1} = -1$,

$$\lambda^n - c_1\lambda^{n-1} + ... + c_r(-1)^{n+r}\lambda^r + ... + c_{n-1}(-1)^{n+1}\lambda + (-1)^n c_n = 0, \tag{18.61}$$

where

$$\begin{aligned} c_1 &= \textstyle\sum_{i=1}^n a_{ii}, \\ c_2 &= \text{sum of all second-order principal minors of the matrix } \mathbf{A}, \\ ... &\quad ... \\ c_r &= \text{sum of all } n!/r!(n-r)! \text{ principal minors of the } r\text{th order}, \\ ... &\quad ... \\ c_n &= |\mathbf{A}|. \end{aligned} \tag{18.62}$$

The solution of Eq. (18.61) will give n roots $\lambda_1, \lambda_2, ..., \lambda_n$, the characteristic roots (or latent roots or eigenvalues) of the matrix $\mathbf{A}$. To each eigenvalue λ_j a corresponding eigenvector (or characteristi vector or latent vector) $\boldsymbol{\alpha}^{(j)}$ can be associated through (18.58), as shown in Sect. 18.1.2 for the 2×2 case.

The solution of the differential equation system under consideration will then have the form

$$\begin{aligned} y_1(t) &= A_1\alpha_1^{(1)}e^{\lambda_1 t} + A_2\alpha_1^{(2)}e^{\lambda_2 t} + ... + A_n\alpha_1^{(n)}e^{\lambda_n t}, \\ y_2(t) &= A_1\alpha_2^{(1)}e^{\lambda_1 t} + A_2\alpha_2^{(2)}e^{\lambda_2 t} + ... + A_n\alpha_2^{(n)}e^{\lambda_n t}, \\ ... &\quad ... \\ y_n(t) &= A_1\alpha_n^{(1)}e^{\lambda_1 t} + A_2\alpha_n^{(2)}e^{\lambda_2 t} + ... + A_n\alpha_n^{(n)}e^{\lambda_n t}, \end{aligned} \tag{18.63}$$

if the eigenvalues are distinct. Note that this expression remains valid also in the case in which some roots are complex (we only have to use the transformation from the exponential to the polar form of complex numbers). In the case of multiple roots we proceed as shown above in relation to the 2×2 system.

To determine the vector of the arbitrary constants $\mathbf{a} = [A_1, A_2, ..., A_n]$ we need n additional conditions, for example $\mathbf{y} = \mathbf{y}_0$ for $t = 0$, where $\mathbf{y}_0$ is a vector of known values, called *initial conditions* (when the side conditions are given at different points of time—for example, some values at $t = t_0$ and other values at $t = t_1$—we speak of *boundary conditions*). From Eqs. (18.63) we then have, since $e^0 = 1$,

$$\mathbf{y}_0 = \mathbf{a}\mathbf{V}, \tag{18.64}$$

where $\mathbf{V}$, the matrix of system (18.63), is the *modal matrix* of $\mathbf{A}$, namely the matrix whose columns are the characteristic vectors of the matrix $\mathbf{A}$. From elementary matrix algebra we know that the characteristic vectors associated with distinct characteristic roots are linearly independent, hence $\mathbf{V}$ is non-singular. Therefore

$$\mathbf{a} = \mathbf{V}^{-1}\mathbf{y}_0 \tag{18.65}$$

yields the required solution for the arbitrary constants.

In the case of multiple roots we can no longer use the property of linear independence, but it is enough to observe that the general solution will consist of the linear combination of n distinct solutions forming a fundamental set, so that the appropriate matrices will always be non singular.

18.2.1.1 The matrix exponential

We have seen in Chap. 12 that the solution of the simple first-order scalar differential equation $y' = ay$ is of the type $y(t) = e^{at}C$, where C is an arbitrary constant. Analogously we would like to be able to say that the solution of the vector differential equation

$$\mathbf{y}' = \mathbf{A}\mathbf{y}, \tag{18.66}$$

i.e., of system (18.55), is

$$\mathbf{y}(t) = e^{\mathbf{A}t}\mathbf{c}, \tag{18.67}$$

where $\mathbf{c}$ is a $(n \times 1)$ vector of arbitrary constants. However, while the scalar exponential e^{at} is well defined, the matrix exponential $e^{\mathbf{A}t}$ needs definition. There is a very natural way of defining $e^{\mathbf{A}t}$ so that it resembles e^{at}. In fact, e^{at} has the well-known power series expansion

$$e^{at} = 1 + at + \frac{a^2t^2}{2} + \frac{a^3t^3}{3!} + ... + \frac{a^nt^n}{n!} + ... \tag{18.68}$$

and similarly we define the matrix exponential as the matrix power series expansion

$$e^{\mathbf{A}t} = \mathbf{I} + \mathbf{A}t + \frac{\mathbf{A}^2t^2}{2} + \frac{\mathbf{A}^3t^3}{3!} + ... + \frac{\mathbf{A}^nt^n}{n!} + ... \tag{18.69}$$

The infinite series (18.69) converges for all t (see, for example, Gantmacher, 1959, p. 137) and can be differentiated term by term to yield

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}e^{\mathbf{A}t} &= \mathbf{A} + \mathbf{A}^2t + \frac{\mathbf{A}^3t^2}{2} + ... + \frac{\mathbf{A}^nt^{n-1}}{(n-1)!} + ... \\ &= \mathbf{A}\left[\mathbf{I} + \mathbf{A}t + \frac{\mathbf{A}^2t^2}{2} + ... + \frac{\mathbf{A}^{n-1}t^{n-1}}{(n-1)!} + ...\right] \\ &= \mathbf{A}e^{\mathbf{A}t}. \end{aligned} \tag{18.70}$$

Therefore, Eq. (18.67) is indeed a solution to Eq. (18.66), since

$$\mathbf{y}' = \frac{\mathrm{d}}{\mathrm{d}t}e^{\mathbf{A}t}\mathbf{c} = \mathbf{A}e^{\mathbf{A}t}\mathbf{c} = \mathbf{A}\mathbf{y}. \tag{18.71}$$

Since $\mathbf{c}$ is a $(n \times 1)$ vector of arbitrary constants, the solution contains exactly n arbitrary constants, hence—according to general principles—it is the general solution.

Of course the formal solution (18.67) is of little practical value unless a closed-form expression for $e^{\mathbf{A}t}$, as distinct from the power series expansion (18.69), can be found. This is not difficult, especially when the matrix $\mathbf{A}$ has distinct eigenvalues. In this case we know from matrix algebra that $\mathbf{A}$ can be diagonalized by a similar transformation with the modal matrix $\mathbf{V}$, namely

$$\mathbf{\Lambda} = \mathbf{V}^{-1}\mathbf{A}\mathbf{V} \quad \text{or} \quad \mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}, \tag{18.72}$$

where

$$\mathbf{\Lambda} = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ ... & ... & ... & ... \\ 0 & 0 & 0 & \lambda_n \end{bmatrix} \tag{18.73}$$

is the diagonal matrix of the eigenvalues of $\mathbf{A}$. Hence

$$e^{\mathbf{A}t} = e^{\mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}t} = e^{\mathbf{V}\mathbf{\Lambda}t\mathbf{V}^{-1}}. \tag{18.74}$$

We now apply Sylvester's theorem (see, for example, Collar and Simpson, 1987, Sect. 1.18) to obtain

$$e^{\mathbf{V}\mathbf{\Lambda}t\mathbf{V}^{-1}} = \mathbf{V}e^{\mathbf{\Lambda}t}\mathbf{V}^{-1}, \tag{18.75}$$

where

$$e^{\mathbf{\Lambda}t} = \begin{bmatrix} e^{\lambda_1 t} & 0 & 0 & 0 \\ 0 & e^{\lambda_2 t} & 0 & 0 \\ ... & ... & ... & ... \\ 0 & 0 & 0 & e^{\lambda_n t} \end{bmatrix}. \tag{18.76}$$

Thus we have

$$\mathbf{y}(t) = e^{\mathbf{A}t}\mathbf{c} = \mathbf{V}e^{\mathbf{\Lambda}t}(\mathbf{V}^{-1}\mathbf{c}) \tag{18.77}$$

and, given the arbitrariness of $\mathbf{c}$, we can set $\mathbf{V}^{-1}\mathbf{c} = \mathbf{a}$, a new vector of arbitrary constants. We finally arrive at the expression

$$\mathbf{y}(t) = \mathbf{V}e^{\mathbf{\Lambda}t}\mathbf{a}, \tag{18.78}$$

that coincides with the closed-form eigenvalue-eigenvector solution previously found: see Eq. (18.63).

In the case of multiple roots the algebra is rather more difficult and lengthy, hence we refer the reader to the literature (see, e.g., Braun, 1986, Sects. 3.11, 3.12; Collar and Simpson, 1987, Sect. 6.1).

18.2.2 Stability conditions

Since the stable or unstable behaviour over time of the solution depends exclusively on the roots $\lambda_1, \lambda_2, ..., \lambda_n$, to analyse the stability of the system we can examine only the nature of such roots, without any need to compute the elements of the eigenvectors, $\alpha_i^{(j)}$. For this purpose we can apply to the characteristic equation (18.61) the stability conditions stated in Chap. 17, that allow us to check whether the roots of a polynomial have negative real parts without calculating them.

However, the expansion of the characteristic determinant to find the polynomial form of the characteristic equation is rather laborious if n is great and if, as is the case in theoretical work, we have to deal with symbolic rather than numeric values of the matrix coefficients. The coefficients of the polynomial are complicated functions of the elements a_{ij} –see Eq. (18.62)—and the Routh-Hurwitz determinants have elements that are themselves sums of determinants. For $n > 2$ the expressions become increasingly cumbersome.

Two related questions then arise:

(i) whether one can obtain alternative determinantal stability criteria in which the elements of the determinants are simpler functions of the a_{ij};

(ii) whether stability conditions exist that can be applied directly to the coefficients a_{ij} of the system without having to expand the determinantal equation (18.59).

As regards the first question, certain 'modified Routh-Hurwitz conditions' have been developed (Fuller, 1968; see also Murata, 1977). However, these modified conditions, though easier to deal with than the conditions obtained using the straightforward method of expanding (18.59) and applying the original Routh-Hurwitz conditions to the resulting polynomial, are by no means simple. For example, in the case of a 3×3 system, they turn out to be

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} < 0, \quad a_{11} + a_{22} + a_{33} < 0,$$

$$\begin{vmatrix} a_{11}+a_{22} & a_{23} & -a_{13} \\ a_{32} & a_{11}+a_{33} & a_{12} \\ -a_{31} & a_{21} & a_{22}+a_{33} \end{vmatrix} < 0, \tag{18.79}$$

and become increasingly cumbersome as the order of the system increases (for the general form of the modified Routh-Hurwitz conditions see Fuller, 1968; Murata, 1977, p. 92).

Therefore, it would be highly desirable to have a positive answer to the second question. Fortunately the stability conditions we are looking for exist, and the most important are listed below. Further results may be found in Barnett and Storey (1970, Chap. 7) and in Murata (1977, Chaps. 3 and 4). Since the proofs require the knowledge of some advanced matrix algebra, they are given in Sect. 18.2.2.4, that can be skipped without loss of continuity.

In what follows, by 'stability conditions' we mean, as usual, 'conditions for the roots of the characteristic equation (18.61), be they real and/or complex, to have negative real parts' or, which is the same thing, for these roots to be strictly lying in the left half of the complex plane. It is important to note that, when these conditions are applied to economic models, one should pay attention to whether the condition being applied is necessary and sufficient, or only necessary, or only sufficient.

I. *Negative definiteness.* If the matrix is symmetric ($a_{ij} = a_{ji}$), then a set of necessary and sufficient stability conditions is given by the following n inequalities

$$a_{11} < 0, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \quad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} < 0,$$

$$\ldots, \operatorname{sgn} \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} = \operatorname{sgn}\,(-1)^n, \tag{18.80}$$

namely the leading principal minors (also called the upper left-hand principal minors) of the matrix $\mathbf{A}$ should alternate in sign, beginning with minus.

II. *Quasi-negative definiteness.* Form the matrix $\mathbf{B} \equiv \frac{1}{2}(\mathbf{A}^{\mathrm{T}} + \mathbf{A})$, i.e.

$$\mathbf{B} \equiv \begin{bmatrix} a_{11} & \frac{1}{2}(a_{12}+a_{21}) & \ldots & \frac{1}{2}(a_{1n}+a_{n1}) \\ \frac{1}{2}(a_{21}+a_{12}) & a_{22} & \ldots & \frac{1}{2}(a_{2n}+a_{n2}) \\ \ldots & \ldots & \ldots & \ldots \\ \frac{1}{2}(a_{n1}+a_{1n}) & \frac{1}{2}(a_{n2}+a_{2n}) & \ldots & a_{nn} \end{bmatrix}.$$

Since $\mathbf{B}$ is symmetric, we can apply conditions I to it. If these conditions are satisfied, then the characteristic roots of $\mathbf{A}$ are stable.

III. *Metzlerian matrix.* Let $a_{ij} \geq 0, i \neq j$, (i.e., all the off-diagonal elements of $\mathbf{A}$ are non-negative). Then a set of necessary and sufficient stability conditions is the same as in number I above. Note that these conditions imply $a_{ii} < 0$. A matrix with $a_{ii} < 0, a_{ij} \geq 0$ is called a Metzlerian matrix.

IV. *Dominant negative diagonal.* A set of sufficient stability conditions is that all the elements on the main diagonal are negative and each is in absolute value greater than the sum of the absolute values of all the other elements belonging to the same line (row or column). Formally,

$$a_{ii} < 0; \quad |a_{ii}| > \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}| \qquad \text{(row dominance)}, \tag{18.81}$$

or

$$a_{ii} < 0; \quad |a_{ii}| > \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ji}| \qquad \text{(column dominance)}. \tag{18.82}$$

Note that conditions (18.92) and (18.82) are alternative, in the sense that each set is by itself sufficient for stability.

Conditions IV can be extended to the case of:

V. *Quasi-dominant negative diagonal.* A set of sufficient stability conditions is that all the elements on the main diagonal are negative and each is in absolute value not smaller than the sum of the absolute values of all the other elements belonging to the same line (row or column), after appropriately weighting the elements involved. Formally,

$$a_{ii} < 0; \quad h_i\,|a_{ii}| \geq \sum_{\substack{j=1 \\ j\neq i}}^{n} h_j\,|a_{ij}| \qquad \text{(at least one strict inequality)}, \tag{18.83}$$

or, alternatively,

$$a_{ii} < 0; \quad h_i\,|a_{ii}| \geq \sum_{\substack{j=1 \\ j\neq i}}^{n} h_j\,|a_{ji}| \qquad \text{(at least one strict inequality)}, \tag{18.84}$$

where the h's are all positive. The 'at least one strict inequality' clause must be further qualified in the case of zero off-diagonal elements. More precisely, when $a_{ij} = 0$ (given $j \in J$ and $i \notin J$ for some set of indices J), the strict inequality must hold for at least one $j \in J$.

It should also be noted that, when an arbitrary matrix has a dominant or quasi-dominant negative diagonal, then this matrix will satisfy conditions (18.80), but the converse is not true.

VI. *Trace.* A necessary (but not sufficient) stability condition is that the trace of $\mathbf{A}$ is negative, namely

$$\sum_{i=1}^{n} a_{ii} < 0. \tag{18.85}$$

VII. *Determinant.* A necessary (but not sufficient) stability condition is that the determinant of $\mathbf{A}$ has the sign of $(-1)^n$.

18.2.2.1 D-stability, and stabilisation of matrices

Consider the matrix $\mathbf{A}$ and a diagonal matrix $\mathbf{D} = \text{diag}\{k_1, k_2, ..., k_n\}$, where the k's are positive constants. Suppose now that $\mathbf{A}$ is a stable matrix, and consider the matrix $\mathbf{DA}$. When is it possible to say that $\mathbf{DA}$ is also stable? This is the problem known as the D-stability problem, and a stable matrix $\mathbf{A}$ such that $\mathbf{DA}$ is also stable is called a D-stable matrix. The interest of this problem lies in the fact that in economic applications one may have to consider a differential system of the type

$$\begin{aligned} y_1' &= k_1(a_{11}y_1 + a_{12}y_2 + ... + a_{1n}y_n), \\ y_2' &= k_2(a_{21}y_1 + a_{22}y_2 + ... + a_{2n}y_n), \\ ... & \; ... \; ... \\ y_n' &= k_n(a_{n1}y_1 + a_{n2}y_2 + ... + a_{nn}y_n), \end{aligned} \tag{18.86}$$

where the k's are positive constants (for example, adjustment speeds), and the question arises whether the system under consideration is stable for any set of these constants. In other words, assuming that the basic system (18.55) is stable, can we say that system (18.86) is stable for any choice of positive k's? Since system (18.86) can be written as

$$\mathbf{y}' = \mathbf{DAy}, \tag{18.87}$$

we are exactly in the context of D-stability. The following results are well-known in the mathematical economics literature (see, for example, Newman, 1959; further results are given in Johnson, 1974, and Khalil, 1980):

(a) if the matrix $\mathbf{A}$ satisfies any one of the stability conditions I through V listed in the previous section, then it is D-stable, namely system (18.87) is also stable;

(b) given an arbitrary matrix $\mathbf{A}$, conditions (18.80) are necessary, though not sufficient, for system (18.87) to be stable for any set of positive k's.

In the study of D-stability the k's are given, at least in principle. Suppose, on the contrary, that they can be chosen at will (for example, they are policy parameters). A very important question is then whether conditions exist under which, given system (18.55) arbitrarily, system (18.87) can be made stable by an appropriate choice of the k's, that may now have any sign. This is known in the mathematical literature as the *stabilization of matrices* problem (Fisher and Fuller, 1958). The following is a slightly strengthened form of the Fisher-Fuller theorem, due to McFadden (1968).

Fisher-Fuller theorem (strengthened form). Suppose $\mathbf{A}$ is a real $n \times n$ matrix with the property that the upper left-hand principal minor A_i of each order $i = 1, ..., n$ is non-zero. Then, there exists a positive scalar ϵ such that if the real diagonal matrix $\mathbf{D} = \text{diag}\{k_1, k_2, ..., k_n\}$ satisfies the conditions

$$\begin{aligned} k_iA_i/A_{i-1} &< 0 \text{ for } i = 1, 2, ..., n \quad (A_0 = 1), \\ &\text{and} \\ |k_i| \,/\, |k_{i-1}| &< \epsilon \text{ for } i = 2, ..., n, \end{aligned} \tag{18.88}$$

then the characteristic roots of $\mathbf{DA}$ are real, negative, and distinct.

Since in economic applications it is usually desirable to be able to give positive values to the k's, it is important to note that $k_i > 0$ *when* $A_i/A_{i-1} < 0$, *namely when conditions* (18.80) *are satisfied.* This result is an obvious consequence of the condition $k_iA_i/A_{i-1} < 0$.

18.2.2.2 Sensitivity analysis

The latent roots of $\mathbf{A}$ are obtained by solving the characteristic equation (18.61), whose coefficients are related to the elements of $\mathbf{A}$ through the relations (18.62). We know that the roots of a polynomial are differentiable functions of the coefficients of the polynomial and so, ultimately, of all the elements of $\mathbf{A}$. These, in turn, may be functions of other relevant parameters: in economic applications, for example, the elements of $\mathbf{A}$ will depend on the underlying model parameters. More precisely, we have

$$\begin{aligned} \lambda_m &= \varphi([a_{ij}]) = \varphi(a_{11}, a_{12}, ..., a_{nn}), \quad m = 1, 2, ..., n, \\ a_{ij} &= \psi(\alpha_1, \alpha_2, ..., \alpha_p), \end{aligned} \tag{18.89}$$

where λ_m is the m-th latent root of $\mathbf{A}$, $[a_{ij}]$ denotes the set of all the elements of $\mathbf{A}$, and α_s, $s = 1, 2, ..., p$, are the parameters of the underlying economic model.

By sensitivity analysis we here mean the computation of the partial derivatives of the m-th latent root with respect to the s-th parameter. This involves the preliminary computation of the partial derivatives of the latent roots with respect to the elements of the matrix, for applying the composite function rule we have

$$\frac{\partial \lambda_m}{\partial \alpha_s} = \frac{\partial \lambda_m}{\partial \mathbf{A}} \frac{\partial \mathbf{A}}{\partial \alpha_s} = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial \lambda_m}{\partial a_{ij}} \frac{\partial a_{ij}}{\partial \alpha_s}. \tag{18.90}$$

The partial derivatives $\partial a_{ij}/\partial \alpha_s$ will depend on the economic model under consideration and their computation need not concern us here. The problem is then to calculate the partial derivatives $\partial \lambda_m/\partial a_{ij}$,which is what is usually meant by *sensitivity analysis* in dynamical system analysis (see, for example, Laughton, 1964). Taking into consideration solely the case of distinct latent roots, that suffices for all practical purposes, it can be shown that the basic formula is

$$\frac{\partial \lambda_m}{\partial \mathbf{A}} = \left[\frac{\partial \lambda_m}{\partial a_{ij}}\right] = (\mathbf{v}_{\overline{m}}^{-1})^{\mathrm{T}} \mathbf{v}_m^{\mathrm{T}}, \tag{18.91}$$

or, equivalently,

$$\frac{\partial \lambda_m}{\partial a_{ij}} = (\mathbf{V}^{-1}\mathbf{I}_{ij}\mathbf{V})_{m,m}, \tag{18.92}$$

where $\mathbf{v}_m$ is the (column) latent vector associated with λ_m, $\mathbf{v}_m^{\mathrm{T}}$ is the transpose of $\mathbf{v}_m$ (i.e., the row vector obtained by transposing the mth column of $\mathbf{V}$), $(\mathbf{v}_{\overline{m}}^{-1})^{\mathrm{T}}$ is the column vector obtained by transposing the mth row of $\mathbf{V}^{-1}$, and $\mathbf{V}$ is the modal matrix, i.e. the matrix whose columns are the latent vectors of $\mathbf{A}$, and $\mathbf{I}_{ij}$ is the matrix that has unity in the $(i.j)$ position and zeros elsewhere.

Formula (18.91), which was proven by Laughton (1964) and Gandolfo (1981), gives the matrix whose elements are the partial derivatives $\partial \lambda_m/\partial a_{ij}$.

Formula (18.92), which was proven by Phillips (1982), gives the single partial derivative $\partial \lambda_m/\partial a_{ij}$, and is equivalent to the previous one because $(\mathbf{V}^{-1}\mathbf{I}_{ij}\mathbf{V})_{m,m}$ is just the (i,j)th element of the outer product of the latent row and column vectors associated with λ_m.

The proof of (18.91) is rather lengthy and complicated, while the proof of (18.92) is much shorter and simpler, hence we shall give the latter.

Since we are assuming distinct latent roots, $\mathbf{A}$ can be diagonalized by a similar transformation through the modal matrix $\mathbf{V}$, namely

$$\mathbf{\Lambda} = \mathbf{V}^{-1}\mathbf{A}\mathbf{V} \quad \text{or} \quad \mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}, \tag{18.93}$$

where $\mathbf{\Lambda}$ is the diagonal matrix of the latent roots of $\mathbf{A}$. Taking differentials we then have

$$\begin{aligned} \mathrm{d}\mathbf{A} &= (\mathrm{d}\mathbf{V})\mathbf{\Lambda}\mathbf{V}^{-1} + \mathbf{V}(\mathrm{d}\mathbf{\Lambda})\mathbf{V}^{-1} - \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-2}\mathrm{d}\mathbf{V} \\ &= (\mathrm{d}\mathbf{V})\mathbf{\Lambda}\mathbf{V}^{-1} + \mathbf{V}(\mathrm{d}\mathbf{\Lambda})\mathbf{V}^{-1} - \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}(\mathrm{d}\mathbf{V})\mathbf{V}^{-1}, \end{aligned} \tag{18.94}$$

from which, premultiplying by $\mathbf{V}^{-1}$ and postmultiplying by $\mathbf{V}$,

$$\mathbf{V}^{-1}(\mathrm{d}\mathbf{A})\mathbf{V} = \mathbf{V}^{-1}(\mathrm{d}\mathbf{V})\mathbf{\Lambda} + \mathrm{d}\mathbf{\Lambda} - \mathbf{\Lambda}\mathbf{V}^{-1}(\mathrm{d}\mathbf{V}). \tag{18.95}$$

We now observe that the diagonal elements of $\mathbf{V}^{-1}(\mathrm{d}\mathbf{V})\mathbf{\Lambda}$ and $\mathbf{\Lambda}\mathbf{V}^{-1}(\mathrm{d}\mathbf{V})$ are identical, so that they cancel out in Eq. (18.95). Hence the (m,m)th element of the diagonal matrix $\mathrm{d}\mathbf{\Lambda}$ (this element is $\mathrm{d}\lambda_m$) equals the (m,m)th element of $\mathbf{V}^{-1}(\mathrm{d}\mathbf{A})\mathbf{V}$, i.e.

$$\mathrm{d}\lambda_m = \left(\mathbf{V}^{-1}(\mathrm{d}\mathbf{A})\mathbf{V}\right)_{m,m}, \tag{18.96}$$

from which we immediately obtain the partial derivative

$$\frac{\partial \lambda_m}{\partial a_{ij}} = (\mathbf{V}^{-1}\mathbf{I}_{ij}\mathbf{V})_{m,m},$$

as stated above.

Sensitivity analysis is very important in the analysis of dynamical systems. It can point out parameters that crucially affect stability, either positively (when an increase in the parameter causes a significant decrease in the m-th latent root) or negatively (when an increase in the parameter causes a significant increase in the m-th latent root). It can help in the study of structural stability: loosely speaking, a system is called structurally stable if slight changes in its parameters do not alter the stability properties of the model (more on this in Chap. 21, Sect. 21.2.3). It is essential in the study of bifurcation points, i.e. of the values of a parameter at which a latent root changes its nature (see Chap. 25).

Sensitivity analysis normally requires numerical values, since the symbolic computation of formulae (18.91) and (18.92), even if now practicable thanks to the help of symbolic computation software, will usually not give meaningful results in the case of $n > 3$. Thus its main domain of application is in the field of numerical simulation and in the field of continuous-time *econometric* models, where its use has been introduced by Wymer (1968, 1987) and Gandolfo (1981, 1990, 1992), amongst others. In the case of $n \leq 3$ simpler formulae can be derived by using the relations between the roots and the coefficients of a polynomial equation. Since the case $n = 2$ is trivial (see exercise 6), we consider the case $n = 3$.

The polynomial form of the characteristic equation, that can easily be obtained by expanding the determinant of the characteristic matrix, is

$$\lambda^3 + a_1\lambda^2 + a_2\lambda + a_3 = 0, \tag{18.97}$$

where the coefficients a_i are related to the principal minors of the system's matrix as shown above, Sect. 18.2.1. The relations between the roots and the coefficients (see Chap. 16, Sect. 16.4) are

$$\begin{aligned} a_1 &= -(\lambda_1+\lambda_2+\lambda_3), \\ a_2 &= \lambda_1\lambda_2+\lambda_1\lambda_3+\lambda_2\lambda_3, \\ a_3 &= -\lambda_1\lambda_2\lambda_3. \end{aligned} \tag{18.98}$$

Suppose now that the elements of the system's matrix, and hence the coefficients a_i of the characteristic equation, are functions of a set of parameters. Then we have, by differentiating (18.98) with respect to any parameter,

$$\begin{aligned} -\frac{\partial\lambda_1}{\partial\alpha_s} & -\frac{\partial\lambda_2}{\partial\alpha_s} & -\frac{\partial\lambda_3}{\partial\alpha_s} &= \frac{\partial a_1}{\partial\alpha_s}, \\ (\lambda_2+\lambda_3)\frac{\partial\lambda_1}{\partial\alpha_s} & +(\lambda_1+\lambda_3))\frac{\partial\lambda_2}{\partial\alpha_s} & +(\lambda_1+\lambda_2)\frac{\partial\lambda_3}{\partial\alpha_s} &= \frac{\partial a_2}{\partial\alpha_s}, \\ -\lambda_2\lambda_3\frac{\partial\lambda_1}{\partial\alpha_s} & -\lambda_1\lambda_3\frac{\partial\lambda_2}{\partial\alpha_s} & -\lambda_1\lambda_2\frac{\partial\lambda_3}{\partial\alpha_s} &= \frac{\partial a_3}{\partial\alpha_s}, \end{aligned} \tag{18.99}$$

where the derivatives $\partial a_i/\partial\alpha_s$ are computed from the expressions that give a_i in terms of the elements of the matrix and hence of the parameter set.

Thus system (18.99) can be solved as a linear system in the unknowns $\partial\lambda_i/\partial\alpha_s$. This may not seem very useful, because in purely qualitative work we do not know the roots λ_i. However, in qualitative work we are interested solely in the signs of $\partial\lambda_i/\partial\alpha_s$. These signs can often be determined by only knowing the signs and nature of the roots, that in turn can be ascertained solely by qualitative methods (see Chap. 16, Sect. 16.4). A case that is of paramount importance in the study of the Hopf bifurcation (see Chap. 25, Sect. 25.2.2) is when λ_1 is real (and negative) and $\lambda_{2,3}=\theta\pm i\omega$. Then Eqs. (18.98) become

$$\begin{aligned} -(\lambda_1+2\theta) &= a_1, \\ 2\theta\lambda_1+\theta^2+\omega^2 &= a_2, \\ -\lambda_1(\theta^2+\omega^2) &= a_3, \end{aligned} \tag{18.100}$$

from which

$$\begin{aligned} -\frac{\partial\lambda_1}{\partial\alpha_s} & -2\frac{\partial\theta}{\partial\alpha_s} & &= \frac{\partial a_1}{\partial\alpha_s}, \\ \cancel{2\theta}\frac{\partial\lambda_1}{\partial\alpha_s} & +2(\lambda_1+\theta)\frac{\partial\theta}{\partial\alpha_s} & +2\omega\frac{\partial\omega}{\partial\alpha_s} &= \frac{\partial a_2}{\partial\alpha_s}, \\ -(\theta^2+\omega^2)\frac{\partial\lambda_1}{\partial\alpha_s} & -2\theta\lambda_1\frac{\partial\theta}{\partial\alpha_s} & -2\omega\lambda_1\frac{\partial\omega}{\partial\alpha_s} &= \frac{\partial a_3}{\partial\alpha_s}. \end{aligned} \tag{18.101}$$

18.2.2.3 A digression on not-wholly-unstable systems

By a 'not wholly unstable system' (or 'conditionally stable' system) we mean a system whose roots are partly unstable (i.e., have positive real parts) and partly stable (negative real parts). Since the presence of even only one root with positive real part causes the overall movement of the system to be divergent, such a system should be classified as unstable.

However, if the initial point happens to be such that only the terms associated with the stable roots are present in the general solution, then it is clear that the movement will be stable. Conditional stability is stability conditional on the initial position of the system.

This can be simply illustrated with reference to the 2×2 homogeneous system, whose solution, given in Eqs. (18.21)-(18.24) above, is reproduced here:

$$\begin{aligned} y &= A_1e^{\lambda_1 t}+A_2e^{\lambda_2 t}, \\ z &= A_1\alpha_2^{(1)}e^{\lambda_1 t}+A_2\alpha_2^{(2)}e^{\lambda_2 t}, \end{aligned} \tag{18.102}$$

where $\alpha_2^{(1)},\alpha_2^{(2)}$ are as given in Eq. (18.24).

Suppose now that $\lambda_1<0,\lambda_2>0$. It is clear that the solution will be stable if $A_2=0$. We want to know the relations that the initial values of y and z must satisfy for A_2 to be zero. For this purpose, consider the equations for the determination of A_1, A_2 :

$$\begin{aligned} y_0 &= A_1+A_2, \\ z_0 &= A_1\alpha_2^{(1)}+A_2\alpha_2^{(2)}, \end{aligned} \tag{18.103}$$

from which

$$A_1=\frac{\alpha_2^{(2)}y_0-z_0}{\alpha_2^{(2)}-\alpha_2^{(1)}}, \qquad A_2=\frac{z_0-\alpha_2^{(1)}y_0}{\alpha_2^{(2)}-\alpha_2^{(1)}}. \tag{18.104}$$

Then a necessary and sufficient condition for $A_2=0$ is

$$z_0-\alpha_2^{(1)}y_0=0, \text{ i.e. } z_0/y_0=\alpha_2^{(1)}. \tag{18.105}$$

Equation (18.105) gives the relation we are looking for. With this relation between z_0 and y_0 it is also easy to see that $A_1=y_0=z_0/\alpha_2^{(1)}$.

When $A_2=0$ the solution of the system will be given by

$$y(t)=A_1e^{\lambda_1 t}, \quad z(t)=A_1\alpha_2^{(1)}e^{\lambda_1 t}.$$

This immediately yields the relation between y, z in the (y,z) plane, namely

$$z=\alpha_2^{(1)}y. \tag{18.106}$$

In the case of a non-homogeneous system, if the particular solution can be interpreted as the equilibrium path of the system, we can consider the deviations of the system from this particular solution. This leaves us with the homogeneous part of the system, and we are back into the case that we have just treated. In the contrary case we shall consider the general solution of the non-homogeneous system and proceed as before. The only difference is that in the formulae derived above we shall replace y_0, z_0 with $(y_0 - \overline{y}_0)$ and $(z_0 - \overline{z}_0)$, where $\overline{y}_0, \overline{z}_0$ are the initial values of the particular solution of the system.

This procedure can easily be generalized to $n \times n$ systems. As shown below —see Eqs. (18.118)—the equations for the determination of the arbitrary constants turn out to be

$$\sum_{j=1}^{n} A_j \alpha_i^{(j)} = y_i(0) - \overline{y}_i(0), \qquad i = 1, 2, ..., n, \tag{18.107}$$

where $\overline{y}_i(0) \equiv 0$ if the system is homogeneous.

Let us first consider the case in which there is only one stable root (which must necessarily be real). Suppose that λ_j is such a root and consider the solution of Eqs. (18.107) by Cramer's rule

$$A_i = \frac{D_i}{D}, \text{ where } D \equiv \begin{vmatrix} \alpha_1^{(1)} & \alpha_1^{(2)} & \dots & \alpha_1^{(n)} \\ \alpha_2^{(1)} & \alpha_2^{(2)} & \dots & \alpha_2^{(n)} \\ \dots & \dots & \dots & \dots \\ \alpha_n^{(1)} & \alpha_n^{(2)} & \dots & \alpha_n^{(n)} \end{vmatrix}, \tag{18.108}$$

and D_i is obtained by substituting the column of the initial conditions in the place of the i-th column of D.

Then, if the initial conditions are proportional to $\alpha_1^{(j)}, \alpha_2^{(j)}, ..., \alpha_n^{(j)}$, namely to the eigenvector associated with λ_j, it follows that each D_i (except D_j) will be zero because it has two proportional columns, and so $A_i = 0, i \neq j$.

This can be summarised in the following theorem:

Theorem 18.1 The behaviour of the general solution of the system $\mathbf{y}' = \mathbf{A}\mathbf{y}$, where the matrix $\mathbf{A}$ has distinct characteristic roots, is determined solely by one particular characteristic root λ_j if, and only if, the initial condition vector $\mathbf{y}(0)$ is proportional to the characteristic vector associated with λ_j (namely $\mathbf{y}(0) = \beta\alpha^{(j)}$ for some non-zero constant β).

Theorem 18.1 implies, in the case there is just one stable root and the rest unstable, that there is a trajectory path defined by the characteristic vector associated with the stable root such that solutions lying on this path converge to the equilibrium point. This can be expressed in terms of the *stable manifold,* which is defined as the set of points for which $\mathbf{y(t)} \to \mathbf{y}_e$ as $t \to \infty$, and is called the *stable arm* when it has dimension one (which is the case with only one stable root). More generally, the *dimension* of the stable manifold is equal to the number of roots with negative real part. Thus we pass on to the case in which there are k stable and $(n - k)$ unstable roots $(0 < k < n)$.

From Eqs. (18.108) and the properties of determinants it is obvious that if the initial condition vector is a linear combination of the characteristic vectors associated with the stable roots, when we compute any A_i corresponding to an *unstable* root, the relative D_i will always have a column which is a linear combination of other columns, hence all these determinants will be zero.

This can be summarised in the following theorem:

Theorem 18.2 The behaviour of the general solution of the system $\mathbf{y}' = \mathbf{A}\mathbf{y}$, where the matrix $\mathbf{A}$ has distinct characteristic roots, is determined solely by a subset of k characteristic roots if, and only if, the initial condition vector is a linear combination of the characteristic vectors associated with these k roots.

In general there is no reason why the initial conditions should satisfy the requirements of either Theorem 18.1 or 18.2. Hence conditional stability might seem of little practical interest. This is not so. If the initial conditions can somehow be selected (see below), then it is clear that the requirements of Theorems 18.1 and 18.2 can always be satisfied. This means that one chooses the initial conditions in such a way that the arbitrary constant(s) associated with the unstable root(s) turn out to be zero, so that *the final solution will only contain the term(s) corresponding to the stable root(s).*

Consider, for example, the simple 2×2 case, and suppose we are free to choose y_0. Then for any given value of z_0 we simply have to set $y_0 = z_0/\alpha_2^{(1)}$ to make $A_2 = 0$. Similarly we shall set $z_0 = y_0\alpha_2^{(1)}$ if we are free too choose z_0 rather than y_0. In the general case it may happen that we are free to choose a certain number of initial conditions arbitrarily, while other initial conditions are given. How can we then ensure the condition required by Theorems 18.1 and 18.2? We now prove the following result:

Theorem 18.3 Given a first-order system in normal form with distinct characteristic roots, partly stable and partly unstable, *we can always make the system stable provided that we are free to choose as many initial conditions* {i.e., the values of $y_{k+1}(0), y_{k+2}(0), ..., y_n(0)$ or of $[y_{k+1}(0) - \overline{y}_{k+1}(0)], [y_{k+2}(0) - \overline{y}_{k+2}(0)], ..., [y_n(0) - \overline{y}_n(0)]$ as the case may be} *as there are unstable roots.*

Assume as before that there are k stable and $(n - k)$ unstable roots $(0 < k < n)$. Without loss of generality we can order them so that the first k are stable. We then proceed as follows:

(i) in (18.107) we impose $A_{k+1} = A_{k+2} = \ldots = A_n = 0$, and use the first k equations to determine $A_1, A_2, \ldots, A_k$ given $y_1(0), y_2(0), \ldots, y_k(0)$ (it goes without saying that $\overline{y}_1(0), \overline{y}_2(0), \ldots, \overline{y}_k(0)$ –if present– are also given);

(ii) these values of $A_1, A_2, \ldots, A_k$ are then substituted into the last $n-k$ equations of (18.107), thus obtaining the required values of $[y_{k+1}(0) - \overline{y}_{k+1}(0)], [y_{k+2}(0) - \overline{y}_{k+2}(0)], \ldots, [y_n(0) - \overline{y}_n(0)]$. If $\overline{y}_{k+1}(0), \overline{y}_{k+2}(0), \ldots, \overline{y}_n(0)$ are given, this will give us the required values of $y_{k+1}(0), y_{k+2}(0), \ldots, y_n(0)$. This proves Theorem 18.3 and also gives a practical way of proceeding.

All this reasoning might seem a curiosum. By definition, the initial conditions are given, and a given is a given. *However, the freedom of choice of an appropriate number of initial conditions is often allowed by the nature of the economic problem under consideration.* As we shall see in Chap. 22, it turns out that optimal growth models and rational-expectation models normally belong to this category. Hence, in economics, a not-wholly-unstable system may actually turn out to be a stable system.

We have so far reasoned on the assumption that we want to eliminate the unstable roots from the solution. But we might, on the contrary, be interested in the opposite case, namely in keeping only one *unstable* term in the final solution, and precisely the term associated with a real positive root. This is obviously the case if we are considering a growth model and want to find the conditions under which the model is capable of *balanced growth*, i.e. a state of growth in which the proportions that the variables bear to each other are constant.

This, of course, is equivalent to the requirement that all variables *actually* (and not only asymptotically) grow at the same proportional rate. Thus balanced growth obtains if, and only if, in the solution of the homogeneous system only one term (containing a positive root) remains. Application of Theorem 18.1 will immediately yield the desired result.

18.2.2.4 Proof of the stability conditions

Conditions I follow from the theory of quadratic forms.

It is a classic result that the quadratic form $Q(x) = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$ is negative definite if, and only if, the leading principal minors of the matrix $\mathbf{A}$ alternate in sign, beginning with minus (this can be proved by an extension of the elementary method of 'completing the square').

Another classic result is that the roots of a real symmetric matrix are all real, and this also can be proved quite easily: given (a) $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$, take complex conjugates, i.e. (b) $\mathbf{A}\overline{\mathbf{x}} = \overline{\lambda}\overline{\mathbf{x}}$. Premultiply (b) by $\mathbf{x}^{\mathrm{T}}$ and take complex conjugates, obtaining (c) $\overline{\mathbf{x}}^{\mathrm{T}}\mathbf{A}\mathbf{x} = \overline{\mathbf{x}}^{\mathrm{T}}\overline{\lambda}\mathbf{x}$. Now premultiply (a) by $\overline{\mathbf{x}}^{\mathrm{T}}$ and subtract the result from (c), obtaining $0 = (\lambda - \overline{\lambda})\overline{\mathbf{x}}^{\mathrm{T}}\mathbf{x}$, which gives $\lambda = \overline{\lambda}$, a contradiction if λ is complex.

Now, diagonalizing the matrix $\mathbf{A}$, the quadratic form can be transformed into a sum of squares of the type $\sum_i \lambda_i p_i^2$, where the λ's are the latent roots of $\mathbf{A}$, so that the quadratic form will be negative definite if, and only if, all the λ's are negative.

Since both sets of conditions are necessary and sufficient, they must be equivalent, and this completes the proof.

Conditions II can be easily proved by an application of Liapunov's second method (see Chap. 23). Given the linear differential system $\mathbf{x}' = \mathbf{A}\mathbf{x}$, consider the Liapunov function $\frac{1}{2}\mathbf{x}^{\mathrm{T}}\mathbf{x}$. Then $\mathrm{d}\left(\frac{1}{2}\mathbf{x}^{\mathrm{T}}\mathbf{x}\right)/dt < 0$ implies stability. It is easy to calculate

$$\begin{aligned} \frac{\mathrm{d}\left(\frac{1}{2}\mathbf{x}^{\mathrm{T}}\mathbf{x}\right)}{dt} &= \frac{1}{2}\frac{\mathrm{d}\mathbf{x}^{\mathrm{T}}}{dt}\mathbf{x} + \frac{1}{2}\mathbf{x}^{\mathrm{T}}\frac{\mathrm{d}\mathbf{x}}{dt} = \frac{1}{2}\mathbf{x}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{x} + \frac{1}{2}\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} \\ &= \mathbf{x}^{\mathrm{T}}\frac{1}{2}(\mathbf{A}^{\mathrm{T}} + \mathbf{A})\mathbf{x} < 0. \end{aligned}$$

Since $\frac{1}{2}(\mathbf{A}^{\mathrm{T}} + \mathbf{A})$ is a symmetric matrix, we can apply conditions I to it, and this gives conditions II.

Conditions III can be proved by an application of a theorem on non-negative matrices. Given a matrix $\mathbf{B} \geq \mathbf{0}$, let λ_M be its dominant root: the theorem states that a set of necessary and sufficient conditions for the real number λ to be greater than λ_M is that all the leading principal minors of the matrix $\lambda\mathbf{I} - \mathbf{B}$ are positive (Gantmacher, 1959, pp. 88-89).

Now, consider a matrix $\mathbf{A}$ with non-negative non-diagonal elements. Obviously for a certain $\lambda > 0$ the matrix $\mathbf{B} = \mathbf{A} + \lambda\mathbf{I}$ is non-negative. Let λ_n be the latent root of $\mathbf{A}$ having the greatest real part and let λ_M be the dominant root of $\mathbf{B}$. Since the roots of $\mathbf{B}$ are the sums $\lambda_i + \lambda$ (λ_i are the roots of $\mathbf{A}$), it follows that $\lambda_M = \lambda_n + \lambda$. Now $\lambda_M < \lambda$ if, and only if, $\lambda_n < 0$, i.e., when all the latent roots of $\mathbf{A}$ have negative real parts. Applying the theorem stated above to the matrix $-\mathbf{A} = \lambda\mathbf{I} - \mathbf{B}$, we obtain conditions III.

To prove conditions V (of which conditions IV are a particular case for $h_i = 1$), the classic result can be used that a quasi-dominant diagonal implies that the matrix is non-singular (this can be proven by showing that the contrary implies a contradiction: see McKenzie, 1960, p. 49). Now consider $\mathbf{A} - \lambda\mathbf{I}$. Since $a_{ii} < 0$, if λ has a non-negative real part , then $|a_{ii} - \lambda| \geq |a_{ii}|$, $\forall i$. These inequalities imply that $\mathbf{A} - \lambda\mathbf{I}$ has a quasi-dominant diagonal and is non-singular: therefore, λ cannot be a latent root of $\mathbf{A}$.

To prove conditions VI and VII we first recall from Eqs. (18.61) and (18.62) that in the polynomial form of the characteristic equation of $\mathbf{A}$ the coefficient of λ^{n-1} equals $-\sum_i a_{ii}$ and the constant term equals $(-1)^n(\det\mathbf{A})$. Now, from elementary algebra, the coefficient of λ^{n-1} is equal to $-\sum_i \lambda_i$, and the constant term is equal to $(-1)^n \prod_i \lambda_i$, hence $\sum_i \lambda_i = \sum_i a_{ii}$, and $\prod_i \lambda_i = \det \mathbf{A}$.

Since $\text{Re}(\lambda_i) < 0$, $\forall i$, implies $\sum_i \lambda_i < 0$, it follows that $\sum_i a_{ii} < 0$ is a necessary (though not sufficient) stability condition. Similarly, $\text{Re}(\lambda_i) < 0$, $\forall i$, implies $\text{sgn}(\prod_i \lambda_i) = \text{sgn}(-1)^n$, hence $\text{sgn}(\det \mathbf{A}) = \text{sgn}(-1)^n$ is another necessary (though not sufficient) stability condition.

18.2.3 Particular solution

The case of a non-homogeneous system is easily exemplified letting $\mathbf{g}(t) = \mathbf{b}$, where $\mathbf{b}$ is a vector of constants. We then try

$$\overline{\mathbf{y}}(t) = \overline{\mathbf{y}} \tag{18.109}$$

as a particular solution, where $\overline{\mathbf{y}}$ is vector of undetermined constants. Substitution in Eq. (18.47) yields

$$\mathbf{0} = \mathbf{A}\overline{\mathbf{y}} + \mathbf{b},$$

whence

$$\overline{\mathbf{y}} = -\mathbf{A}^{-1}\mathbf{b}. \tag{18.110}$$

This of course requires that zero is not a characteristic root, for in the contrary case we would have $|\mathbf{A}| = 0$. In such a case we try $\overline{\mathbf{y}}(t) = \boldsymbol{\mu}_1 + \boldsymbol{\mu}_2 t$, where $\boldsymbol{\mu}_1, \boldsymbol{\mu}_2$ are vectors of undetermined constants, etcetera.

18.2.3.1 Variation of parameters

The principles of this method have already been illustrated in Chap. 12, Sect. 12.2.6. Let us recall that this is a general method of solving a differential equation (in our case a system) by considering the arbitrary constants that appear in the known solution of a simpler system, as variable (i.e., as functions of t), and so determining them that the more general system is identically satisfied.

In our case the simpler system with known solution is the homogeneous system. Here the matrix exponential solution is particularly useful. Hence we start from Eq. (18.55) and its matrix exponential solution (18.67), and posit

$$\mathbf{y}(t) = e^{\mathbf{A}t}\mathbf{c}(t), \tag{18.111}$$

where $\mathbf{c}(t)$ is a vector of undetermined functions. Hence

$$\mathbf{y}'(t) = \mathbf{A}e^{\mathbf{A}t}\mathbf{c}(t) + e^{\mathbf{A}t}\mathbf{c}'(t). \tag{18.112}$$

Substituting (18.111) and (18.112) into Eq. (18.47) we have

$$\mathbf{A}e^{\mathbf{A}t}\mathbf{c}(t) + e^{\mathbf{A}t}\mathbf{c}'(t) = \mathbf{A}e^{\mathbf{A}t}\mathbf{c}(t) + \mathbf{g}(t),$$

from which

$$\mathbf{c}'(t) = e^{-\mathbf{A}t}\mathbf{g}(t). \tag{18.113}$$

Hence

$$\mathbf{c}(t) = \int e^{-\mathbf{A}t}\mathbf{g}(t)\mathrm{d}t + \mathbf{b}, \tag{18.114}$$

where $\mathbf{b}$ is a vector of arbitrary constants of integration. Going back to Eq. (18.111) we have

$$\mathbf{y}(t) = e^{\mathbf{A}t}\left(\int e^{-\mathbf{A}t}\mathbf{g}(t)\mathrm{d}t + \mathbf{b}\right) = e^{\mathbf{A}t}\mathbf{b} + e^{\mathbf{A}t}\int e^{-\mathbf{A}t}\mathbf{g}(t)\mathrm{d}t. \tag{18.115}$$

Let us note that Eq. (18.115) also contains the general solution of the homogeneous equation and the vector $\mathbf{b}$ of arbitrary constants, hence it is the general solution of the non-homogeneous equation. If we are only interested in the particular solution, then we set $\mathbf{b} = \mathbf{0}$ in (18.114) and consider

$$\overline{\mathbf{y}}(t) = e^{\mathbf{A}t}\int e^{-\mathbf{A}t}\mathbf{g}(t)\mathrm{d}t \tag{18.116}$$

as the particular solution we are looking for.

18.2.4 Determination of the arbitrary constants

Finally, the arbitrary constants that appear in the general solution can be determined as usual by means of a number of additional conditions equal to the number of arbitrary constants. These conditions give the information that, for a given value of t (usually for $t = 0$, whence the name *initial conditions*), the values of the various functions are known values. Substituting in the general solution of the system under consideration, we obtain a system of linear equations which can be solved for the values of the arbitrary constants.

Consider, for example, the case in which the characteristic roots are all distinct. Then the equations for the determination of the arbitrary constants turn out to be

$$\sum_{j=1}^{n} A_j \alpha_i^{(j)} + \overline{y}_i(0) = y_i(0), \qquad i = 1, 2, ..., n, \tag{18.117}$$

where $\overline{y}_i(0) \equiv 0$ if the system is homogeneous (see above Eq. (18.89)). It is a well-known theorem in matrix algebra that characteristic vectors associated with distinct characteristic roots are linearly independent. Therefore, the matrix $[\alpha_i^{(j)}]$ is non-singular and system (18.117) can be solved. Some complications may arise when there are multiple roots, but it is enough to observe that, since we start from a fundamental set, the relevant matrices will always be non singular according to general principles.

Note, finally, that system (18.117) can also be written as

$$\sum_{j=1}^{n} A_j\alpha_i^{(j)} = y_i(0) - \overline{y}_i(0), \qquad i = 1, 2, ..., n, \tag{18.118}$$

where the r.h.s. can be interpreted as giving the *initial deviations* of the system from its equilibrium solution $\overline{y}_i(0)$.

In the case of the matrix exponential solution (18.115), given $\mathbf{y}(t) = \mathbf{y}_0$ for $t = t_0$, we first observe that when initial conditions are given the solution takes the form

$$\mathbf{y}(t) = e^{\mathbf{A}t}\mathbf{b} + e^{\mathbf{A}t}\int_{t_0}^{t} e^{-\mathbf{A}s}\mathbf{g}(s)dt = e^{\mathbf{A}t}\mathbf{b} + \int_{t_0}^{t} e^{\mathbf{A}(t-s)}\mathbf{g}(s)ds, \tag{18.119}$$

since the integral has to be taken between t_0 and t, and $e^{\mathbf{A}t}$ can be brought under the integral sign since it is a constant with respect to the variable of integration s. Then by letting $t = t_0$ we have

$$\mathbf{y}_0 = e^{\mathbf{A}t_0}\mathbf{b}, \tag{18.120}$$

whence

$$\mathbf{b} = e^{-\mathbf{A}t_0}\mathbf{y}_0 \tag{18.121}$$

and substituting back into the solution (18.119) we obtain

$$\mathbf{y}(t) = e^{\mathbf{A}t}e^{-\mathbf{A}t_0}\mathbf{y}_0 + \int_{t_0}^{t} e^{\mathbf{A}(t-s)}\mathbf{g}(s)ds = e^{\mathbf{A}(t-t_0)}\mathbf{y}_0 + \int_{t_0}^{t} e^{\mathbf{A}(t-s)}\mathbf{g}(s)ds. \tag{18.122}$$

This formula further simplifies to

$$\mathbf{y}(t) = e^{\mathbf{A}t}\mathbf{y}_0 + \int_{0}^{t} e^{\mathbf{A}(t-s)}\mathbf{g}(s)ds \tag{18.123}$$

when, as is usual in economic applications, we take $t_0 = 0$.

18.3 General systems

The systems hitherto examined have the peculiarity that in each equation only first-order derivatives appear, and each equation involves the derivative of only one unknown function in turn. This is why they are called first-order and normal, as stated at the beginning of this chapter. But, in general, in each equation the first-order derivatives of two or more unknown functions as well as higher-order derivatives might appear.

For didactic purposes we distinguish between first-order systems not in normal form and higher-order systems.

18.3.1 First-order systems not in normal form

In this case, the derivatives are still first-order, but in each equation the derivatives of two (or more) functions appear:

$$\mathbf{A}\mathbf{y}' + \mathbf{B}\mathbf{y} = \mathbf{g}(t), \tag{18.124}$$

where $\mathbf{A}, \mathbf{B}$ are $n \times n$ square matrices, $\mathbf{y} = [y_1, y_2, ..., y_n]$ is a $n \times 1$ vector of unknown functions of time, and $\mathbf{g}(t)$ is a vector of known functions. The corresponding homogeneous system is

$$\mathbf{A}\mathbf{y}' + \mathbf{B}\mathbf{y} = \mathbf{0}. \tag{18.125}$$

To solve system (18.125) we can use the same direct method explained above in relation to first-order systems in normal form, namely we start by trying as a solution $y_i = \alpha_i e^{\lambda t}$:

$$\mathbf{y} = [\alpha_1 e^{\lambda t}, \alpha_2 e^{\lambda t}, ..., \alpha_n e^{\lambda t}] = \boldsymbol{\alpha} e^{\lambda t}, \tag{18.126}$$

where the elements of the vector $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, ..., \alpha_n]$ are not all zero. Substituting in system (18.125) we have

$$\mathbf{A}\boldsymbol{\alpha}\lambda e^{\lambda t} + \mathbf{B}\boldsymbol{\alpha} e^{\lambda t} = e^{\lambda t}(\mathbf{A}\lambda + \mathbf{B})\boldsymbol{\alpha} = \mathbf{0}, \tag{18.127}$$

which will be identically satisfied if, and only if

$$(\mathbf{A}\lambda + \mathbf{B})\boldsymbol{\alpha} = \mathbf{0}. \tag{18.128}$$

System (18.128), in turn, admits of a non-trivial solution for $\boldsymbol{\alpha}$ if, and only if, its determinant is zero, namely

$$|\mathbf{A}\lambda + \mathbf{B}| = 0. \tag{18.129}$$

The determinantal equation (18.129) and its expanded form are called the characteristic equation of system (18.125). The expansion of the determinant gives rise to a n-th order polynomial equation in the unknown λ. From this point on, the procedure is the same as that explained in detail in relation to first-order systems in normal form. If, for example, the characteristic equation has n distinct real roots, to each root λ_j we can associate a vector $\alpha^{(j)} = [\alpha_1^{(j)}, \alpha_2^{(j)}, ..., \alpha_n^{(j)}]$ derived from the solution of set (18.128) when $\lambda = \lambda_j$. Then the solution of system (18.125) will have the usual form

$$\begin{aligned}
y_1(t) &= A_1\alpha_1^{(1)}e^{\lambda_1 t} + A_2\alpha_1^{(2)}e^{\lambda_2 t} + ... + A_n\alpha_1^{(n)}e^{\lambda_n t},\\
y_2(t) &= A_1\alpha_2^{(1)}e^{\lambda_1 t} + A_2\alpha_2^{(2)}e^{\lambda_2 t} + ... + A_n\alpha_2^{(n)}e^{\lambda_n t},\\
\cdots &\quad \cdots\cdots\cdots\cdots\cdots\cdots\\
y_{nt}(t) &= A_1\alpha_n^{(1)}e^{\lambda_1 t} + A_2\alpha_n^{(2)}e^{\lambda_2 t} + ... + A_n\alpha_n^{(n)}e^{\lambda_n t},
\end{aligned} \tag{18.130}$$

where $A_1, A_2, ..., A_n$ are arbitrary constants to be determined by means of n additional conditions in the usual way.

Alternatively, we can *reduce system* (18.125) *to an equivalent first-order system in normal form* by simple manipulations. If the matrix $\mathbf{A}$ is non-singular, we immediately obtain the normal form

$$\mathbf{y}' = \mathbf{C}\mathbf{y}, \qquad \mathbf{C} = -\mathbf{A}^{-1}\mathbf{B}. \tag{18.131}$$

If $\mathbf{A}$ is singular, we must first reduce the number of equations and unknowns to r (the rank of $\mathbf{A}$) by suitable substitutions, and then express the remaining $y_i'(t), i = 1, 2, ..., r$; $(r < n)$, in terms of the $y_i(t)$ as before, obtaining the normal form. For the details of this procedure see Collar and Simpson (1986, Sect. 6.2). Note that when $\mathbf{A}$ is singular the order of the system is no longer n, but r.

The reason why the transformation of a general first-order system into the normal form (18.131) might be preferable to the direct solution is of course that by so doing we can apply the stability conditions given above (see Sect. 18.2.2).

Finally, a particular solution of the non-homogeneous system (18.124) can be found by the method of undetermined coefficients in the usual way, and the standard procedure can be applied to the determination of the arbitrary constants.

18.3.2 Higher-order systems

As we said above, in each equation *each* unknown function might appear with derivatives of any order. Let us begin with a simple example before going on to the general case.

18.3.2.1 An example

Let us consider the non-homogeneous system

$$\begin{aligned} a_3 y''' + a_2 y'' \quad &+ b_1 z' + b_0 z &= g_1(t), \\ c_1 y' + c_0 y \quad &+ d_2 z''' &= g_2(t), \end{aligned} \tag{18.132}$$

and its homogeneous counterpart

$$\begin{aligned} a_3 y''' + a_2 y'' \quad &+ b_1 z' + b_0 z &= 0, \\ c_1 y' + c_0 y \quad &+ d_2 z''' &= 0. \end{aligned} \tag{18.133}$$

To solve system (18.133) we can proceed as with first-order systems in normal form, namely either reduce it to a single differential equation or apply

the direct method, trying as a solution $y = \alpha_1 e^{\lambda t}, z_t = \alpha_2 e^{\lambda t}$, where α_1, α_2 are not both zero. We then have

$$\begin{aligned} a_3\alpha_1\lambda^3 e^{\lambda t} + a_2\alpha_1\lambda^2 e^{\lambda t} \quad &+ b_1\alpha_2\lambda e^{\lambda t} + b_0\alpha_2 e^{\lambda t} &= 0, \\ c_1\alpha_1\lambda e^{\lambda t} + c_0\alpha_1 e^{\lambda t} \quad &+ d_2\alpha_2\lambda^3 e^{\lambda t} &= 0, \end{aligned}$$

from which

$$\begin{aligned} e^{\lambda t}[(a_3\lambda^3 + a_2\lambda^2)\alpha_1 \quad &+ (b_1\lambda + b_0)\alpha_2] &= 0, \\ e^{\lambda t}[(c_1\lambda + c_0)\alpha_1 \quad &+ d_2\lambda^3\alpha_2] &= 0. \end{aligned} \tag{18.134}$$

If $\alpha_1 e^{\lambda t}, \alpha_2 e^{\lambda t}$ is a solution, system (18.134) must be satisfied for any t and this is possible—apart from the trivial case $\lambda = 0$—if, and only if

$$\begin{aligned} (a_3\lambda^3 + a_2\lambda^2)\alpha_1 \quad &+ (b_1\lambda + b_0)\alpha_2 &= 0, \\ (c_1\lambda + c_0)\alpha_1 \quad &+ d_2\lambda^3\alpha_2 &= 0. \end{aligned} \tag{18.135}$$

System (18.135) will yield a non-trivial solution for α_1, α_2 if, and only if, its determinant is zero, namely

$$\begin{vmatrix} a_3\lambda^3 + a_2\lambda^2 & b_1\lambda + b_0 \\ c_1\lambda + c_0 & d_2\lambda^3 \end{vmatrix} = 0. \tag{18.136}$$

The polynomial form of the determinantal equation (18.136) is

$$a_3 d_2\lambda^6 + a_2 d_2\lambda^5 - b_1 c_1\lambda^2 - (b_1 c_0 + b_0 c_1)\lambda - b_0 c_0 = 0. \tag{18.137}$$

From this point on, the procedure is the same as that explained above for first-order systems in normal form. If, for example, the characteristic equation (18.137) yields six real and distinct roots, to each of them we can associate a couple of values $\alpha_1^{(j)}, \alpha_2^{(j)}$, and the solution will have the form

$$\begin{aligned} y(t) = \; & A_1\alpha_1^{(1)}e^{\lambda_1 t} + A_2\alpha_1^{(2)}e^{\lambda_2 t} + A_3\alpha_1^{(3)}e^{\lambda_3 t} + A_4\alpha_1^{(4)}e^{\lambda_4 t} \\ & + A_5\alpha_1^{(5)}e^{\lambda_5 t} + A_6\alpha_1^{(6)}e^{\lambda_6 t}, \\ z(t) = \; & A_1\alpha_2^{(1)}e^{\lambda_1 t} + A_2\alpha_2^{(2)}e^{\lambda_2 t} + A_3\alpha_2^{(3)}e^{\lambda_3 t} + A_4\alpha_2^{(4)}e^{\lambda_4 t} \\ & + A_5\alpha_2^{(5)}e^{\lambda_5 t} + A_6\alpha_2^{(6)}e^{\lambda_6 t}, \end{aligned} \tag{18.138}$$

where $A_1, A_2, ..., A_6$ are arbitrary constants.

The stability of the system can be examined by applying the stability conditions explained in Sect. 16.4 to Eq. (18.137); see, however, below, for a more general treatment of stability.

In the case of a non-homogeneous system, a particular solution can be found by the method of undetermined coefficients. The arbitrary constants, finally, can be determined given a sufficient number of additional conditions, for example that $y_0, y_0'; z_0, z_0', z_0''$ are all known values.

18.3.2.2 The general case

The determinantal form of the characteristic equation of a general higher order system can be easily found if we introduce the differential operator $D \equiv \mathrm{d}/dt$, namely the operator such that

$$y' = Dy, \qquad D^n y = y^{(n)}, \tag{18.139}$$

and the polynomial operator

$$P(D) = a_0 D^n + a_1 D^{n-1} + ... + a_{n-1} D + a_n. \tag{18.140}$$

Note that D^k simply means that the operator D has been applied k times in succession.

Then a general homogeneous differential equation system can be written as

$$\begin{array}{lllll} P_{11}(D)y_1 & +P_{12}(D)y_2 & +...+ & P_{1n}(D)y_n & = 0, \\ P_{21}(D)y_1 & +P_{22}(D)y_2 & +...+ & P_{2n}(D)y_n & = 0, \\ ... & ... & ... & ... & ... \\ P_{n1}(D)y_1 & +P_{n2}(D)y_2 & +...+ & P_{nn}(D)y_n & = 0, \end{array} \tag{18.141}$$

where the $P_{ij}(D)$ are polynomial operators of the appropriate orders. For example, the homogeneous part of system (18.132) can be written as

$$\begin{array}{lll} (a_3 D^3 + a_2 D^2)y & + \ (b_1 D + b_0)z & = 0, \\ (c_1 D + c_0)y & + \ d_2 D^3 z & = 0, \end{array} \tag{18.142}$$

so that $P_{11}(D) = a_3 D^3 + a_2 D^2$, etcetera.

Now, if we apply the usual direct method of solution to system (18.141), it can easily be seen that the determinantal form of the characteristic equation of this system can be written simply by considering the determinant whose elements are the polynomials obtained by replacing D with λ in the polynomial operators $P_{ij}(D)$. That is to say,

$$\begin{vmatrix} P_{11}(\lambda) & P_{12}(\lambda) & ... & P_{1n}(\lambda) \\ P_{21}(\lambda) & P_{22}(\lambda) & ... & P_{2n}(\lambda) \\ ... & ... & ... & ... \\ P_{n1}(\lambda) & P_{n2}(\lambda) & ... & P_{nn}(\lambda) \end{vmatrix} = 0 \tag{18.143}$$

is the characteristic equation of system (18.141).

18.3.2.3 Transformation of a higher-order system into a first-order system in normal form

A general higher-order system can always be written as

$$[\mathbf{M}_0 D^m + \mathbf{M}_1 D^{m-1} + \mathbf{M}_2 D^{m-2} + ... + \mathbf{M}_m]\mathbf{y}(t) = \mathbf{g}(t), \tag{18.144}$$

where $\mathbf{y}$ is the n-dimensional vector $[y_1(t), y_2(t), ..., y_n(t)]$ of unknown functions of time, and $\mathbf{M}_i$, $i = 0, 1, 2, ..., m$ are $n \times n$ matrices with zeros in the appropriate places. This system can always be reduced into first-order normal form by means of standard transformations (see, e.g., Frazer, Duncan, Collar, 1938, Sect. 5.5; Collar and Simpson, 1987, Sect. 6.2). These amount to defining the new variables

$$\mathbf{y}_1 \equiv D\mathbf{y}, \quad \mathbf{y}_2 \equiv D\mathbf{y}_1, \quad \mathbf{y}_3 \equiv D\mathbf{y}_2, ..., \mathbf{y}_{m-1} \equiv D\mathbf{y}_{m-2}, \tag{18.145}$$

from which $D^j\mathbf{y} = \mathbf{y}_j, j = 1, 2, ..., m-1$, and $D^m\mathbf{y} = D\mathbf{y}_{m-1}$. Thus we can rewrite system (18.144) as

$$\mathbf{M}_0 D\mathbf{y}_{m-1} + \mathbf{M}_1\mathbf{y}_{m-1} + \mathbf{M}_2\mathbf{y}_{m-2} + ... + \mathbf{M}_{m-1}\mathbf{y}_1 + \mathbf{M}_m\mathbf{y} = \mathbf{g}(t). \tag{18.146}$$

Considering for example system (18.133), we have

$$\begin{gathered} \begin{bmatrix} a_3 & 0 \\ 0 & d_2 \end{bmatrix} \begin{bmatrix} y''' \\ z''' \end{bmatrix} + \begin{bmatrix} a_2 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} y'' \\ z'' \end{bmatrix} + \begin{bmatrix} 0 & b_1 \\ c_1 & 0 \end{bmatrix} \begin{bmatrix} y' \\ z' \end{bmatrix} \\ + \begin{bmatrix} 0 & b_0 \\ c_0 & 0 \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} g_1(t) \\ g_2(t) \end{bmatrix}, \end{gathered} \tag{18.147}$$

and if we now define $y' = y_1, y_1' = y_2; z' = z_1, z_1' = z_2$, we obtain

$$\begin{gathered} \begin{bmatrix} a_3 & 0 \\ 0 & d_2 \end{bmatrix} \begin{bmatrix} y_2' \\ z_2' \end{bmatrix} + \begin{bmatrix} a_2 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} y_2 \\ z_2 \end{bmatrix} + \begin{bmatrix} 0 & b_1 \\ c_1 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ z_1 \end{bmatrix} \\ + \begin{bmatrix} 0 & b_0 \\ c_0 & 0 \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} g_1(t) \\ g_2(t) \end{bmatrix}, \end{gathered} \tag{18.148}$$

which is in the form (18.146).

We can assume that $\mathbf{M}_0$ is non singular –in the opposite case, we first reduce the system to the rank of $\mathbf{M}_0$, as explained in Sect. 18.3.1– so that we can premultiply Eq. (18.144) by $\mathbf{M}_0^{-1}$. Defining $\mathbf{N}_i \equiv \mathbf{M}_0^{-1}\mathbf{M}_i, i = 1, 2, ..., m$; $\mathbf{G}(t) \equiv \mathbf{M}_0^{-1}\mathbf{g}(t)$, and using the definitions (18.145), we have

$$
\begin{array}{lllllll}
y'_{m-1} = & -N_1 y_{m-1} & -N_2 y_{m-2} & -\ldots & -N_{m-1} y_1 & -N_m y & +G(t), \\
y'_{m-2} = & y_{m-1}, & & & & & , \\
y'_{m-3} = & & y_{m-2}, & & & & , \\
\vdots & & & & & & \vdots \\
y' = & & & & y_1. & & .
\end{array}
\tag{18.149}
$$

System (18.149) can be written in the composite matrix form

$$
\mathbf{Y}' = \mathbf{N}\mathbf{Y} + \mathbf{b}(t), \tag{18.150}
$$

where

$$
\mathbf{Y} \equiv \begin{bmatrix} \mathbf{y}_{m-1} \\ \mathbf{y}_{m-2} \\ \vdots \\ \mathbf{y}_1 \\ \mathbf{y} \end{bmatrix}, \quad \mathbf{b} \equiv \begin{bmatrix} \mathbf{G}(t) \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \tag{18.151}
$$

and

$$
\mathbf{N} \equiv \begin{bmatrix}
-\mathbf{N}_1 & -\mathbf{N}_2 & -\mathbf{N}_3 & \ldots & -\mathbf{N}_{m-1} & -\mathbf{N}_m \\
\mathbf{I} & 0 & 0 & \ldots & 0 & 0 \\
0 & \mathbf{I} & 0 & \ldots & 0 & 0 \\
0 & 0 & \mathbf{I} & \ldots & 0 & 0 \\
\vdots & \vdots & \vdots & \ldots & \vdots & \vdots \\
0 & 0 & 0 & \ldots & \mathbf{I} & 0
\end{bmatrix} \tag{18.152}
$$

is a square $(nm) \times (nm)$ matrix. System (18.150) is a first-order system in normal form.

For example, take system (18.147) and first put it into the form (18.149). It is easy to check that

$$
\mathbf{M}_0^{-1} = \begin{bmatrix} 1/a_3 & 0 \\ 0 & 1/d_2 \end{bmatrix},
$$

so that

$$
\begin{aligned}
-\mathbf{N}_1 &= \begin{bmatrix} -a_2/a_3 & 0 \\ 0 & 0 \end{bmatrix}, -\mathbf{N}_2 = \begin{bmatrix} 0 & -b_1/a_3 \\ -c_1/d_2 & 0 \end{bmatrix}, \\
-\mathbf{N}_3 &= \begin{bmatrix} 0 & -b_0/a_3 \\ -c_0/d_2 & 0 \end{bmatrix}; \quad \mathbf{G}(t) = \begin{bmatrix} g_1(t)/a_3 \\ g_2(t)/d_2 \end{bmatrix}.
\end{aligned} \tag{18.153}
$$

Hence the matrix $\mathbf{N}$ turns out to be

$$
\mathbf{N} = \left[\begin{array}{cc|cc|cc}
-a_2/a_3 & 0 & 0 & -b_1/a_3 & 0 & -b_0/a_3 \\
0 & 0 & -c_1/d_2 & 0 & -c_0/d_2 & 0 \\
\hline
1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 \\
\hline
0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0
\end{array}\right]. \tag{18.154}
$$

The reader may wish to check as an exercise that the characteristic equation of the matrix (18.154) gives rise to the polynomial equation

$$
\lambda^6 + \frac{a_2}{a_3}\lambda^5 - \frac{b_1 c_1}{a_3 d_2}\lambda^2 - \frac{b_1 c_0 + b_0 c_1}{a_3 d_2}\lambda - \frac{b_0 c_0}{a_3 d_2} = 0,
$$

which coincides with Eq. (18.137).

18.3.2.4 Stability conditions for higher-order systems

The transformation of a higher-order system into an equivalent first-order system in normal form might seem a roundabout and complicated way of solving it. This is certainly true if one is interested in finding the explicit form of the solution. But it is no longer true if one is interested in checking the stability of the system without solving it. In such a case the transformation explained in the previous section gives us the possibility of applying to the matrix $\mathbf{N}$—defined in (18.152)—the stability conditions valid for first-order systems in normal form (see above, Sect. 18.2.2).

18.4 Exercises

18.4.1 Example

Consider the first-order system whose matrix is

$$
\mathbf{A} = \begin{bmatrix}
-\epsilon & -1 & 0 & 0 \\
0 & -\epsilon & -1 & 0 \\
0 & 0 & -\epsilon & -1 \\
1 & -1 & 1 & -1-\epsilon
\end{bmatrix},
$$

where ϵ is an arbitrary positive number. It can be checked that the principal minors of $\mathbf{A}$ alternate in sign, beginning with minus, for any positive ϵ. However, the system's matrix does not belong to any of the types for which this alternation is a stability condition (see Sect. 18.2.2). Hence we shall have to examine the characteristic equation of **A** to proceed further. This equation is

$$\begin{vmatrix} -\epsilon-\lambda & -1 & 0 & 0 \\ 0 & -\epsilon-\lambda & -1 & 0 \\ 0 & 0 & -\epsilon-\lambda & -1 \\ 1 & -1 & 1 & -1-\epsilon-\lambda \end{vmatrix} = 0,$$

and its polynomial form turns out to be, after rearrangement of terms,

$$(\epsilon+\lambda)^4 + (\epsilon+\lambda)^3 + (\epsilon+\lambda)^2 + (\epsilon+\lambda) + 1 = 0.$$

The transformation of variables $x = \epsilon+\lambda$ gives rise to the reciprocal equation

$$x^4 + x^3 + x^2 + x + 1 = 0, \qquad (18.155)$$

that can be solved by standard procedures (see, for example Turnbull, 1957, pp. 114-115). Since $x = 0$ is not a root, we can divide through by $x^2 \neq 0$; rearranging terms we get

$$x^2 + x^{-2} + x + 1 = 0.$$

The transformation

$$z = x + x^{-1}, \quad z^2 = x^2 + x^{-2} + 2$$

puts the equation into the form

$$z^2 + z - 1 = 0,$$

which yields

$$z_1, z_2 = -\frac{1}{2} \pm \frac{1}{2}\sqrt{5}.$$

Since the transformation $z = x + x^{-1}$can also be written as

$$x^2 - zx + 1 = 0,$$

we have

$$x_1, x_2 = \frac{1}{2}z \pm \frac{1}{2}\sqrt{z^2-4}.$$

If we finally set z equal to its two values found before, we have the four values of x that solve the original equation. It turns out that these roots are all complex, and precisely

$$x_{1,2} = \cos 72^\circ \pm \mathrm{i}\sin 72^\circ \quad x_{3,4} = -\cos 36^\circ \pm \mathrm{i}\sin 36^\circ. \qquad (18.156)$$

A computationally simpler, though conceptually harder, way of obtaining the same result is to observe that the polynomial (18.155) can be considered as a divisor of $x^5 - 1$, since

$$x^5 - 1 = (x-1)(x^4 + x^3 + x^2 + x + 1),$$

so that the roots of Eq. (18.155) will be given by the roots of $x^5 - 1 = 0$ after discarding the spurious root $x = 1$. These roots are

$$x = \sqrt[5]{1}.$$

From complex number theory (see, for example, Courant, 1937, p. 73) we know that the n roots of unity are given by the formula

$$\cos\frac{2k\pi}{n} + \mathrm{i}\frac{2k\pi}{n},$$

where k is given the values $0, 1, ..., n-1$. In our case $n = 5$, and after discarding the spurious root (that comes out of the formula for $n = 0$) we have

$$\begin{aligned} x_1 &= \cos 72^\circ + \mathrm{i}\sin 72^\circ, x_2 = \cos 144^\circ + \mathrm{i}\sin 144^\circ, \\ x_3 &= \cos 216^\circ + \mathrm{i}\sin 216^\circ, x_4 = \cos 288^\circ + \mathrm{i}\sin 288^\circ. \end{aligned}$$

Now from elementary trigonometry we know that $\cos 72^\circ = \cos 288^\circ$, $\sin 72 = -\sin 288^\circ$, $\cos 144^\circ = \cos 216^\circ = -\cos 36^\circ$, $\sin 144^\circ = \sin 36^\circ = -\sin 216^\circ$, hence we get the same roots given in (18.155).

Since $x = \lambda + \epsilon$, we have $\lambda = x - \epsilon$ and so the roots of the characteristic equation are

$$\lambda_{1,2} = (\cos 72^\circ - \epsilon) \pm \sin 72^\circ \quad \lambda_{3,4} = (-\cos 36^\circ - \epsilon) \pm \sin 36^\circ. \qquad (18.157)$$

It is now easy to see that, given the arbitrariness of ϵ, we can set $0 < \epsilon < \cos 72^\circ$, so that the first pair of roots has a positive real part. Hence the solution will display an unstable oscillatory movement.

18.4.2 Other exercises

1. Solve the homogeneous first order system whose matrix is

$$\begin{bmatrix} -2 & 4 \\ -1 & 1 \end{bmatrix},$$

and show that the approach to equilibrium is damped oscillatory.

2. Solve the homogeneous first order system whose matrix is

$$\begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix}.$$

3. Solve the following systems and determine the arbitrary constants:

(3.i)

$$\mathbf{y}' = \begin{bmatrix} 1 & -1 \\ 5 & -3 \end{bmatrix}\mathbf{y}, \quad \mathbf{y}(0) = \begin{bmatrix} 1 \\ 2 \end{bmatrix},$$

(3.ii)

$$\mathbf{y}' = \begin{bmatrix} 2 & -6 \\ -1 & 1 \end{bmatrix}\mathbf{y}, \quad \mathbf{y}(0) = \begin{bmatrix} 0 \\ 5 \end{bmatrix},$$

(3.iii)

$$\mathbf{y}' = \begin{bmatrix} 3 & -4 \\ 1 & -1 \end{bmatrix}\mathbf{y}, \quad \mathbf{y}(0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

(3.iv)

$$\mathbf{y}' = \begin{bmatrix} 4 & 5 \\ -2 & -2 \end{bmatrix}\mathbf{y} + \begin{bmatrix} 4e^t \cos t \\ 0 \end{bmatrix}, \quad \mathbf{y}(0) = \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

(3.v)

$$\mathbf{y}' = \begin{bmatrix} -1 & -1 & 2 \\ 1 & 1 & 1 \\ 2 & 1 & 3 \end{bmatrix}\mathbf{y} + \begin{bmatrix} e^t \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{y}(0) = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

4. Transform the following first-order system into normal form and solve it:

$$\begin{aligned} 2y_1' + y_2' - 4y_1 &= e^t, \\ y_1' + 3y_1 + y_2 &= 0. \end{aligned}$$

5. Solve the following higher-order non-homogeneous system:

$$\begin{aligned} (D^2 - 2)y_1 - 3y_2 &= e^{2t}, \\ y_1 + (D^2 + 2)y_2 &= 0, \end{aligned}$$

with the initial conditions $y_1 = y_2 = 1$ and $y_1' = y_2' = 0$ for $t = 0$. (Hint: apply the roots of unity formula given in the Example).

6. Consider the 2×2 first-order system in normal form $\mathbf{y}' = \mathbf{A}\mathbf{y}$, where the elements of $\mathbf{A}$ are continuously differentiable functions of a parameter α:

$$\mathbf{A} \equiv \begin{bmatrix} a_{11}(\alpha) & a_{12}(\alpha) \\ a_{21}(\alpha) & a_{22}(\alpha) \end{bmatrix}.$$

Write down the polynomial form of the characteristic equation of $\mathbf{A}$ and show how to calculate the derivatives of the characteristic roots with respect to the parameter (i.e., $\partial\lambda_i/\partial\alpha, i = 1, 2$) both by starting from the standard formula for the roots of a quadratic equation and by starting from the relations between the roots and the coefficients.

18.5 References

Arrow, K.J., 1974, Stability Independent of Adjustment Speeds.

Arrow, K.J. and M. McManus, 1958, A Note on Dynamic Stability.

Barnett, S. and C. Storey, 1970, *Matrix Methods in Stability Theory,* Chap.7.

Bellman, R., 1953, *Stability Theory of Differential Equations,* Chap. 1, §§ 1, 7, 12.

Braun, M., 1986, *Differential Equations and their Applications,* Chap. 3, Sects. 3.8-3.12.

Collar, A.R. and A. Simpson, 1987, *Matrices and Engineering Dynamics,* Chap. 6.

Courant, R., 1961 (1937), *Differential and Integral Calculus,* Vol. I, p. 75.

Fisher, M.E., and A.T. Fuller, 1958, On the Stabilization of Matrices and the Convergence of Linear Iterative Processes.

Frazer, R.A., W.J. Duncan and A.R. Collar, 1938, *Elementary Matrices and Some Applications to Dynamics and Differential Equations,* Chaps. V, VI.

Fuller, A.T., 1968, Conditions for a Matrix to Have Only Characteristic Roots with Negative Real Parts.

Gandolfo, G., 1981, *Qualitative Analysis and Econometric Estimation of Continuous Time Dynamic Models,* Chap. 2, Sect. 2.2.4.

Gandolfo, G. and P.C. Padoan, 1990, The Italian Continuous Time Model: Theory and Empirical Results, pp. 115-116.

Gandolfo, G., 1992, Sensitivity Analysis in Continuous Time Econometric Models.

Gantmacher, F.R., 1959, *Applications of the Theory of Matrices,* pp. 88-89, 137.

Ince, E.L., 1956, *Ordinary Differential Equations,* Chap. V, Sect. 5.23; Chap. VI, Sect. 6.4.

Johnson, C.R., 1974, Sufficient Conditions for D-stability.

Khalil, H.K., 1980, A New Test for D-stability.

Laughton, M.A., 1964, Sensitivity in Dynamical System Analysis.

McFadden, D., 1968, On Hicksian Stability.

McKenzie, L., 1960, Matrices with Dominant Diagonals and Economic Theory, pp. 47-49.

Murata, Y., 1977, *Mathematics for Stability and Optimization of Economic Systems,* Chap. 2, Sect. 2.5; Chap. 3, Sects. 3.1, 3.4; Chap. 4, Sects.

4.2, 4.5.
Newman, P.K., 1959, Some Notes on Stability Conditions.
Phillips, P.C.B., 1982, A Simple Proof of the Latent Root Sensitivity Formula.
Quirk, J.P., 1974, A Class of Generalized Metzlerian Matrices.
Samuelson, P.A., 1947, *Foundations of Economic Analysis,* Appendix A, Sect. V; Appendix B, §§ 30-35.
Turnbull, H.W., 1957, *Theory of Equations,* pp. 114-115.
Wymer, C.R., 1968, The Computation of the Eigensystem of an Econometric Model.
Wymer, C.R., 1987, Sensitivity Analysis of Economic Policy.

Chapter 19

Differential Equation Systems in Economic Models

A few moments of reflection will suffice to convince us that practically all economic models, even the simplest ones, are simultaneous systems. The fact that many of them are usually examined as giving rise to a single functional equation is due to the fact that the reduction to a single equation involves only direct substitutions and is actually easier than simultaneous methods. For example, consider the homogeneous part of the Walrasian demand and supply adjustment model (see Chap. 12, Sect. 12.1):

$$\begin{aligned} D &= a + bp, \\ S &= a_1 + b_1 p, \\ p' &= c(D - S), \end{aligned}$$

which has the characteristic equation

$$\begin{vmatrix} 1 & 0 & -b \\ 0 & 1 & -b_1 \\ c & -c & -\lambda \end{vmatrix} = -\lambda - b_1 c + bc = 0,$$

so that $\lambda = c(b - b_1)$, etcetera. However, direct substitution from the first two equations into the third suggests itself quite naturally and is actually simpler.

On the other hand, there are economic models in which, after all direct substitutions have been made, two or more simultaneous equations still remain, to which simultaneous methods can be profitably applied.

19.1 Stability of Walrasian general equilibrium of exchange

The problem of the stability of demand and supply equilibrium was examined in Chap. 12, Sect. 12.1, in the context of a partial equilibrium analysis. In the case of a general equilibrium analysis the demand and supply functions relative to each good are in principle functions of the prices of all goods, and not only of the price of the good to which they refer. Thus we can write

$$E_j = E_j(p_1, p_2, ..., p_m), \quad j = 1, 2, ..., m, \tag{19.1}$$

where E_j is the excess demand for the j-th good. The equilibrium point $(p_1^e, p_2^e, ..., p_m^e)$ is determined by the conditions that all excess demands are simultaneously zero. We assume that this point exists and is economically meaningful, and turn to the study of its stability.

As in Sect. 13.1 we distinguish between static and dynamic stability. Although the relevant concept is the dynamic one, we shall examine static stability as well, in order to show that as soon as we pass from one to two or more markets the static stability conditions do *not* coincide any more (except for some particular cases) with the 'true' dynamic stability conditions.

The behaviour assumption that we make is that the price of a good varies in relation to the good's excess demand, i.e. the price increases (decreases) if excess demand is positive (negative). This corresponds to the mechanism conceived by Walras, the famous 'tâtonnement', of which a brief account is the following (Walras, 1954, Lesson 12, §§124-130; Jaffé, 1967). Assume that the market starts with arbitrary initial prices (*'prix criés au hasard'*, in Walras' terminology). An 'auctioneer' collects supply offers and demand requests for each good, and if a good is in excess demand (excess supply) he raises (lowers) the price of the good, and communicates the new set of prices to the market. The process goes on until, according to Walras, equilibrium is reached. It must be noted that in this idealized process no exchanges take place until equilibrium is reached (i.e. contracts are provisional, and binding only if prices turn out to be the equilibrium prices). The reason for this assumption is that, if exchanges take place out of equilibrium, then the initial endowments of the various goods owned by the single traders would change; these changes would modify the form of the excess demand functions, and so the position of the equilibrium point itself would change. Later studies have considered the so called 'non-tâtonnement' processes, which allow trading out of equilibrium according to certain rules, but we shall not examine these processes (for an introduction see Negishi, 1962).

Now, Walras did not give a rigorous proof that the tâtonnement process was indeed convergent (see however below, Sect. 19.1.2), nor did his successors. The first to tackle the stability problem rigorously was J. R. Hicks (1939), who extended the static stability condition of partial equilibrium to general equilibrium.

19.1.1 Static stability

The static stability condition is that a change in price causes a change in the opposite direction of the 'own' excess demand, i.e.

$$\mathrm{d}E_j/\mathrm{d}p_j < 0. \tag{19.2}$$

Of course, account must be taken that we are in a general equilibrium context, so that the change in a price influences not only its own excess demand, but also, in principle, all the other excess demands. Thus, as Hicks observes, conditions (19.2) must be qualified. Hicks' qualifications consist in the distinction between *imperfect* and *perfect* stability.

Stability is imperfect when (19.2) holds only in the case in which, given a change in the j-th price, all the other prices have adjusted in such a way that all the other markets are again in equilibrium.

Stability is perfect when (19.2) holds in any case, that is when (1) all the other prices have adjusted as in the previous case; (2) all the other prices have remained constant; (3) any subset of k other prices have varied so that equilibrium in the respective markets has been restored, while the remaining $(m-k)$ prices have remained constant (if we let k vary from zero to $m-1$, cases (1) and (2) are also included here).

To find the conditions for two kinds of stability we begin with the total differentials of the excess demand functions, $\mathrm{d}E_i$, evaluated at the equilibrium point. Note that the total differential can approximate the variation in a function only in a sufficiently small neighbourhood of the starting point, hence the conditions that we shall obtain are *local* stability conditions. Thus we have

$$\mathrm{d}E_i = a_{i1}\mathrm{d}p_1 + a_{i2}\mathrm{d}p_2 + ... + a_{im}\mathrm{d}p_m, \quad i = 1, 2, ..., m, \tag{19.3}$$

where $a_{ik} \equiv \partial E_i/\partial p_k$ evaluated at the equilibrium point $(k = 1, 2, ..., m)$.

Let us now assume that the price of the j-th good has changed and that all the other prices have varied as required by the definition of imperfect stability. Thus we have

$$\mathrm{d}E_j = a_{j1}\mathrm{d}p_1 + a_{j2}\mathrm{d}p_2 + ... + a_{jm}\mathrm{d}p_m, \tag{19.4}$$

while the remaining total differentials are all zero since equilibrium has been restored:

$$\mathrm{d}E_k = 0 = a_{k1}\mathrm{d}p_1 + a_{k2}\mathrm{d}p_2 + ... + a_{km}\mathrm{d}p_m, \quad k = 1, 2, ..., m; k \neq j. \tag{19.5}$$

Equations (19.4) and (19.5) form a set of m equations in the m unknowns $\mathrm{d}p_i$. Solving with respect to $\mathrm{d}p_j$ we have

$$\mathrm{d}p_j = \mathrm{d}E_j \frac{D_{jj}}{D}, \tag{19.6}$$

where

$$D \equiv \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mm} \end{vmatrix}, \tag{19.7}$$

the determinant of the system, is the Jacobian of the excess demand functions, and D_{jj} is the cofactor of a_{jj} in D. Note that, since the sums of the subscripts of a_{jj} is even, the cofactor is the same as the corresponding minor; moreover, this minor, by its very definition, is a principal minor of order $m-1$.

From Eq. (19.7) we immediately have

$$\frac{dE_j}{dp_j} = \frac{D}{D_{jj}}, \tag{19.8}$$

so that inequality (19.2) will be satisfied if, and only if, D and D_{jj} are of opposite sign. This condition must hold for any j from 1 to m, since any market may be involved. It follows that necessary and sufficient conditions for *imperfect stability* are that all the principal minors of order $m-1$ of D have a sign opposite to that of D.

To obtain the perfect stability conditions let us consider first the case in which the price of good j varies while all the other prices remain constant. Thus we have

$$\begin{aligned} dE_j &= a_{jj}dp_j, \\ dE_k &= a_{kj}dp_j, \quad k = 1, 2, ..., m;\ k \neq j. \end{aligned} \tag{19.9}$$

From the first equation in (19.9) we have

$$dE_j/dp_j = a_{jj}, \tag{19.10}$$

and so

$$a_{jj} < 0 \tag{19.11}$$

for all j is the first condition for perfect stability. That is, the partial derivative of the j-th excess demand with respect to the 'own' price p_j, must be negative. Thus Eq. (19.10) implies that all markets must be stable when considered in isolation: $a_{jj} < 0$ is, in fact, equivalent to the static stability condition in the case of a partial equilibrium analysis.

Consider now the case in which p_j varies and another price, say p_h, adjusts in such a way that equilibrium is restored in the h-th market, while all the remaining prices are constant. The relevant equations are

$$\begin{aligned} dE_j &= a_{jj}dp_j + a_{jh}dp_h, \\ dE_h &= 0 = a_{hj}dp_j + a_{hh}dp_h, \end{aligned} \tag{19.12}$$

so that, solving for dp_j, we have

$$dp_j = dE_j \frac{a_{hh}}{\begin{vmatrix} a_{jj} & a_{jh} \\ a_{hj} & a_{hh} \end{vmatrix}}, \tag{19.13}$$

from which

$$\frac{dE_j}{dp_j} = \frac{\begin{vmatrix} a_{jj} & a_{jh} \\ a_{hj} & a_{hh} \end{vmatrix}}{a_{hh}}. \tag{19.14}$$

It follows that $dE_j/dp_j < 0$ if, and only if,

$$\begin{vmatrix} a_{jj} & a_{jh} \\ a_{hj} & a_{hh} \end{vmatrix} > 0, \tag{19.15}$$

since a_{hh} must be negative from (19.11). Condition (19.15) must hold for any combination of subscripts h and j ($h \neq j$), since any two markets may be involved. From its definition, the determinant that appears in (19.15) is a second-order principal minor of D. Thus condition (19.15) states that all the second-order principal minors of D must be positive.

By similarly considering adjustments in two, three,..., markets we obtain the complete necessary and sufficient stability conditions for *perfect stability*, which are that all the principal minors of order r of D have the sign of $(-1)^r$, $r = 1, 2, ..., m$. This is the same as saying that such principal minors taken in ascending order alternate in sign beginning with minus:

$$a_{jj} < 0, \begin{vmatrix} a_{jj} & a_{jh} \\ a_{hj} & a_{hh} \end{vmatrix} > 0, \begin{vmatrix} a_{jj} & a_{jh} & a_{js} \\ a_{hj} & a_{hh} & a_{hs} \\ a_{sj} & a_{sh} & a_{ss} \end{vmatrix} < 0, ..., \tag{19.16}$$

for any $j, h, s, ...$(not equal).

This completes the examination of the Hicksian *static stability* conditions, so that we can now study the *dynamic stability* of the model.

19.1.2 Dynamic stability

The dynamic formalization of the Walrasian behaviour assumption is the following simultaneous system of differential equations

$$p_j' = F_j[E_j(p_1, p_2, ..., p_m)], \quad j = 1, 2, ..., m, \tag{19.17}$$

where the F_j's are sign-preserving functions, namely

$$\text{sgn}F_j[...] = \text{sgn}[...], \quad F_j[0] = 0, \quad dF_j[0]/dE_j > 0.$$

Equations (19.17) are an obvious extension to a multi-market context of the dynamic formalization for a single market (see Chap. 13, Sect. 13.1). Performing the linearisation of the F's and of the E's at the equilibrium point and considering the deviations, we have

$$(p_j - p_j^e)' = k_j a_{j1}(p_1 - p_1^e) + k_j a_{j2}(p_2 - p_2^e) + ... + k_j a_{jm}(p_m - p_m^e),\ j = 1, 2, ..., m, \tag{19.18}$$

where $k_j \equiv \mathrm{d}F_j[0]/\mathrm{d}E_j$, $a_{jk} \equiv \partial E_j/\partial p_k$, and p_j^e indicates the equilibrium value of the j-th price.

For the moment we make the simplifying assumption that $k_j = 1$ for all j, so that our system can be written as

$$\overline{\mathbf{p}}' = \mathbf{A}\overline{\mathbf{p}}, \tag{19.19}$$

where

$$\overline{\mathbf{p}} = \begin{bmatrix} p_1 - p_1^e \\ p_2 - p_2^e \\ \dots \\ p_m - p_m^e \end{bmatrix}, \ \mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mm} \end{bmatrix}. \tag{19.20}$$

Dynamic stability means $\lim_{t\to\infty}(p_j - p_j^e) = 0$, and this in turn is equivalent to all the latent roots of $\mathbf{A}$ having negative real parts. The elements of $\mathbf{A}$ are the same as those of the Jacobian appearing in the static stability analysis, but the static stability conditions are not, in general, either necessary or sufficient for the roots of $\mathbf{A}$ to be stable. The numerical counter-examples built by Samuelson have been examined as exercises in the previous chapter. The example exercise shows that the perfect stability conditions are not sufficient. Exercise 1 shows that both imperfect and perfect stability conditions are not necessary. Exercise 2 shows that imperfect stability conditions are not sufficient. We just add that the reason for building that complicated counter-example to the static perfect stability conditions lies in the fact that with 2×2 systems the static perfect stability conditions are indeed sufficient for dynamic stability. In a 2×2 system, in fact, we have

$$a_{11} < 0, a_{22} < 0, \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0$$

as static perfect stability conditions. Since the characteristic equation of $\mathbf{A}$ is

$$\lambda^2 - (a_{11} + a_{22})\lambda + (a_{11}a_{22} - a_{12}a_{21}) = 0,$$

the static perfect stability conditions are indeed sufficient for the roots of the characteristic equation to have negative real parts. This is no longer generally true in higher-dimensional systems, except for a few particular cases:

(i) the first case, pointed out by Samuelson, is the case of symmetry, i.e. $a_{ij} = a_{ji}$. In this case—see Chap. 18, Sect. 18.2.2, conditions I—the alternation in the signs of the principal minors of $\mathbf{A}$ is necessary and sufficient for the latent roots to be stable. Symmetry means that for all (i, j) the effect of the i-th price on the excess demand for the j-th good is the same as the effect of the j-th price on the excess demand for the i-th good. This is a rather strong requirement and we cannot expect it to be generally satisfied. It would be satisfied, for example, if in the Slutsky decomposition of the price effect into an income effect and a substitution effect, the income effect were negligible (substitution effects are, in fact, symmetric).

(ii) the second case was pointed out by Metzler (1945), and is the gross substitutability case. By 'gross substitutability' we mean that each excess demand responds negatively to an increase in its 'own' price and positively to an increase in the price of any other good. In formal terms this means that $a_{ii} < 0$, $a_{ij} > 0$. We know from conditions III in Sect. 18.2.2 that also in this case the alternation in the signs of the principal minors of $\mathbf{A}$ is necessary and sufficient for the latent roots to be stable. The Metzler conditions can be weakened to $a_{ij} \geq 0$. However, gross substitutability seems, also in its weakened form, a rather strong requirement. It should also be pointed out that, when we are in a general equilibrium system, gross substitutability is by itself sufficient to guarantee stability of equilibrium, without having to check conditions (19.16.). This depends on certain properties of general equilibrium systems (see below, Chap. 22) and cannot be generalized (i.e., conditions $a_{ii} < 0$, $a_{ij} > 0$ are not sufficient stability conditions unless we are in a general equilibrium setting).

(iii) the third case is perhaps closest to what Walras had in mind when he tried to show that his tâtonnement process would indeed bring the system to equilibrium. He assumed that the influence of the 'own' price was equilibrating (this corresponds to $a_{jj} < 0$ in modern terms). He was well aware of the indirect influences of the changes in the other prices, but he assumed that these influences were some equilibrating and some disequilibrating, so that up to a certain point they cancelled each other out; hence the prevailing effect was the stabilizing one of the 'own' price. This sounds very close to the modern case of negative dominant diagonal that, taken as row dominance, implies that the equilibrating effect of the element on the main diagonal (the 'own' price effect) somehow dominates all the effects of the other elements (representing the influences of the other prices). A negative dominant diagonal is by itself sufficient for stability of equilibrium and implies conditions (19.16), as we know from conditions IV and V in Sect. 18.2.2. A dominant negative diagonal is by no means a preposterous assumption, hence Walras' intuition was right.

Let us now turn to the question of adjustment speeds. From the results on D-stability (see Sect. 18.2.2.1) we know that in the cases of stability examined here (symmetry, Metzlerian, dominant diagonal) the system remains stable for any positive set of adjustments speeds k_j. From the same results we also know that, given an arbitrary matrix $\mathbf{A}$, the conditions for perfect stability, i.e., conditions (19.16), are necessary (but not sufficient) if equilibrium is to be stable for any choice of positive adjustment speeds (this was already pointed out by Metzler, 1945).

Finally, let us suppose that the adjustment speeds are 'small' and that the markets can be ranked in order of decreasing speed of adjustment. Then, it follows from the Fisher-Fuller theorem on the stabilization of matrices (see Sect. 18.2.2.1) that Hicks' perfect stability conditions are sufficient for dynamic stability: in fact, the assumptions of 'small' speeds of adjustment and of their ranking are the economic counterpart of the conditions $k_i/k_{i-1} < \epsilon$ (a matrix satisfying the perfect stability conditions is called a Hicksian matrix). For further evaluation of Hicks' contribution see Hahn (1991).

19.2 Human capital in a growth model

We have already examined neoclassical growth theory in Chap. 13, Sect. 13.2. Let us recall that growth in the basic neoclassical model is exogenous: the steady state path, in fact, depends on factors such as the rate of growth of the labour force and technical progress. Both are exogenous: the labour force grows according to exogenous demographic factors, and technical progress is no more than an exogenous time trend.

Technical progress can be endogenized in several ways, for example by introducing R&D activity, or investment in human capital. A simple neoclassical model in which human capital plays a crucial role has been presented by Mankiw et al. (1992) with a view to empirical testing of the convergence hypothesis (see Chap. 13, Sect. 13.2.4.2) and without presenting a formal analysis of the model. Here we shall study the formal stability properties of the model in depth, which will allow us to illustrate a few interesting mathematical points concerning stability.

Let the production function be first-degree homogeneous in physical capital, human capital, and effective labour, i.e., assuming a Cobb-Douglas specification,

$$Y(t) = K(t)^{\alpha} H(t)^{\beta} (A(t)L(t))^{1-\alpha-\beta}, \tag{19.21}$$

where $Y(t)$ is output, $K(t)$ the stock of physical capital, $H(t)$ the stock of human capital, $L(t)$ the labour force assumed to grow exogenously at the rate n, $A(t)$ the level of labour-augmenting technical progress assumed to grow exogenously at the rate g (so that the number of effective units of labour grows at rate $n+g$). The authors assume that the same production function holds for physical capital, human capital, and consumption. In addition, they assume that human capital depreciates at the same rate δ as physical capital. More sophisticated models in which the production function for human capital is different from that for other goods are described in Barro and Sala-i-Martin (1995).

The model assumes that a constant fraction of output, s, is saved and invested, partly in physical and partly in human capital, according to given proportions, s_k, s_h, such that $s_k + s_h = s$. Thus, with an obvious extension of the basic model in which only physical capital is present (see Chap. 13, Sect. 13.2), the evolution of the economy is determined by

$$\begin{aligned} k'(t) &= s_k y(t) - (n+g+\delta)k(t), \\ h'(t) &= s_h y(t) - (n+g+\delta)h(t), \end{aligned} \tag{19.22}$$

where $y = Y/AL, k = K/AL, h = H/AL$ are quantities per effective unit of labour (we have changed the symbology used in Sect. 13.2 to conform with that used by Mankiw et al.).

From the production function (19.21) we obtain

$$y = \frac{K^{\alpha} H^{\beta} (AL)^{1-\alpha-\beta}}{AL} = k^{\alpha} h^{\beta}, \tag{19.23}$$

so that the basic differential system becomes

$$\begin{aligned} k' &= s_k k^{\alpha} h^{\beta} - (n+g+\delta)k, \\ h' &= s_h k^{\alpha} h^{\beta} - (n+g+\delta)h. \end{aligned} \tag{19.24}$$

The steady state equilibrium path will be determined by the conditions $k' = h' = 0$, from which

$$\begin{aligned} s_k k^{\alpha} h^{\beta} - (n+g+\delta)k = 0, \\ s_h k^{\alpha} h^{\beta} - (n+g+\delta)h = 0. \end{aligned} \tag{19.25}$$

Taking logarithms and rearranging terms we have the log-linear system

$$\begin{aligned} (\alpha-1)\ln k + \beta \ln h &= \ln(n+g+\delta) - \ln s_k, \\ \alpha \ln k + (\beta-1)\ln h &= \ln(n+g+\delta) - \ln s_h, \end{aligned} \tag{19.26}$$

from which we get, solving for $\ln k, \ln h$ and then reverting to k, h, the equilibrium values

$$\begin{aligned} k^* &= \left(\frac{s_k^{1-\beta} s_h^{\beta}}{n+g+\delta}\right)^{1/(1-\alpha-\beta)}, \\ h^* &= \left(\frac{s_k^{\alpha} s_h^{1-\alpha}}{n+g+\delta}\right)^{1/(1-\alpha-\beta)}. \end{aligned} \tag{19.27}$$

From Eqs. (19.27) and (19.23) we obtain $y^* = (k^*)^{\alpha}(h^*)^{\beta}$.

To study the stability of the equilibrium point we have to consider system (19.24), which is non-linear. It is however possible to perform a graphical analysis which will show that the equilibrium is stable. For convenience we introduce the variables

$$z_1 = \frac{k}{k^*}, \quad z_2 = \frac{h}{h^*}, \tag{19.28}$$

and, since from (19.25) we have

$$\begin{aligned} k^* &= \frac{s_k}{n+g+\delta} k^{*\alpha} h^{*\beta}, \\ h^* &= \frac{s_h}{n+g+\delta} k^{*\alpha} h^{*\beta}, \end{aligned} \tag{19.29}$$

we can rewrite (19.24) as

$$\begin{aligned} z_1' &= c(z_1^{\alpha} z_2^{\beta} - z_1), \\ z_2' &= c(z_1^{\alpha} z_2^{\beta} - z_2), \end{aligned} \tag{19.30}$$

where $c = n + g + \delta$.

We now study system (19.30) graphically to show that the equilibrium point—that in terms of z_1, z_2 is $(1, 1)$—is stable. For this purpose let

$$\begin{aligned} \varphi_1(z_1, z_2) &= c(z_1^{\alpha} z_2^{\beta} - z_1), \\ \varphi_2(z_1, z_2) &= c(z_1^{\alpha} z_2^{\beta} - z_2), \end{aligned} \tag{19.31}$$

and begin with the study of the curves $\varphi_1 = 0, \varphi_2 = 0$ in the positive quadrant.

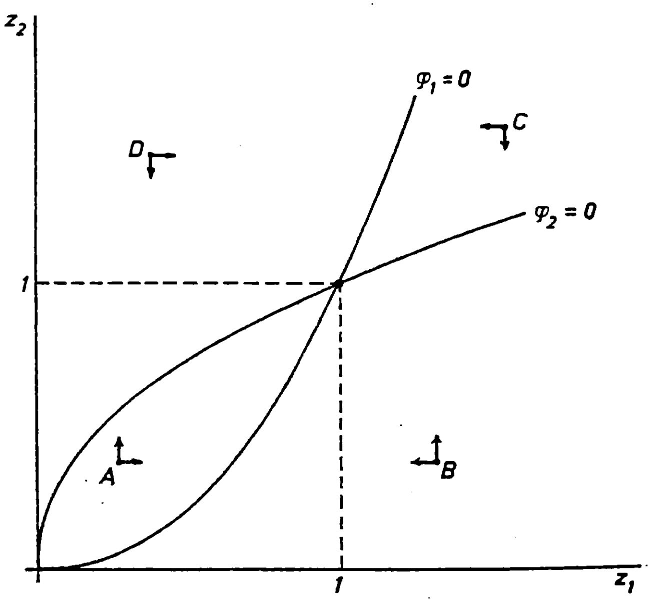


Figure 19.1: Stability of growth equilibrium with human capital

From $\varphi_1 = 0$ we can express z_2 as a function of z_1, obtaining

$$\left(z_2 = z_1^{(1-\alpha)/\beta}\right)_{\varphi_1=0}, \tag{19.32}$$

whence

$$\begin{aligned} \left(\frac{dz_2}{dz_1}\right)_{\varphi_1=0} &= \frac{1-\alpha}{\beta} z_1^{(1-\alpha-\beta)/\beta}, \\ \left(\frac{d^2 z_2}{dz_1^2}\right)_{\varphi_1=0} &= \frac{(1-\alpha)(1-\alpha-\beta)}{\beta^2} z_1^{(1-\alpha-2\beta)/\beta}. \end{aligned} \tag{19.33}$$

Both derivatives are positive for $z_1 > 0$, since $1 - \alpha > 0$ and $1 - \alpha - \beta > 0$. Besides,

$$\begin{aligned} \lim_{z_1 \to 0} \left(\frac{dz_2}{dz_1}\right)_{\varphi_1=0} &= 0, \\ \lim_{z_1 \to \infty} \left(\frac{dz_2}{dz_1}\right)_{\varphi_1=0} &= \infty. \end{aligned} \tag{19.34}$$

From (19.33) and (19.34) it follows that the curve $\varphi_1 = 0$ has the shape shown in Fig. 19.1.

It is now important to note that we have $\varphi_1 \gtrless 0$ respectively above (below) the curve $\varphi_1 = 0$. This follows from the fact that

$$\frac{\partial \varphi_1}{\partial z_2} = c\beta z_1^{\alpha} z_2^{\beta-1} > 0, \tag{19.35}$$

whence, taking an arbitrary point on the curve $\varphi_1 = 0$ and considering any point vertically above it, the value of φ_1 in the new point will be $\varphi_1 + d\varphi_1 = d\varphi_1 = \frac{\partial \varphi_1}{\partial z_2} dz_2 > 0$.

Let us now consider the $\varphi_2 = 0$ curve. By expressing z_2 in terms of z_1 and proceeding as before we have

$$\left(z_2 = z_1^{\alpha/(1-\beta)}\right)_{\varphi_2=0}, \tag{19.36}$$

whence

$$\left(\frac{dz_2}{dz_1}\right)_{\varphi_2=0} = \frac{\alpha}{1-\beta} z_1^{(\alpha+\beta-1)/(1-\beta)}, \tag{19.37}$$

$$\left(\frac{d^2 z_2}{dz_1^2}\right)_{\varphi_2=0} = \frac{\alpha(\alpha+\beta-1)}{\beta^2} z_1^{(\alpha+2\beta-2)/(1-\beta)}, \tag{19.38}$$

$$\lim_{z_1 \to 0} \left(\frac{dz_2}{dz_1}\right)_{\varphi_2=0} = \infty, \tag{19.39}$$

$$\lim_{z_1 \to \infty} \left(\frac{dz_2}{dz_1}\right)_{\varphi_2=0} = 0, \tag{19.40}$$

and so the function $\varphi_2 = 0$ has the shape shown in Fig. 19.1.

It is also important to note that

$$\frac{\partial \varphi_2}{\partial z_1} = c\alpha z_1^{\alpha-1} z_2^{\beta} > 0, \tag{19.41}$$

so that $\varphi_2 \gtrless 0$ respectively to the right (left) of the curve $\varphi_2 = 0$.

We now have all the ingredients required. The two curves obviously intersect at the equilibrium point $(1,1)$. Let us now consider disequilibrium points. Representative points are A, B, C, D.

Point A lies *above* the curve $\varphi_1 = 0$, therefore (since $z_1' = \varphi_1$ by definition, see above) at point A we have $z_1' > 0$, and z_1 increases. This is shown by the horizontal arrow pointing rightwards from A. We also see that point A lies *to the right* of the curve $\varphi_2 = 0$, therefore we have $z_2' = \varphi_2 > 0$, and z_2 increases as well (upwards pointing arrow). The student may check in a similar way the directions of the arrows originating from points B, C, D.

Thus the arrows point toward equilibrium in any case, and this gives an intuitive idea of the stability properties of the model, since the trajectory of any point will be included between the two arrows originating from it. However, as we shall see in the next section, the use of such 'arrow diagrams' *alone* cannot demonstrate stability rigorously. Since the dynamic model under consideration is non-linear, let us consider a linear approximation around the equilibrium point (1,1). This is performed by expanding in Taylor's series and neglecting all terms of order higher than the first (for a general treatment of this procedure see Chap. 21, Sect. 21.4.21).

Thus we have

$$\begin{aligned}\varphi_1(z_1,z_2) &\simeq \varphi_1(1,1) + \left(\frac{\partial \varphi_1}{\partial z_1}\right)_{(1,1)}(z_1-1) + \left(\frac{\partial \varphi_1}{\partial z_2}\right)_{(1,1)}(z_2-1)\\ &= 0 + \left(c(\alpha z_1^{\alpha-1}z_2^{\beta}-1)\right)_{(1,1)}(z_1-1) + \left(c\beta z_1^{\alpha}z_2^{\beta-1}\right)_{(1,1)}(z_2-1)\\ &= c(\alpha-1)\overline{z}_1 + c\beta\overline{z}_2, \end{aligned} \tag{19.42}$$

where the overbar denotes the deviations from equilibrium (i.e., $\overline{z}_i = z_i - 1$). Proceeding in a similar way for $\varphi_2(z_1,z_2)$, and noting that $\overline{z}_i' = z_i'$, we have the linear system

$$\begin{aligned}\overline{z}_1' &= c(\alpha-1)\overline{z}_1 + c\beta\overline{z}_2,\\ \overline{z}_2' &= c\alpha\overline{z}_1 + c(\beta-1)\overline{z}_2.\end{aligned} \tag{19.43}$$

This is a first-order system in normal form, whose matrix

$$\begin{bmatrix} c(\alpha-1) & c\beta \\ c\alpha & c(\beta-1)\end{bmatrix} \tag{19.44}$$

has a negative dominant diagonal. In fact, from the production function, $\alpha < 1, \beta < 1, 1-\alpha-\beta > 0$. This gives $\alpha - 1 < 0, \beta - 1 < 0, |\alpha-1| = 1-\alpha > \beta$, $|\beta - 1| = 1-\beta > \alpha$ (row dominance). Hence, according to conditions IV in Chap. 18, Sect. 18.2.2, the latent roots of matrix (19.44) have negative real parts and the equilibrium point is stable.

19.3 A digression on 'arrow diagrams'

In the previous section we have shown the use of arrow diagrams (we have called them so—see Gandolfo, 1971, Part II, Chap. 9, § 4—for lack of a better name) as a visual aid for the study of dynamic stability. These diagrams were actually introduced by an economist and precisely by A. Marshall in his study of the stability of equilibrium in international trade (Marshall, 1877; Gandolfo, 1994). The first modern writer to have used arrow diagrams again seems to be Metzler (1951). They have been popularized in international monetary economics by Mundell in the 1960s and since then widely used in economics in all cases in which there is a 2×2 dynamic model to be analysed.

We have stated in the previous section that such arrow diagrams, *by themselves alone*, cannot be used to study the dynamic stability of the equilibrium point, since they cannot tell us whether the initial point actually converges, and how, to the equilibrium point. Here we generically call y_1, y_2 the two variables concerned.

The precise trajectory of the initial non-equilibrium point can be drawn *only* after having formalized the problem by means of a system of functional equations, whose solution yields the time path of each variable. Knowing these time paths, we can draw the trajectory of the system as a whole (as represented by the movement of a point); this trajectory is actually a curve expressed in parametric form. In the terminology of non-linear analysis we are dealing with a phase path (see Part III, Chap. 21, Sect. 21.3.2, where we shall also show how a trajectory can be obtained by finding the integral curves of the system without actually solving it with respect to time), but for the present purposes it is enough to recall that the parametric equations of a curve simply express the coordinates of each point of the curve in the (y_1, y_2) plane as functions of a parameter (in our case the parameter is time). Formally we have

$$y_1 = y_1(t), \quad y_2 = y_2(t), \tag{19.45}$$

where the functions $y_1(t), y_2(t)$, are given by the solution of the system of functional equations.

In Fig. 19.2 we have schematically drawn the equilibrium point E and the initial point P: from the economic point of view, the system under examination can be any dynamic model involving only two variables to be expressed as functions of time.

In the second and fourth quadrants we have drawn the functions $y_1(t)$, $y_2(t)$ obtained as the solution of a functional system; these functions have been assumed both monotonically convergent for simplicity. Given the unit of time on the relevant axes, we can find, by means of the self-explanatory graphical procedure depicted in the diagram, the successive points P_1, P_2, P_3, ..., corresponding to times $t = 1, 2, 3, \ldots$, and so build the trajectory of point P with the desired degree of approximation. This trajectory will be monotonically convergent to E, given the monotonic convergence of $y_1(t), y_2(t)$.

It should now be stressed that by using the arrows alone we could not draw this trajectory. Even if both arrows point toward equilibrium, we cannot be sure that the equilibrium point is stable. To draw the trajectory from the arrows alone we would need *not only the direction, but also the length of the arrows, that represent vectors.* In the physicists' parlance, to apply the

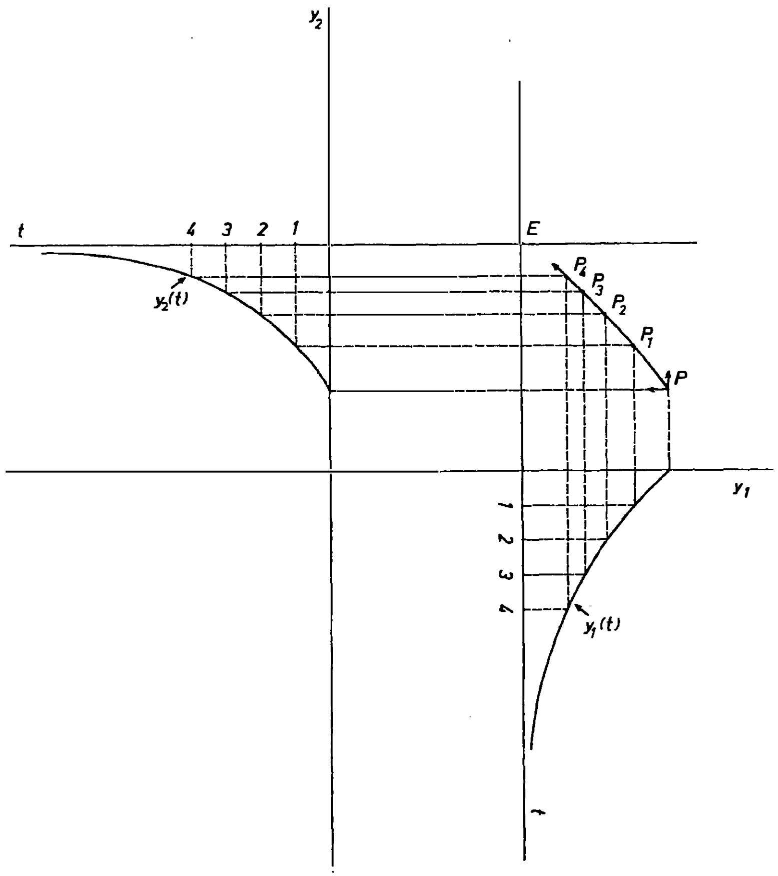


Figure 19.2: Construction of an arrow diagram

elementary 'composition of forces' or 'parallelogram' rule we need to know not only the direction but also the length of the vectors involved, and this is equivalent to the parametric equations that we have used.

The main purpose of this digression is to warn the student against using

arrow diagrams *alone* to analyse a stability problem. Such diagrams can at most be used to obtain a first idea of the situation, but to obtain rigorous results the problem must always be analysed by means of functional equations: if not, wrong conclusions are likely to be drawn, especially in non-normal cases[1].

Of course, it is perfectly legitimate to use arrow diagrams after the problem has been formally solved: then they are a useful expository device, and very convenient in explaining the results in an intuitive manner for the non-mathematical reader.

Let us note, finally, that when the dynamic system to be analysed is linear, no problem exists in finding its explicit solution. When the system is non-linear, the usual procedure—an application of which has been shown in the previous section—is to linearise the system around the equilibrium point and examine the resulting linear system. By this procedure we shall obtain only *local* stability results. More advanced methods are described in Chaps. 21-23.

19.4 Balanced growth in a multi-sector economy

Contrary to the current boom of research on aggregate growth models, multi-sector models of economic growth have not yet drawn the attention of the new growth economists—actually, true multi-sector models (as distinct from simple two-sector models) are not even mentioned in the current standard reference work on economic growth (Barro and Sala-i-Martin, 1995).

However, the predominant interest in aggregate models should not obscure the fact that reality *is* multi-sectoral, and that the structure of the economy plays a crucial role in the phenomenon of economic growth. Actually, mathematical models of multi-sector growth antedate the Harrod-Domar and Solow-Swan aggregate models. The celebrated von Neumann (1937) general equilibrium model and Leontief's (1953) dynamic input-output model are the precursors.

In this section we give a brief account of the dynamic input-output model, which also serves to illustrate some interesting mathematical points treated in Chap. 18.

Let us first recall that in the basic static input-output model the interdependency among the various sectors of the economy is represented by a set of constant technical coefficients $a_{ik} \geq 0$. Each of these is the quantity of the i-th good currently used up in the production of one unit of the k-th

[1] Even Marshall—who had an engineer's background, but believed that arrow diagrams were not only a perfect substitute for differential equations, but actually a better tool—was led by them to make some wrong statements. See Gandolfo, 1971, note * on p. 312.

good. Thus the coefficients a_{ik} refer to flows of good currently used up by the various industries and say nothing about capital formation.

In addition to the flows just mentioned, there are other flows of goods that are not embodied in current output, but serve to increase the stock of capital, both fixed and circulating. The basic assumption is that the technology requires fixed coefficients, say $b_{ik} \geq 0$, also as regards capital requirements (this is called a Leontief production function). Each b_{ik} represents the stock of the i-th good (in its quality of capital good) that the k-th industry must have at hand for each unit of its output. Of course, not each good can simultaneously be an intermediate good and a capital good, so that some of the b's as well as some of the a's will be zero.

If we denote by S_{ik} the stock of the i-th good *qua* capital good owned by the k-th industry, and by X_k total output of the k-th good (industry), we have the relations

$$S_{ik} = b_{ik}X_k, \ i, k = 1, 2, ..., m, \tag{19.46}$$

which incorporates the assumption of instantaneous adjustment, i.e. the capital stock is always the appropriate one for producing the current output. By differentiating with respect to time it follows that

$$S'_{ik} = b_{ik}X'_k, \tag{19.47}$$

which links the variation in the capital stocks (i.e., investment) to the output variation through the appropriate capital coefficients. This is a multi-sector extension of the acceleration principle.

The basic equations of the model are given by

$$X_i = \sum_{k=1}^{m} a_{ik}X_k + \sum_{k=1}^{m} b_{ik}X'_k + Y_i, \ i = 1, 2, ..., m, \tag{19.48}$$

which state the equilibrium conditions that total current output of the i-th good should equal total current demand. This latter is made up of three components:

(i) the quantity of the i-th good currently used as an intermediate input by all industries, that is $\sum_{k=1}^{m} a_{ik}X_k$;

(ii) the quantity of the i-th good currently demanded for investment purposes by all industries, that is $\sum_{k=1}^{m} b_{ik}X'_k$;

(iii) the quantity of the i-th good currently used to satisfy final demand, that is Y_i.

As in the static model (which simply is the same model without the second component) we can distinguish an *open* model, in which final demands are given exogenously, and a *closed model,* in which final demands are made endogenous by the introduction of an n-th sector ($n = m+1$), the household sector. The 'output' of this sector is labour, and its 'inputs' are the various

consumption goods, durable and non-durable. In this case Eqs. (19.48) become

$$X_i = \sum_{k=1}^{n} a_{ik}X_k + \sum_{k=1}^{n} b_{ik}X'_k, \ i = 1, 2, ..., m, n, \tag{19.49}$$

where

$$a_{in}X_n + b_{in}X'_n = Y_i.$$

We shall deal with the closed model, since the solution of the corresponding open model—apart from the number of equations—simply requires to find a particular solution, that depends on the assumed form of the final demands.

Two related questions can now be asked. The first is whether the model is capable of *balanced growth*, that is, a state of growth in which the proportions that the variables bear to each other are constant. This, of course, is equivalent to the requirement that all variables grow at the same rate. It should be stressed that we are requiring *actual* balanced growth, not *asymptotic* balanced growth, namely all variables should *actually grow* at the same rate, and not only *tend to grow* at the same rate as $t \to \infty$. This, in fact, is the subject of the second question.

To answer these questions let us first write the basic dynamic system in matrix form:

$$\mathbf{X} = \mathbf{AX} + \mathbf{BX}', \quad \text{or} \quad -\mathbf{BX}' + (\mathbf{I} - \mathbf{A})\mathbf{X} = \mathbf{0}, \tag{19.50}$$

where $\mathbf{A} = [a_{ik}]$ and $\mathbf{B} = [b_{ik}]$. We assume that the input-output matrix $\mathbf{A}$ is indecomposable and that its dominant root μ_{MAX} is smaller than unity. Indecomposability (see Chap. 9, Sect. 9.2.1) means that each good *qua* intermediate good is directly or indirectly required in the production of all the other goods, and this seems a reasonable assumption. $\mu_{MAX} < 1$ means that the static open model has an economically meaningful solution. In fact, the solution of the static open model

$$\mathbf{X} = \mathbf{AX} + \mathbf{Y}$$

is

$$\mathbf{X} = [\mathbf{I} - \mathbf{A}]^{-1}\mathbf{Y},$$

and $[\mathbf{I} - \mathbf{A}]^{-1} > \mathbf{0}$ if, and only if, $\mu_{MAX} < 1$ (Gantmacher, 1959, Chap. III, Sect. 2).

System (19.50) is a first-order system not in normal form, whose characteristic equation is

$$|-\lambda\mathbf{B} + (\mathbf{I} - \mathbf{A})| = |\mathbf{I} - (\mathbf{A} + \lambda\mathbf{B})| = 0. \tag{19.51}$$

To examine our questions we must first of all show that this characteristic equation has (at least) one positive root, since in the contrary case the model would not be capable of growth. Now we observe that Eq. (19.51) can be written as

$$|\rho\mathbf{I}-(\mathbf{A}+\lambda\mathbf{B})| = |(\mathbf{A}+\lambda\mathbf{B})-\rho\mathbf{I}| = 0, \quad \rho = 1, \tag{19.52}$$

which can be interpreted as the characteristic equation of the matrix $\mathbf{C} \equiv (\mathbf{A}+\lambda\mathbf{B})$ having a characteristic root equal to unity.

Since $\mathbf{A}$ is non-negative indecomposable, and $\mathbf{B}$ is non-negative, we can apply the theorem according to which the dominant latent root of a non-negative indecomposable matrix is a continuous and strictly increasing function of its elements (Schwartz, 1961, p.24). Since the dominant root of $\mathbf{A}$ has been assumed smaller than unity, there will exist a unique *positive* λ, say λ_1, such that the dominant root of $\mathbf{C} \equiv (\mathbf{A}+\lambda\mathbf{B})$ is exactly unity.

Thus for $\lambda = \lambda_1 > 0$, Eq. (19.52) will be indeed satisfied with $\rho = \rho_{MAX} = 1$. Since Eq. (19.52) is the same as Eq. (19.51), this shows that a positive root of the characteristic equation of our differential equation system does exit.

But there is more to it than that. Let us consider the equations for the determination of the vector $\boldsymbol{\alpha}$ associated with the root $\lambda = \lambda_1$ of the characteristic equation (19.51) of the dynamic system (19.50), which are (see Chap. 18, Sect. 18.3.1)

$$[-\lambda_1\mathbf{B}+(\mathbf{I}-\mathbf{A})]\boldsymbol{\alpha} = [\mathbf{I}-(\mathbf{A}+\lambda_1\mathbf{B})]\boldsymbol{\alpha} = \mathbf{0}. \tag{19.53}$$

Equations (19.53) coincide with those for the determination of the latent vector associated with the dominant latent root of $\mathbf{C} \equiv (\mathbf{A}+\lambda\mathbf{B})$, namely

$$[\rho\mathbf{I}-(\mathbf{A}+\lambda_1\mathbf{B})]\boldsymbol{\alpha} = \mathbf{0}, \quad \rho = \rho_{MAX} = 1. \tag{19.54}$$

Now, since $\mathbf{A}$ is indecomposable, $\mathbf{C}$ is also indecomposable, hence by the Frobenius theorem the characteristic vector associated with the dominant characteristic root of $\mathbf{C}$, $\rho_{MAX} = 1$ is strictly positive, i.e.

$$\boldsymbol{\alpha} > \mathbf{0}. \tag{19.55}$$

This proves that, in the solution of the differential system (19.50), the terms containing the positive root yield an economically meaningful growth path. Hence the system under consideration is capable of economic growth.

The solution of our differential system is then—see Chap. 18, Eq. (18.130)—

$$\begin{aligned}
X_1(t) &= A_1\alpha_1^{(1)}e^{\lambda_1 t} + A_2\alpha_1^{(2)}e^{\lambda_2 t} + ... + A_m\alpha_1^{(m)}e^{\lambda_m t},\\
X_2(t) &= A_1\alpha_2^{(1)}e^{\lambda_1 t} + A_2\alpha_2^{(2)}e^{\lambda_2 t} + ... + A_m\alpha_2^{(m)}e^{\lambda_m t},\\
\cdots\cdots & \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
X_{nt}(t) &= A_1\alpha_m^{(1)}e^{\lambda_1 t} + A_2\alpha_m^{(2)}e^{\lambda_2 t} + ... + A_m\alpha_n^{(m)}e^{\lambda_m t},
\end{aligned} \tag{19.56}$$

where $\lambda_2, ..., \lambda_m$ are the other characteristic roots, which may also include multiple roots, but we shall neglect this complication, that in any case would not change the results.

Since we want the system to exhibit actual balanced growth, in the solution (19.56) only the terms containing $e^{\lambda_1 t}$ should remain. This is the case treated in Chap. 18, Sect. 18.2.2.3. Let us consider the equations for the determination of the arbitrary constants, which—in the case of system (19.56)—turn out to be

$$X_j(0) = \sum_{i=1}^{m} A_i\alpha_j^{(i)}, \qquad j = 1, 2, ..., m, \tag{19.57}$$

whose solution by Cramer's rule is

$$A_i = \frac{D_i}{D}, \text{where } D \equiv \begin{vmatrix} \alpha_1^{(1)} & \alpha_1^{(2)} & ... & \alpha_1^{(m)} \\ \alpha_2^{(1)} & \alpha_2^{(2)} & ... & \alpha_2^{(m)} \\ ... & ... & ... & ... \\ \alpha_m^{(1)} & \alpha_m^{(2)} & ... & \alpha_m^{(m)} \end{vmatrix}, \tag{19.58}$$

and D_i is obtained by substituting the column of $X_j(0)$ in the place of the ith column of D.

Then, if we choose $X_1(0), X_2(0), ..., X_m(0)$ proportional to $\alpha_1^{(1)}, \alpha_2^{(1)}, ..., \alpha_m^{(1)}$, it follows that each D_i (except D_1) will have two proportional columns, and so $A_i = 0, i \neq 1$.

Let us now consider the case of asymptotic balanced growth. In formal terms,

$$\lim_{t\to\infty}\frac{X_i'(t)}{X_i(t)} = \lambda_1, \quad \text{or} \quad \lim_{t\to\infty}\frac{X_i(t)}{A_1\alpha_i^{(1)}e^{\lambda_1 t}} = 1, \tag{19.59}$$

namely all outputs tend to grow at the same rate λ_1, i.e. the actual path converges to the balanced growth path. A system satisfying (19.59) is called *relatively stable*. Given the properties of exponential functions, conditions (19.59) will be satisfied if, and only if, λ_1 is the dominant root, i.e., if, and only if

$$\lambda_1 > |\lambda_j|, \quad j = 2, 3, ..., m. \tag{19.60}$$

Up to this point we have considered only the quantity side of the model. We now turn to the determination of prices. It has been shown by Solow that, in formulating the price equations, account must be taken of capital losses or gains, i.e. of the variations in the value of the stocks of capital goods brought about by price changes. Thus the price level in each sector equals the cost of the current inputs plus interest charges plus capital losses (or minus capital gains). The price equations are then

$$\mathbf{p} = \mathbf{A}^{\mathrm{T}}\mathbf{p} + r\mathbf{B}^{\mathrm{T}}\mathbf{p} - \mathbf{B}^{\mathrm{T}}\mathbf{p}', \tag{19.61}$$

where r is the interest rate and $\mathbf{p}$ the vector of prices; the superscript T denotes transposition. The characteristic equation of system (19.61) is

$$\begin{aligned} \left|\eta\mathbf{B}^{\mathrm{T}} - r\mathbf{B}^{\mathrm{T}} + (\mathbf{I} - \mathbf{A}^{\mathrm{T}})\right| &= \left|-(r-\eta)\mathbf{B}^{\mathrm{T}} + (\mathbf{I} - \mathbf{A}^{\mathrm{T}})\right| \\ &= \left|\mathbf{I} - (\mathbf{A}^{\mathrm{T}} + (r-\eta)\mathbf{B}^{\mathrm{T}})\right| = 0. \quad (19.62) \end{aligned}$$

Given that a matrix and its transpose have the same latent roots, it is apparent that the roots η of (19.62) and the roots λ of (19.51) are related by

$$\lambda = r - \eta, \quad \text{or} \quad \eta = r - \lambda. \quad (19.63)$$

This allows us to prove the so-called *dual stability theorem* (Jorgenson, 1960): if the output system is relatively stable (in the sense defined above), the price system is relatively unstable, and vice versa.

In fact, by the same procedure followed in the study of the output system, we find a $\eta_1 > 0$ such that the associated vector $\boldsymbol{\alpha}$ is positive, hence the condition for relative stability of the price system is

$$\eta_1 > |\eta_j| \quad j = 2, 3, ..., m, \quad (19.64)$$

i.e., given Eq. (19.63),

$$r - \lambda_1 > |r - \lambda_i| \geq r - |\lambda_i|, \quad (19.65)$$

from which

$$\lambda_1 < |\lambda_i|. \quad (19.66)$$

This is exactly the opposite of Eq. (19.60), hence the theorem is proved.

19.5 Exercises

1. Suppose that in a general equilibrium system excess demand for the j-th good also depends on the expected price of the same good, $\hat{p}_j$, so that $E_j = E_j(p_1, p_2, ..., p_m, \hat{p}_j)$. In equilibrium, $E_j = 0$ and $\hat{p}_j = p_j$. Assume that expectations are formed according to the relations $\hat{p}_j = p_j + \rho_j p_j'$, where $\rho_j \gtreqless 0$ is an expectation coefficient. Consider the linearised dynamic system

$$\overline{p}_j' = k_j \left(\sum_{i=1}^{m} a_{ji}\overline{p}_j + b_j \overline{\hat{p}}_j \right),$$

and assume that the system with static expectations ($\rho_j = 0$) fulfils the Metzler conditions. Then prove the following:

(1.a) Extrapolative expectations ($\rho_j > 0$) coupled with a negative effect of $\hat{p}_j$ on E_j (i.e., $b_j < 0$) maintain stability.

(1.b) Conservative expectations ($\rho_j < 0$) coupled with a positive effect of $\hat{p}_j$ on E_j (i.e., $b_j > 0$) maintain stability.

(1.c) In the other cases, the system remains stable when the speed of adjustment is sufficiently low or, given the speed of adjustment, when the expectations coefficient is sufficiently small in absolute value.

(Hint: the general condition is $1/k_j > b_j \rho_j$).

2. Consider a single market in disequilibrium. Then forces will operate on both price and quantity simultaneously. Assume that prices adjust according to the Walrasian hypothesis while quantities adjust according to the Marshallian hypothesis (see Chap. 13, Sect. 13.1). Thus we have the 2×2 system

$$\begin{aligned} p' &= c[D(p) - q], \\ q' &= k[p - p_s(q)], \end{aligned}$$

where q is the quantity supplied and $p_s(q)$ is the supply price as defined in Sect. 13.1. Assuming linear demand and supply functions, i.e. $D = a + bp$, $p_s = -\frac{a_1}{b_1} + \frac{1}{b_1}q$, examine the stability of the system both in the normal ($b < 0$, $b_1 > 0$) and in the various abnormal cases. (Hint: see Beckmann and Ryder, 1969. For a multi-market generalization see Mas-Colell, 1986).

3. Examine the price system in Leontief's multi-sector growth model and show that the prices corresponding to the balanced growth path for outputs are constant if, and only if, the rate of interest is equal to the rate of growth of outputs on the balanced growth path, i.e. $r = \lambda_1$.

4. Consider a simple two-sector growth model with Leontief technology (fixed technical coefficients) and no technical progress. Sector 1 produces the capital good K (assumed to be non-depreciating), which is used as a factor, together with labour, in both sector 1 and sector2 according to a Leontief technology (a_1, a_2 are the fixed capital/labour ratios in the two sectors). Sector 2 produces the consumption good X_2, which is entirely consumed by workers (capitalist save and invest all their income, workers do not save). The labour force N grows at the exogenous rate n and is paid a fixed wage rate W. The basic equations of the model are

$$\begin{aligned} K &= a_1 N_1 + a_2 N_2, \\ X_2 &= W(N_1 + N_2), \\ N_0 e^{nt} &= N_1 + N_2, \\ K' &= X_1, \end{aligned}$$

which express, in the order, full employment of the capital stock, equilibrium in the consumption good market, full employment of the labour force, equilibrium in the capital good market.

(4.a) Show that the model gives rise to the following first-order system of differential equations not in normal form

$$\begin{aligned} a_1N_1' + a_2N_2' - \tfrac{1}{b_1}N_1 &= 0, \\ N_1' + N_2' - N_1 - N_2 &= 0. \end{aligned}$$

(4.b) Show that the roots of the characteristic equation are $\lambda_1 = n, \lambda_2 = 1/(a_1 - a_2)b_1$. What happens when $a_1 = a_2$?

(4.c) Determine the arbitrary constants, given the initial conditions $K = K_0, N = N_0$ for $t = 0$.

(4.d) Show that the model is capable of actual balanced growth.

(4.e) Prove the *capital-intensity condition,* i.e., the balanced growth path is stable (both in the absolute and in the relative sense) if, and only if, the consumption good sector is *more* capital intensive than the capital good sector ($a_2 > a_1$). (Hint: see Shinkai, 1960).

5. Consider the following typical target-instrument policy problem

$$\mathbf{x} = \mathbf{A}\mathbf{y}, \tag{19.67}$$

where $\mathbf{x}$ is the vector of the deviations of the n targets from equilibrium, $\mathbf{y}$ is the vector of the deviations of the n instruments from equilibrium, and $\mathbf{A}$ is the $n \times n$ matrix representing the reduced form of the underlying model. The number of independent instruments has been assumed equal to the number of targets. Assume now that the instruments are changed in response to the disequilibria in the targets according to the following general dynamic scheme

$$\mathbf{y}' = \mathbf{K}\mathbf{x}, \tag{19.68}$$

where the elements of the matrix $\mathbf{K} = [k_{ij}]$ represent the weight of the j-th target in the determination of the adjustment of the i-th instrument. The assignment scheme (or decentralized policy making) is a particular case, in which each instrument is managed in relation to only one target. In this case it is always possible, by appropriately renumbering the variables and rearranging the equations if the case, to write the dynamic system (19.69) as

$$\mathbf{y}' = \mathbf{k}\mathbf{x}, \tag{19.69}$$

where $\mathbf{k} = \mathrm{diag}\{k_{11}, k_{22}, ..., k_{mm}\}$ is a diagonal matrix. Substituting from Eq. (19.67) into (19.69) we have

$$\mathbf{y}' = \mathbf{k}\mathbf{A}\mathbf{y}. \tag{19.70}$$

The problem is then to choose the policy parameters k_{ii} in an economically meaningful way so that system (19.70) is stable. (Hint: use the Fisher-Fuller theorem. See also Petit, 1990).

6. As a specification of the previous problem, consider the simple Mundellian extension of the standard IS-LM model to an open-economy under fixed exchange rates:

$$\begin{aligned} A(Y,r) + X_0 - M(Y) + G - Y &= 0, \\ X_0 - M(Y) + E_0 + K(r) - B &= 0, \end{aligned} \tag{19.71}$$

where A is aggregate expenditure by residents, depending on income and the interest rate ($A_y > 0$, $A_r < 0$), X_0 exports (exogenously given), M imports ($0 < M_y < 1$), G government expenditure, Y national income (output), E_0 exogenous elements in the balance of payments (transfers, etc.), K international capital flows ($0 < K_r < \infty$, i.e. capital mobility is not infinite), and B the balance of payments. G and r are instruments (to simplify the analysis we assume that the money supply is always adequate to the level of the rate of interest chosen, hence we can neglect the monetary equilibrium condition). The targets are a given 'full employment' level of Y, say Y_F, and a balance of payments in equilibrium, $B = 0$.

(6.a) Linearise Eqs. (19.71) at the equilibrium point ($Y = Y_F, B = 0$) so as to obtain Eqs. (19.67).

(6.b) Suppose that the government wants to pair instruments and targets, i.e. to assign each instruments to the pursuance of one target only. There are two possible pairings (government expenditure with the internal and the interest rate with the external target, or the other way round). Determine the pairing and the rules for the management of each instrument (e.g., government expenditure increases if there is unemployment, the interest rate increases if $B < 0$—formally, $G' = k_G(Y_F - Y)$, $k_G > 0$; $r' = k_rB, k_r < 0$, etc.) so that equilibrium is stable.

19.6 References

Arrow, K.J. and F.H. Hahn, 1971, *General Competitive Analysis.*

Barro, R.J. and X. Sala-i-Martin, 1995, *Economic Growth.*

Beckmann, M.J. and H.E. Ryder, 1969, Simultaneous Price and Quantity Adjustment in a Single Market, Sect. 1.

Courant, R. 1961(1937), *Differential and Integral Calculus,* Vol. I, p. 75.

Enthoven, A.C. and K.J. Arrow, 1956, A Theorem on Expectations and the Stability of Equilibrium.

Gandolfo, G., 1971, *Mathematical Methods and Models in Economic Dynamics,* Part II, Chap. 9, §§ 3, 4.

Gandolfo, G., 1994, *International Economics I: The Pure Theory of International Trade,* Chap. 3, Sects. 3.4.2 and A.3.4.

Gantmacher, F.R., 1959, *Applications of the Theory of Matrices,* pp. 83, 88.

Hababigahi, H. and J. Quirk, 1973, Hicksian Stability and Walras' Law.

Hahn, F.H., 1991, Hicksian Themes on Stability.
Hicks, J.R., 1939, 1946, *Value and Capital,* Chaps. IV, V and relative Mathematical Appendices; Additional Note C.
Jaffé, W., 1967, Walras' Theory of *Tâtonnement*: A Critique of Recent Interpretations.
Jorgenson, D.W., 1960, A Dual Stability Theorem.
Jorgenson, D.W., 1961a, Stability of a Dynamic Input-Output System.
Jorgenson, D.W., 1961b, The Structure of Multi-Sector Dynamic Models.
Lange, O., 1952, *Price Flexibility and Employment,* Appendix, pp. 91-99.
Leontief, W. et al., 1953, *Studies in the Structure of the American Economy,* Chap. 3.
Mankiw, N.G., D. Romer and D.N. Weil, 1992, A Contribution to the Empirics of Economic Growth.
Marshall, A., 1937 (1877), *The Pure Theory of Foreign Trade.*
Mas-Colell, A., 1986, Notes on Price and Quantity Tâtonnement Dynamics.
McFadden, D., 1968, On Hicksian Stability.
McManus, M., 1963, Notes on Jorgenson's Model.
Metzler, L.A., 1945, Stability of Multiple Markets: The Hicks Conditions.
Metzler, L.A., 1951, Wealth, Saving and the Rate of Interest.
Morishima, M., 1964, *Equilibrium Stability, and Growth.*
Mundell, R.A., 1960, The Monetary Dynamics of International Adjustment under Fixed and Flexible Exchange Rates.
Murata, Y., 1977, *Mathematics for Stability and Optimization of Economic Systems,* Chap. 4, Sects. 4.2, 4.5.
Negishi, T., 1962, The Stability of a Competitive Economy: A Survey Article.
Newman, P.K., 1959, Some Notes on Stability Conditions.
Patinkin, D., 1965, *Money, Interest, and Prices,* Supplementary Note B.
Petit, M.L., 1990, *Control Theory and Dynamic Games in Economic Policy Analysis,* Chap. 4, Sect. 4.4.
Samuelson, P.A., 1944, The Relation between Hicksian Stability and True Dynamic Stability.
Samuelson, P.A., 1947, *Foundations of Economic Analysis,* Chap. IX, pp. 269-274.
Schwartz, J.T., 1961, *Lectures on the Mathematical Method in Analytical Economics*, p.24.
Solow, R.M., 1959, Competitive Valuation in a Dynamic Input-Output System.
Tarr, D.G., 1978, Expectations and Stability with Gross Complements.
Turnbull, H.W., 1957, *Theory of Equations,* pp. 114-115.
von Neumann, J., 1945 (1937), A Model of General Equilibrium.
Walras, L., 1954 (1874, 1926), *Elements of Pure Economics,* Lesson 12, §§ 124-130.

Part III

ADVANCED TOPICS

Chapter 20

Comparative Statics and the Correspondence Principle

20.1 Introduction

Let there be given a situation of static equilibrium, formally described by the solution of a system of static equations in which, in addition to the variables (whose equilibrium values are given by the solution of the system), various parameters also appear, considered as exogenously given.

Comparative statics studies the displacement of equilibrium, namely it purposes to examine how the equilibrium values of the variables respond to a change in one or more parameters, that is in which direction they change and establish a new equilibrium to match the new configuration of the parameters.

The traditional example of demand-and-supply equilibrium may be useful to clarify such concepts. Let us suppose that in the demand function there is a parameter representing, for example, consumers' tastes: call it α. An increase in α means that consumer tastes have changed in favour of the good in question, i.e. that the demand curve shifts upwards (more is demanded at any given price). The intersection of the demand curve in the new position with the supply curve (which we assume to have remained in the same position) determines the new equilibrium point. We now want to know whether in the new equilibrium the quantity exchanged and the price are greater or smaller than before. That is, how have the variable q^e and the variable p^e reacted to a change in the parameter α? The task of comparative statics is to answer such questions.

It must be stressed that comparative statics does not say anything about the time path of the variables from the initial to the final equilibrium point; *nor can it say whether the new equilibrium point will actually be approached.* However, as we shall see, there exists a strict connection between comparative statics and dynamics. This connection is expressed by the principle that Samuelson called the '*correspondence principle*'.

20.2 The method of comparative statics

The mathematics of comparative statics consists of two well-known theorems: the implicit function theorem and the chain rule for the differentiation of composite functions.

Since the implicit function theorem can be formulated in various ways with different degrees of generality, we give here the version which interests us.

Implicit function theorem. The functions

$$f^i(x_1, x_2, ..., x_n; \alpha_1, \alpha_2, ..., \alpha_m) = 0, \quad i = 1, 2, ..., n,$$

together with their first-order partial derivatives, are continuous with respect to the $n+m$ variables $(x_1, x_2, ..., x_n; \alpha_1, \alpha_2, ..., \alpha_m)$ in a neighbourhood N of the point $(x_1^0, x_2^0, ..., x_n^0; \alpha_1^0, \alpha_2^0, ..., \alpha_m^0)$, which satisfies the relations

$$f^i\left(x_1^0, x_2^0, ..., x_n^0; \alpha_1^0, \alpha_2^0, ..., \alpha_m^0\right) = 0, \quad i = 1, 2, ..., n.$$

Moreover, the value of the Jacobian of the f^i with respect to the x_i is non-zero at the point $(x_1^0, x_2^0, ..., x_n^0; \alpha_1^0, \alpha_2^0, ..., \alpha_m^0)$. If such conditions are satisfied, there exists a set of functions

$$x_i = x_i\left(\alpha_1^0, \alpha_2^0, ..., \alpha_m^0\right), \quad i = 1, 2, ...n,$$

that in a neighbourhood of the point $(x_1^0, x_2^0, ..., x_n^0; \alpha_1^0, \alpha_2^0, ..., \alpha_m^0)$ are single-valued, continuous, are the only functions to satisfy the relations

$$f^i\left(x_1(\alpha_1, \alpha_2, ..., \alpha_m), x_2(\alpha_1, \alpha_2, ..., \alpha_m), ..., x_n(\alpha_1, \alpha_2, ..., \alpha_m); \alpha_1, \alpha_2, ..., \alpha_m\right)$$
$$= 0, \quad i = 1, 2, ..., n,$$

and have continuous partial derivatives of as many orders as are possessed by the f^i in the neighbourhood N.

For the proof of the theorem see e.g. Hobson (1957, Vol. I, §319), or any advanced calculus textbook. It must be noted that the above is the traditional implicit function theorem, which ensures local univalence. Subsequently Gale and Nikaidô (1965; see also Nikaidô, 1968) have proved a global version of this theorem, according to which sufficient conditions for global univalence are that the Jacobian and all its principal minors be everywhere positive.

Let us now come to the economics. We are given n equations in implicit form, whose solution determines the equilibrium point which we were discussing in the previous section. Of course, such equations will be obtained

from economic considerations, which for the moment need not interest us. In these equations n variables and m $\left(m \gtreqless n\right)$ parameters appear. In symbols,

$$f^i(x_1, x_2, ..., x_n; \alpha_1, \alpha_2, ..., \alpha_m) = 0, \quad i = 1, 2, ..., n. \tag{20.1}$$

Given a certain configuration of the parameters, say $(\alpha_1^0, \alpha_2^0, ..., \alpha_m^0)$, the solution of system (20.1) determines the corresponding equilibrium values $(x_1^0, x_2^0, ..., x_n^0)$ of the variables x_i. Since the comparative statics method takes the equilibrium point as given, we assume that this solution exists and is economically meaningful.

Now, if the conditions required by the implicit function theorem are satisfied, we can express the x_i as differentiable functions of the α's in a neighbourhood of the point $(x_1^0, x_2^0, ..., x_n^0; \alpha_1^0, \alpha_2^0, ..., \alpha_m^0)$, i.e.

$$x_i = x_i(\alpha_1, \alpha_2, ..., \alpha_m), \quad i = 1, 2, ..., n. \tag{20.2}$$

The problem of comparative statics would seem therefore solved: substituting the new values of the parameters $\left(\alpha_1', \alpha_2', ..., \alpha_m'\right)$ in (20.2) we can immediately obtain the new equilibrium point $\left(x_1', x_2', ..., x_n'\right)$. Of course we could also substitute the new set $(\alpha_1', \alpha_2', ..., \alpha_m')$ directly in (20.1) and solve them again for the x_i. This, however, would not change the difficulties that we are going to examine.

Unfortunately, the fact is that usually we must be content with the knowledge that Eqs. (20.2) exist, without being able to express them in terms of known functions. This impossibility may be purely mathematic, in the sense that, given the f^i, it is not always possible to solve for the x_i; but it usually derives from the fact that in economic theory the desire to reach general results, not depending on a particular form of the f^i, leads us to *leave this form unspecified and to make only 'qualitative' assumptions, consisting of assumptions on the signs of the partial derivatives of the* f^i *at the most.*

These difficulties, however, can be overcome if we are satisfied with qualitative results as well. Suppose that we succeed in determining the *signs* of the partial derivatives

$$\left(\frac{\partial x_i}{\partial \alpha_j}\right)^0, \quad i = 1, 2, ...n; \quad j = 1, 2, ...m, \tag{20.3}$$

where with the notation $()^0$ we want to specify that the partial derivatives are evaluated at the point $(x_1^0, x_2^0, ..., x_n^0; \alpha_1^0, \alpha_2^0, ..., \alpha_m^0)$. From now on for simplicity we shall drop this notation, it being clear from the context that the various derivatives are evaluated at the said point.

Now, when we know the signs of the partial derivatives (20.3), we have determined the *direction* in which the new equilibrium value of the i-th variable lies as a result of a sufficiently small change in the j-th parameter, that is, *we know whether at new equilibrium point the i-th variable has a value* *which is greater than, smaller than or equal to the value that it had at the initial equilibrium point,* although we cannot say by how much. Such ordinal comparison is the most than we can hope for in 'qualitative' economic theory.

If all parameters change simultaneously, we only have to evaluate, instead of the partial differentials $(\partial x_i/\partial \alpha_j)\mathrm{d}\alpha_j$ (it must be noted that the variations $\mathrm{d}\alpha_j$ are known exogenously), the total differentials

$$\frac{\partial x_i}{\partial \alpha_1}\mathrm{d}\alpha_1 + \frac{\partial x_i}{\partial \alpha_2}\mathrm{d}\alpha_2 + \ldots + \frac{\partial x_i}{\partial \alpha_m}\mathrm{d}\alpha_m,$$

and similarly if any group of k parameters $(k < m)$ vary simultaneously.

Now, the comparative statics method suggests how to compute the partial derivatives (20.3). This method is the one 'codified' by Samuelson. Of course earlier applications of the method can be found in the mathematical economics literature. Without going back to the last century, we can indicate Slutsky's (1915) famous article and, of course, Hicks' (1939) *Value and Capital.* However, Samuelson was the first to give a systematic account of this method and to point out its general nature. It can also be noted, from the purely mathematical point of view, that the same result—i.e. Eqs. (20.6) below—could be reached by linearising the f^i at the equilibrium point, solving explicitly for the x_i in terms of the α_j and computing the partial derivatives of the resulting linear expressions.

Since (20.2) satisfies (20.1), that is

$$\begin{aligned} &f^i(x_1(\alpha_1, \alpha_2, ..., \alpha_m), x_2(\alpha_1, \alpha_2, ..., \alpha_m), \\ &\quad ..., x_n(\alpha_1, \alpha_2, ..., \alpha_m); \alpha_1, \alpha_2, ..., \alpha_m) = 0, \quad i = 1, 2, ..., n, \end{aligned}$$

we can compute, by means of the chain rule theorem, the total derivative of each f^i with respect to any parameter α_j obtaining

$$\sum_{s=1}^{n} \frac{\partial f^i}{\partial x_s}\frac{\partial x_s}{\partial \alpha_j} + \frac{\partial f^i}{\partial \alpha_j} = 0, \qquad i = 1, 2, ..., n. \tag{20.4}$$

The left-hand side of (20.4) is the total derivative of the i-th function with respect to the j-th parameter; this derivative is evaluated at the point

$$\left(x_1^0, x_2^0, ..., x_n^0; \alpha_1^0, \alpha_2^0, ..., \alpha_m^0\right).$$

Since each f^i is identically zero in a neighbourhood N of this point and so is stationary there, the above total derivative must be zero, from which (20.4) follow.

By shifting $\partial f^i/\partial \alpha_j$ to the right-hand side and writing the equations in

extended form we get

$$\begin{aligned}
\frac{\partial f^1}{\partial x_1}\frac{\partial x_1}{\partial \alpha_j} + \frac{\partial f^1}{\partial x_2}\frac{\partial x_2}{\partial \alpha_j} + \ldots + \frac{\partial f^1}{\partial x_n}\frac{\partial x_n}{\partial \alpha_j} &= -\frac{\partial f^1}{\partial \alpha_j}, \\
\frac{\partial f^2}{\partial x_1}\frac{\partial x_1}{\partial \alpha_j} + \frac{\partial f^2}{\partial x_2}\frac{\partial x_2}{\partial \alpha_j} + \ldots + \frac{\partial f^2}{\partial x_n}\frac{\partial x_n}{\partial \alpha_j} &= -\frac{\partial f^2}{\partial \alpha_j}, \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & \quad \ldots\ldots \\
\frac{\partial f^n}{\partial x_1}\frac{\partial x_1}{\partial \alpha_j} + \frac{\partial f^n}{\partial x_2}\frac{\partial x_2}{\partial \alpha_j} + \ldots + \frac{\partial f^n}{\partial x_n}\frac{\partial x_n}{\partial \alpha_j} &= -\frac{\partial f^n}{\partial \alpha_j}.
\end{aligned} \tag{20.5}$$

Eqs. (20.5) are a linear system of n unknowns $\partial x_i/\partial \alpha_j$, $i = 1, 2, ..., n$. Solving such a system we have

$$\frac{\partial x_i}{\partial \alpha_j} = \frac{\Delta_i}{\Delta}, \tag{20.6}$$

where

$$\Delta \equiv \begin{vmatrix} \frac{\partial f^1}{\partial x_1} & \frac{\partial f^1}{\partial x_2} & \cdots & \frac{\partial f^1}{\partial x_n} \\ \frac{\partial f^2}{\partial x_1} & \frac{\partial f^2}{\partial x_2} & \cdots & \frac{\partial f^2}{\partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial f^n}{\partial x_1} & \frac{\partial f^n}{\partial x_2} & \cdots & \frac{\partial f^n}{\partial x_n} \end{vmatrix}$$

is the determinant of the system and Δ_i is the determinant we obtain by substituting in Δ the column of the known terms $-\partial f^k/\partial \alpha_j$ to the column of the coefficients concerning the i-th unknown, which are $\partial f^k/\partial x_i$, $k = 1, 2, ..., n$. We can immediately note that Δ is the Jacobian of the functions f^i with respect to the x_i, which we have assumed to be non-vanishing for the functions (20.2) to exist; therefore system (20.5) is always solvable.

We must now cope with the problem of determining the sign of both Δ and Δ_i. We have said above that we know solely the signs of the derivatives $\partial f^k/\partial \alpha_j$ and $\partial f^k/\partial x_i$, and it is obvious that such knowledge is not, except for particular cases, sufficient to determine the sign of Δ and Δ_i.

However, this difficulty can be (partially) overcome. It is well known that two powerful tools in the economic theorist's tool-kit are *the assumption of maximizing behaviour* and *the assumption that the equilibrium point is dynamically stable.* By means of such assumptions it is often possible to determine the sign of Δ; sometimes they also help in determining the sign of Δ_i.

Although a book on dynamics, such as this is, need only present the relation between comparative statics and the stability assumption (which is the correspondence principle), the relation between comparative statics and the maximising assumption is also given, to complete the description of the comparative statics method.

20.2.1 Purely qualitatively comparative statics

We have said above that in some cases the signs of the relevant determinants can be ascertained starting from the mere knowledge of the signs of the partial derivatives which appear in them. For example, if a determinant has the following sign pattern

$$\begin{vmatrix} + & + & + & + & + & + \\ - & + & + & + & + & + \\ 0 & - & + & + & + & + \\ 0 & 0 & - & + & + & + \\ 0 & 0 & 0 & - & + & + \\ 0 & 0 & 0 & 0 & - & + \end{vmatrix},$$

then it is clearly positive.

Research has been carried out to find those special cases where the knowledge of the signs of the elements of a determinant is sufficient to determine the sign of the determinant and an algorithm has been devised for this determination (see Lancaster, 1966). This algorithm is a method of obtaining in a more efficient way the *same* information (i.e., whether the sign of the determinant is determinate or not) which could be obtained more laboriously by explicit expansion of the determinant. However, Basset (1968) has provided a counter-example, in which the explicit expansion of the determinant allows the determination of its sign, whereas Lancaster's algorithm leaves it indeterminate. This is due to the fact that in economic problems we usually know not only the sign but also the form of the element—for example, that it is a marginal propensity to spend, etc.—so that in the expansion some terms may cancel out.

The Lancaster algorithm has been improved by Ritschard (1983), who presents an algorithm that much increases the efficiency of the Samuelson-Lancaster elimination principle.

These algorithms cannot, of course, make sign-determinate a determinant which is not so; therefore, they cannot substitute the additional information that we obtain from the maximizing assumption and from the stability assumption. Farley and Lin (1990) present new techniques that were originally discussed in the context of qualitative physics and adapt them to qualitative reasoning in economics.

20.2.2 The inverse comparative statics problem

We may also be interested in the *inverse* comparative statics problem (or backward reasoning), which goes from the endogenous variables to the parameters. This is the well known targets-instruments context in economic policy analysis, where we are interested in obtaining a certain change in the endogenous variables (e.g., an increase in the quantity produced q^e) by changing the parameter(s) under the control of the policy authorities (e.g. government demand α).

In other words, we are interested in the functions

$$\alpha_j = \alpha_j(x_1, x_2, ..., x_n), \quad j = 1, 2, ..., m, \tag{20.7}$$

rather than in the functions (20.2). This requires the appropriate Jacobian of the functions (20.1) with respect to the α_j to be non-vanishing. We say 'appropriate' because, as the Jacobian is by definition a square determinant, we must be careful in checking the condition. Let us assume that the n targets are independent from one another and that the m instruments are also independent from one another, so that no problem of rank arises.

Now, if $m = n$, i.e. if the number of instruments and targets are equal (this is the well-known Tinbergen case), the Jacobian is unique. If $m < n$ (number of instruments smaller than the number of targets) we cannot solve the problem unless we give up some targets, namely a subset of $k = n - m$ targets. Finally, if the number of instruments is greater than the number of targets ($m > n$) we have a redundancy of instruments, and the problem has multiple solutions, since we can use any subset of $l = m - n$ instruments to write the Jacobian. This illustrates the well-known Tinbergen principle, according to which—in a static context—in order to reach a certain number of independent targets, the policy maker must have at least as many independent instruments as targets (Petit, 1990, Chap. 3, Sect. 3.2).

Once the appropriate Jacobian is non vanishing, we can proceed exactly as in the 'direct' comparative statics problem to find the relevant derivatives.

20.3 Comparative statics and optimizing behaviour: an example from traditional demand theory.

By optimizing behaviour we mean both maximizing and minimizing behaviour. The reader is assumed to be familiar with the first and second-order conditions for a (free or constrained) maximum or minimum (in case of need, consult Appendix A of Samuelson, 1947, or any text of mathematics for economists).

The optimizing behaviour assumption is one of the basic principles of neoclassical economic theory, in which the principle of rational behaviour is equivalent to the assumption that economic agents maximize or minimize something. The best-known basic examples are the theory of consumer's behaviour and the theory of cost and production.

From the formal point of view, let us assume that Eqs. (20.1) of the previous section are the first-order conditions for an extremum (maximum or minimum, free or constrained). In other words, the f^i are the n first-order partial derivatives of a function F to be maximized or minimized (in the F, and so in the f^i also, one or more Lagrange multipliers will appear if the extremum is constrained). For the point determined by the solution of (20.1) to be a solution of the extremum problem and so to establish an equilibrium point, the second-order conditions[1] must, of course, be satisfied.

Now, the satisfaction of the second-order conditions implies, among other things, that the *Hessian* of the F is non-zero and has a precise sign depending on the number of variables and on the type of extremum. By definition, the Hessian of a function coincides with the Jacobian of the first-order partial derivatives of the function, so that we can immediately draw two important conclusions.

The first is that, since the Jacobian is different from zero, the essential condition required by the implicit function theorem is satisfied, so that the functions (20.2) of the previous section exist.

The second is that, when the sign of the said Jacobian is determined, the sign of Δ is automatically determined, since, as we have seen in the previous section, Δ and the Jacobian are the same. From this it follows the importance, to which Samuelson has called attention, of the second-order conditions for an extremum, not only to establish whether the stationary point is a maximum or minimum, but also to have useful information for comparative statics purposes.

As a simple illustration we shall examine the consumer's choice problem. Let $U = U(x_1, x_2)$ be an (ordinal) utility function that the consumer is assumed to maximize subject to the budget constraint $R - p_1x_1 - p_2x_2 = 0$. Let us form the function

$$F = U(x_1, x_2) + \lambda (R - p_1x_1 - p_2x_2),$$

where λ is a Lagrange multiplier. The first-order conditions for a maximum are

$$\begin{aligned}
\frac{\partial F}{\partial x_1} &= \frac{\partial U}{\partial x_1} - \lambda p_1 = 0,\\
\frac{\partial F}{\partial x_2} &= \frac{\partial U}{\partial x_2} - \lambda p_2 = 0,\\
\frac{\partial F}{\partial \lambda} &= R - p_1x_1 - p_2x_2 = 0,
\end{aligned}$$

[1]In the older literature the second-order conditions for an extremum are often called 'stability conditions' (see e.g. Hicks, 1939, mathematical appendixes). They must not be confused with dynamic stability conditions, although they are not complețly unrelated, as we shall see in Sect. 20.5 below.

and the second-order conditions are that the following bordered[2] Hessian

$$\begin{vmatrix}
\dfrac{\partial^2 U}{\partial x_1^2} & \dfrac{\partial^2 U}{\partial x_1 x_2} & -p_1\\[2ex]
\dfrac{\partial^2 U}{\partial x_1 x_2} & \dfrac{\partial^2 U}{\partial x_2^2} & -p_2\\[2ex]
-p_1 & -p_2 & 0
\end{vmatrix}$$

should be positive. The first-order conditions are a system of three equations in implicit form in the three variables x_1, x_2, λ and in the three parameters p_1, p_2, R, whose solution yields the point of the consumer's equilibrium, provided that the second-order conditions are satisfied. It can easily be checked that the Jacobian of the first-order equations with respect to x_1, x_2, λ is the same as the above-written Hessian. Thus the following functions exist:

$$\begin{aligned}
x_1 &= x_1(p_1, p_2, R),\\
x_2 &= x_2(p_1, p_2, R),\\
\lambda &= \lambda(p_1, p_2, R),
\end{aligned}$$

where the first two functions are the consumer's demand functions. We now want to examine, for example, how the quantity demanded of x_1 responds to a change in p_1. Differentiating totally the first-order conditions with respect to p_1 we have

$$\begin{aligned}
\frac{\partial^2 U}{\partial x_1^2}\frac{\partial x_1}{\partial p_1} + \frac{\partial^2 U}{\partial x_1 \partial x_2}\frac{\partial x_2}{\partial p_1} - p_1\frac{\partial \lambda}{\partial p_1} &= \lambda,\\
\frac{\partial^2 U}{\partial x_1 \partial x_2}\frac{\partial x_1}{\partial p_1} + \frac{\partial^2 U}{\partial x_2^2}\frac{\partial x_2}{\partial p_1} - p_2\frac{\partial \lambda}{\partial p_1} &= 0,\\
-p_1\frac{\partial x_1}{\partial p_1} - p_2\frac{\partial x_2}{\partial p_2} &= x_1,
\end{aligned}$$

so that

$$\frac{\partial x_1}{\partial p_1} = x_1 \begin{vmatrix}
\dfrac{\partial^2 U}{\partial x_1 \partial x_2} & -p_1\\[2ex]
\dfrac{\partial^2 U}{\partial x_2^2} & -p_2
\end{vmatrix} \Delta^{-1} + \lambda\frac{-p_2^2}{\Delta} = -x_1\frac{\partial x_1}{\partial R} + \lambda\frac{-p_2^2}{\Delta},$$

[2]Note that the Hessian is bordered if we start from the Hessian of the utility function, but is a normal Hessian if we consider it as the Hessian of the F function.

where[3] Δ is the determinant of the system, which coincides with the above-written Hessian. Thus we know that $\Delta > 0$. Therefore, if we also take into account the fact that $\lambda > 0$, it follows that the second term in the last expression is negative (substitution effect) whereas the first term (income effect) remains with an uncertain sign.

The above succinct treatment seems sufficient to illustrate the relation between comparative statics and the maximizing behaviour assumption, so we pass on to the treatment of the correspondence principle.

20.4 Comparative statics and dynamic stability of equilibrium: the 'correspondence principle'

The study of the dynamic stability of equilibrium is very important on its own. In fact, an equilibrium point which exists in principle, but which cannot be approached and which (supposing that it has been hit by chance) is such that the slightest disturbance starts a movement away from it—that is, an unstable equilibrium point—is obviously not very relevant from an economic point of view (for a discussion of these problems see Chap. 21). Now, it turns out that the study of the dynamic stability of equilibrium may also be important for obtaining determinate comparative statics results, through Samuelson's correspondence principle between statics and dynamics (Samuelson, 1941b, 1947).

Let the equilibrium point be determined by relations (20.2) of Sect. 20.2. In the study of the dynamic stability of this point it often happens that plausible behaviour assumptions lead to dynamic equations or systems of the type

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = k_i f^i\left(x_1, x_2, ..., x_n\right), \tag{20.8}$$

where the k_i are positive constants. The substance of the following treatment would not be significantly changed if we considered difference equations instead of differential equations. It can also be observed that the right-hand side of Eqs. (20.8) can be considered as originating from linearisation at the equilibrium point of the functions $h_i\left[f^i(...)\right]$, where h_i are sign-preserving

[3]In the last equality the following relation is implicit

$$\frac{\partial x_1}{\partial R} = -\begin{vmatrix} \frac{\partial^2 U}{\partial x_1 \partial x_2} & -p_1 \\ \frac{\partial^2 U}{\partial x_2^2} & -p_2 \end{vmatrix} \Delta^{-1}.$$

In fact, differentiating with respect to R the first-order conditions and solving for $\partial x_1/\partial R$, we obtain the above relation.

functions and $h_i'[0] \equiv k_i$. Equations like (20.8) are frequently used in dynamical mechanics, where the laws of motion often make the movement of a point depend on the current position of the point itself. In economics it is enough to recall, for example, the study of the stability of Walrasian general equilibrium of exchange (see Part II, Chap. 19, Sect. 19.1), where the equilibrium point is determined by the set of equations $E_i(p_1, p_2, ...p_m) = 0$ and the *tâtonnement* process can be formalized as $\mathrm{d}p_i/\mathrm{d}t = k_i E_i$.

Not knowing the form of the functions f^i, we can make a linear approximation at the equilibrium point. This implies that we are studying local stability. What if we study global stability? No linearisation is required applying Liapunov's second method (see Chap. 23), so that it would seem that we cannot obtain any information from global stability considerations. Of course, this is not true, since if global stability obtains, local stability obtains *a fortiori*.

Considering the deviations from equilibrium, $\bar{x}_i = x_i - x_i^0$, we obtain

$$\frac{\mathrm{d}\,\bar{x}_i}{\mathrm{d}t} = k_i \sum_{j=1}^{n} \frac{\partial f^i}{\partial x_j} \bar{x}_j, \quad i = 1, 2, ..., n, \tag{20.9}$$

where the $\partial f^i/\partial x_j$ are evaluated at the equilibrium point. The characteristic equation of system (20.9) is

$$\begin{vmatrix} k_1\frac{\partial f^1}{\partial x_1} - \lambda & k_1\frac{\partial f^1}{\partial x_2} & \dots & k_1\frac{\partial f^1}{\partial x_n} \\ k_2\frac{\partial f^2}{\partial x_1} & k_2\frac{\partial f^2}{\partial x_2} - \lambda & \dots & k_2\frac{\partial f^2}{\partial x_n} \\ \dots & \dots & \dots & \dots \\ k_n\frac{\partial f^n}{\partial x_1} & k_n\frac{\partial f^n}{\partial x_2} & \dots & k_n\frac{\partial f^n}{\partial x_n} - \lambda \end{vmatrix} = 0.$$

Expanding the determinant we obtain an equation of the type

$$(-1)^n \lambda^n + (-1)^{n-1} c_1 \lambda^{n-1} + ... + (-1)^{n-r} c_r \lambda^{n-r} + ... + c_n = 0, \tag{20.10}$$

where the coefficients c_1 are expressed in terms of the elements of the matrix

$$\begin{bmatrix} k_1\frac{\partial f^1}{\partial x_1} & k_1\frac{\partial f^1}{\partial x_2} & \dots & k_1\frac{\partial f^1}{\partial x_n} \\ k_2\frac{\partial f^2}{\partial x_1} & k_2\frac{\partial f^2}{\partial x_2} & \dots & k_2\frac{\partial f^2}{\partial x_n} \\ \dots & \dots & \dots & \dots \\ k_n\frac{\partial f^n}{\partial x_1} & k_n\frac{\partial f^n}{\partial x_2} & \dots & k_n\frac{\partial f^n}{\partial x_n} \end{bmatrix},$$

and are defined as follows:

c_1 = sum of all principal minors of the first order
c_2 = sum of all principal minors of the second order
....
c_r = sum of all $\frac{n!}{r!(n-r)!}$ principal minors of the r-th order
....
c_n = determinant of the matrix.
Let us note that, from the properties of determinants,

$$c_n = k_1k_2...k_n \begin{vmatrix} \frac{\partial f^1}{\partial x_1} & \frac{\partial f^1}{\partial x_2} & \cdots & \frac{\partial f^1}{\partial x_n} \\ \frac{\partial f^2}{\partial x_1} & \frac{\partial f^2}{\partial x_2} & \cdots & \frac{\partial f^2}{\partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial f^n}{\partial x_1} & \frac{\partial f^n}{\partial x_2} & \cdots & \frac{\partial f^n}{\partial x_n} \end{vmatrix},$$

that is

$$c_n = k_1k_2...k_n\Delta,$$

where Δ is the same determinant which appears in the comparative statics results (see above, Sect. 20.2).

Let us now examine the stability conditions. We know (see Chap. 18, Sect. 18.2.2) that a necessary (although not sufficient) stability condition, is that $c_n \neq 0$ (this excludes any zero real root) and that all the coefficients $(-1)^n, (-1)^{n-1} c_1, ..., c_n$ have the same sign (this excludes, by Descartes' theorem, any positive root). Therefore c_n, and consequently Δ (since the k_i are positive constants), must have the sign of $(-1)^n$. Thus, by means of stability considerations, we are able to determine the sign of the determinant which appears in the denominators of the comparative statics results. Moreover, the fact that also the signs of the other coefficients $(c_1, c_2..., c_{n-1})$ are determined, may be useful to determine the sign of the Δ_i.

The determination of the sign of Δ (and where possible also of Δ_i) by means of dynamic stability considerations is the essence of the *correspondence principle.*

20.4.1 Criticism and qualifications

Of course, this principle is not without limitations, some of which were already pointed out by Patinkin (1952). Cases may occur in which the perfect 'correspondence' between comparative statics and the dynamic system—perfect in the sense that the comparative statics Δ is the same, apart from the positive multiplicative constants, as the constant term in the characteristic equation of the dynamic system—is not possible. Such cases are all those in which the behaviour assumptions made in the study of the dynamic stability do not lead to equations of type (20.8).

Patinkin illustrates the case in which the dynamic equations for the study of the stability of equilibrium are of the type

$$\frac{dx_i}{dt} = \sum_{j=1}^{n} k_{ij} f^j (x_1, x_2, ..., x_n; \alpha_1, \alpha_2, ..., \alpha_m), \ i = 1, 2, ..., n.$$

Yet in this case too it is possible to obtain some information about comparative statics (note 2, p. 42, in Patinkin's (1952) article).

However, the strongest critique to the correspondence principle comes from general equilibrium theory, where it was already shown by Arrow and Hahn (1971, Chap. 12, Sect. 11) that there is little hope of obtaining useful restrictions on $\partial p_i/\partial \alpha_j$. But the most destructive criticism in the field of general equilibrium theory is a consequence of the Sonnenschein (1972) theorem (see also Mantel, 1974, and Debreu, 1974), according to which the axioms of Arrow-Debreu-McKenzie general equilibrium theory are sufficiently general to allow *any* continuous function $E(\mathbf{p})$ to be an excess demand function for an economy populated by agents with perfectly well-behaved utility functions. Hence, starting from the general-equilibrium system of equations $\mathbf{E}(\mathbf{p}, \boldsymbol{\alpha}) = \mathbf{0}$, arbitrary $\partial \mathbf{p}/\partial \boldsymbol{\alpha}$ can be obtained although $\partial \mathbf{E}/\partial \mathbf{p}$ is a stable matrix and $\partial \mathbf{E}/\partial \boldsymbol{\alpha}$ has a priori sign restrictions. Of course, if one is willing somehow to restrict the Jacobian of the excess demand functions, attenuated forms of the Principle can be found (Kemp et al., 1990). But the fact remains that the Principle is generally invalid.

This is indeed the epitaph of the Correspondence Principle *in the field of general equilibrium theory.* Fortunately for the correspondence principle (and for economists of a less abstract bent), there does not only exist abstract general equilibrium theory. In other fields of economic theory (as, for example, in optimal aggregate-growth theory: see Brock, 1986) that criticism does not apply, hence it seems fair to say that the correspondence principle, while not being a panacea, can be a useful tool in many instances. In fact, it often turns out that equations of type (20.8) are indeed appropriate for the study of the dynamic stability of the equilibrium point, and with this study it may be possible to obtain some information about comparative statics from dynamic considerations.

Let us note that by applying the correspondence principle we kill two birds with one stone, since in addition to obtaining information about comparative statics, we check at the same time that the system will actually approach the new equilibrium point, a result that cannot be derived from comparative statics alone.

As we have shown, the sign of Δ has been determined *assuming* that the equilibrium is stable. It must be noted that when different behaviour assumptions are possible (for some general considerations on the 'relativity' of stability conditions, see Chap. 13, Sect. 13.1), different stability conditions (and so different comparative statics results) may be obtained.

Apart from this, the objection could be raised that to *impose* on the model from outside that the equilibrium point should be stable is not legitimate, or at least not logically satisfactory. A possible answer to this objection is that, if we take the view that unstable equilibria are not meaningful from the economic point of view (more on this in Chap. 21), then we are interested in analysing the comparative statics properties of stable equilibria only. It would be completely useless to know where the new equilibrium point lies if this point cannot be approached. A methodologically sounder way out is to drop the stability *assumption*, and to make conditional (or taxonomic) statements of the type '*if* the equilibrium point is stable (unstable), then the comparative statics results are so and so'.

This, however, does not solve a problem that comes from the very presence of parameters in the equations. It is sufficient that a system of differential equations contains just one parameter for the *possibility* to arise that changes in the value of the parameter give rise to *bifurcations*, namely to points where the qualitative behaviour of the system changes dramatically (on bifurcation theory see Chap. 25). For example, at $\alpha = \alpha_0$ the system is stable, but becomes unstable at $\alpha_1 = \alpha_0 + \mathrm{d}\alpha_0$.

Some types of bifurcations can be excluded, which are all those that require the Jacobian to be zero at the bifurcation point—the Jacobian, in fact, has to be non-zero for the very possibility of carrying out a comparative statics exercise. For continuity reasons a negative real characteristic root cannot become positive without passing through zero, which implies the Jacobian becoming zero, and this is excluded. However, the negative real part of a pair of complex conjugate roots may become positive passing through zero without causing the Jacobian to become zero, which is exactly the case of the Hopf bifurcation (see Chap. 25, Sect. 25.2.2) from which a limit cycle arises. It follows that if one performs a comparative statics exercise *cum* correspondence principle one must be very careful if the system is not structurally stable (on the notion of structural stability see Chap. 21, Sect 21.2.3).

It must be noted, finally, that when different behaviour assumptions are possible, different stability conditions may hold (this is the principle of relativity of stability conditions, see Chap. 13, Sect. 13.1), and so different comparative statics results may be obtained.

20.5 Extrema and dynamic stability

The 'static' considerations concerning the second-order conditions for an extremum and the dynamic considerations concerning stability can sometimes be put together. This is another aspect of the relations between statics and dynamics, which can be considered as a special case of a generalized correspondence principle.

Formally, if the equations

$$f^i(x_1, x_2, ..., x_n; \alpha_1, \alpha_2, ..., \alpha_m) = 0, \qquad i = 1, 2, ..., n,$$

determine a stationary point (maximum, minimum, or a point not an extremum) of a function F, of which the f^i are the first-order partial derivatives, and if the behaviour assumptions concerning the dynamic stability of the same point give rise to a dynamic system of the type

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = k_i f^i(x_1, x_2, ..., x_n; \alpha_1, \alpha_2, ..., \alpha_m), \quad i = 1, 2, ..., n, \quad k_i > 0,$$

then:

(I) if the stationary point is a maximum, it is locally stable;

(II) if the stationary point is a minimum, it is unstable;

(III) if the stationary point is not an extremum, it is stable or unstable according to the initial position of the system (one-sided stability-instability in Samuelson's (1947) terminology or, more synthetically, semi-stability).

The above theorem can easily be proved if we consider the simple case in which there is only one function in one independent variable. Let $y = F(x)$ be such a function, and let x^0 be the stationary point determined by the equation

$$f(x) = F'(x) = 0.$$

From the theory of maxima and minima, we know that, if $F'(x^0) = F''(x^0) = ... = F^{(n-1)}(x^0) = 0$, with $F^{(n)}(x^0) \neq 0$, then:

(1) if n is even, the stationary point is an extremum, and precisely a maximum (minimum) if $F^{(n)}(x^0) \lessgtr 0$;

(2) if n is odd, the stationary point is not an extremum but a horizontal inflection point, in the neighbourhood of which the function is increasing (decreasing) if $F^{(n)}(x^0) \gtrless 0$.

Let us begin with the special cases in which $n = 2$ or 3. Let us consider the differential equation

$$\frac{\mathrm{d}x}{\mathrm{d}t} = kf(x), \quad k > 0,$$

and expand the right-hand side in Taylor's series at the point x^0. Omitting all non-linear terms we have

$$\frac{\mathrm{d}\overline{x}}{\mathrm{d}t} = kf'\left(x^0\right)\overline{x},$$

where $\overline{x} = x - x^0$ denotes deviations from equilibrium. The solution of the last differential equation is

$$\overline{x}(t) = \overline{x}(0)\exp\left[kf'\left(x^0\right)t\right],$$

where $\overline{x}(0)$ indicates the initial deviation $x(0) - x^0$ for $t = 0$. Since $f'(x^0) \equiv F''(x^0)$, we have

$$\overline{x}(t) = \overline{x}(0) \exp\left[kF''\left(x^0\right)t\right],$$

from which it can be immediately seen that the point x^0 is stable or unstable if $F''(x^0) \lessgtr 0$; that is if x^0 affords a maximum (minimum) to the function $F(x)$.

Let us now suppose that $f'(x^0) = 0$, but that $f''(x^0) \equiv F'''(x^0) \neq 0$; in Taylor's expansion of $f(x)$ the first-order term is zero, so that we must consider the second-order term (higher-order terms are neglected). Thus we have

$$\frac{d\overline{x}}{dt} = kf''\left(x^0\right)\frac{1}{2}\overline{x}^2.$$

This can be considered as a Bernoulli equation (see Chap. 24, Sect. 24.2.3, for a general treatment of these differential equations). In the case under consideration, let $z = 1/\overline{x}$, whence $dz/dt = -1/\overline{x}^2$, and, of course $\overline{x} = 1/z$, obtaining

$$-\overline{x}^{-2}\left(\frac{d\overline{x}}{dt}\right) = -\frac{1}{2}kf''\left(x^0\right),$$

whence

$$\frac{dz}{dt} = -\frac{1}{2}kf''\left(x^0\right),$$

and so, integrating,

$$z(t) = -\frac{1}{2}kf''\left(x^0\right)t + z(0),$$

where, given the definition of z, the arbitrary constant $z(0)$ is equal to $1/\overline{x}(0)$. Going back from z to $\bar{x}$, one obtains

$$\overline{x}(t) = \frac{1}{z(t)} = \frac{1}{-\frac{1}{2}kf''(x^0)t + \frac{1}{\overline{x}(0)}} = \frac{1}{-\frac{1}{2}kF'''(x^0)t + \frac{1}{\overline{x}(0)}}.$$

It can easily be checked that, if $\overline{x}(0)$ has the opposite sign of $-F'''(x^0)$, there will be a positive value of t for which the denominator will vanish, that is $\overline{x}$ becomes infinite and the equilibrium is not stable. On the other hand, if $\overline{x}(0)$ has the same sign as $-F'''(x^0)$, then the denominator has always the same sign for $t \geq 0$ and tends in absolute value to infinity as $t \to +\infty$, so that $\overline{x}(t) \to 0$ and the equilibrium is stable. Therefore, we have stability or instability according to the position of the initial point $x(0)$ which determines the initial deviation $\overline{x}(0)$.

What we have shown for $n = 2, 3$, can be generalized. In what follows we put $m \equiv n - 1$ and we assume that $m > 2$ (and so that $n > 3$) since the cases up to $n = 3$ have just been examined. Let us then suppose that the

first m derivatives of F, that is the first $m - 1$ derivatives of f are zero at $x = x^0$, whereas $F^{(n)}(x^0) \equiv f^{(m)}(x^0) \neq 0$. The differential equation is

$$\frac{d\overline{x}}{dt} = \frac{k}{m!}f^{(m)}x^0\,\bar{x}^m,$$

which again is a Bernoulli equation. In the case under consideration let $z = \overline{x}^{(1-m)}$, whence $dz/dt = (1-m)\overline{x}^{-m}\,(d\overline{x}/dt)$, and, of course, $\overline{x} = z^{\frac{1}{(1-m)}} = 1/z^{\frac{1}{(m-1)}}$. Multiply both members of the equation by $(1-m)\overline{x}^{-m}$, obtaining

$$(1-m)\overline{x}^{-m}\frac{d\overline{x}}{dt} = \frac{k}{m!}(1-m)f^{(m)}\left(x^0\right),$$

whence

$$\frac{dz}{dt} = \frac{k}{m!}(1-m)f^{(m)}\left(x^0\right),$$

and so, integrating,

$$z(t) = \frac{k}{m!}(1-m)f^{(m)}\left(x^0\right)t + z(0),$$

where, given the definition of z, the arbitrary constant $z(0)$ is equal to $[\overline{x}(0)]^{1-m} = 1/[\overline{x}(0)]^{m-1}$. Going back from z to $\bar{x}$, one obtains

$$\begin{aligned}\overline{x}(t) &= \frac{1}{z(t)^{\frac{1}{(m-1)}}} = \frac{1}{\left\{\frac{k}{m!}(1-m)f^{(m)}(x^0)t + \frac{1}{[\bar{x}(0)]^{m-1}}\right\}^{\frac{1}{(m-1)}}} \\ &= \frac{1}{\left\{\frac{k}{m!}(1-m)F^{(n)}(x^0)t+\right\}^{\frac{1}{(n-2)}}}.\end{aligned}$$

We must now distinguish two cases:

(1) n is even, so that also $n-2$ is even and so $\frac{1}{[\bar{x}(0)]^{n-2}} > 0$. Since $1-m < 0$ we have the following results:

(a) if $F^n(x^0) < 0$, that is if the stationary point is a maximum, the denominator is positive for any $t \geq 0$ and tends to infinity as $t \to +\infty$, so that $\overline{x}(t) \to 0$ and the equilibrium is stable;

(b) if $F^n(x^0) > 0$ that is, if the stationary point is a minimum, the denominator becomes zero for a positive value of t, that is, $\bar{x}$ becomes infinite and the equilibrium is not stable.

(2) n is odd, so that also $n - 2$ is odd and $\left(1/[\bar{x}(0)]^{n-2}\right) \gtrless 0$ if $\overline{x} \gtrless 0$. It can easily be checked that the denominator always maintains the same sign (for any $t \geq 0$) and tends in absolute value to infinity as $t \to +\infty$, or vanishes for a certain (positive) value of t according as $\overline{x}(0)$ has the same or the opposite sign of $(1-m)F^{(n)}(x^0)$. Thus we have stability or instability according to the position of the initial point.

Thus we have proved proposition (I), (II) and (III) for the case in which only one independent variable is involved. We can also say that, even when the relation $f(x) = 0$ has been derived from considerations other than the maximizing behaviour assumption, if the point x^0 is stable, unstable, stable-unstable, then it must necessarily correspond respectively to a maximum, minimum, horizontal inflection point, of a function $F(x) \equiv \int f(x)dx$. However, this is a purely formal result, since the *a posteriori* construction of the $F(x)$ generally has no economic meaning.

Propositions (I), (II) and (III) can be extended to the general case of a function in n independent variables. As an illustration, we shall treat the case in which the stationary point is a maximum. If $y = F(x_1, x_2, ..., x_n)$ is the function to maximize, the first-order conditions are

$$\frac{\partial F}{\partial x_i} = f^i(x_1, x_2, ..., x_n) = 0, \quad i = 1, 2, ..., n,$$

and the second-order conditions are that the principal minors of the following Hessian matrix

$$\begin{bmatrix} \frac{\partial^2 F}{\partial x_1^2} & \frac{\partial^2 F}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 F}{\partial x_1 \partial x_n} \\ \frac{\partial^2 F}{\partial x_2 \partial x_1} & \frac{\partial^2 F}{\partial x_2^2} & \cdots & \frac{\partial^2 F}{\partial x_2 \partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial^2 F}{\partial x_n \partial x_1} & \frac{\partial^2 F}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 F}{\partial x_n^2} \end{bmatrix} \equiv \begin{bmatrix} \frac{\partial f^1}{\partial x_1} & \frac{\partial f^1}{\partial x_2} & \cdots & \frac{\partial f^1}{\partial x_n} \\ \frac{\partial f^2}{\partial x_1} & \frac{\partial f^2}{\partial x_2} & \cdots & \frac{\partial f^2}{\partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial f^n}{\partial x_1} & \frac{\partial f^n}{\partial x_2} & \cdots & \frac{\partial f^n}{\partial x_n} \end{bmatrix} \qquad (20.11)$$

alternate in sign, beginning with minus.

Let us now examine the dynamic system

$$\frac{dx_i}{dt} = f^i(x_1, x_2, ..., x_n), \quad i = 1, 2, ..., n, \qquad (20.12)$$

where for the moment we have put $k_i = 1$. Performing a linear approximation at the equilibrium point and considering the deviations $\bar{x}_i = \bar{x}_i - x_i^0$, we have

$$\frac{d\bar{x}_i}{dt} = \frac{\partial f^1}{\partial x_1}\bar{x}_1 + \frac{\partial f^1}{\partial x_2}\bar{x}_2 + ... + \frac{\partial f^1}{\partial x_n}\bar{x}_n, \quad i = 1, 2, ..., n. \qquad (20.13)$$

The matrix of the coefficients of system (20.13) coincides with the Hessian (20.11) and is symmetrical, given the commutative property of partial differentiation ($\partial^2 F/\partial x_i \partial x_j = \partial^2 F/\partial x_j \partial x_i$). Now, when the matrix of the coefficients of a differential system is symmetric, the alternation (starting with minus) of the signs of the principal minors is a necessary and sufficient stability condition (see Part II, Chap. 18, Sect. 18.2.2, Condition I). It follows that

the maximum point is stable, of course if the dynamic behaviour assumptions give rise to the differential system (20.12).

We finally recall that if a symmetric matrix $\mathbf{A}$ is stable, also the matrix $\mathbf{DA}$ (where $\mathbf{D}$ is the diagonal matrix $[k_1, k_2, ..., k_n]$; $k_i > 0$) is stable (see Sect. 18.2.2.1). Thus the above argument is also valid if the k_i are not unity.

20.5.1 An application to the theory of the firm

We conclude our treatment with a simple economic application to the elementary theory of the monopolistic firm. The 'traditional' monopolist wants to maximize profit, so that, calling q the quantity produced and sold, $R(q)$ total revenue, and $C(q)$ total cost, the first- and second-order conditions for maximum profit are

$$\begin{aligned} R'(q) - C'(q) &= 0, \\ R''(q) - C''(q) &< 0. \end{aligned}$$

Let us now make the following behaviour assumption: the monopolist increases (decreases) the quantity produced and sold if marginal revenue is greater (smaller) than marginal cost. This look like a very plausible assumption, since if marginal revenue is greater (smaller) than marginal cost, the monopolist's profit increases (decreases) as the quantity produced and sold increases (decreases). Thus the assumption reflects the presumed behaviour of a monopolist in his search for the maximum profit. The formal counterpart of this assumption is

$$\frac{dq}{dt} = \varphi[R'(q) - C'(q)], \quad sgn\,\varphi[...] = sgn\,[...], \quad \varphi[0] = 0, \quad \varphi'[0] > 0.$$

Performing a linear approximation of the function φ at the equilibrium point and putting $\varphi'[0] \equiv k$, we have

$$\frac{dq}{dt} = k[R'(q) - C'(q)].$$

We must now linearise the function $R'(q) - C'(q)$ at the equilibrium point, after which we obtain

$$\frac{d\bar{q}}{dt} = k[R''(q^0) - C''(q^0)]\bar{q},$$

where $\bar{q} = q - q^0$. The solution is

$$\bar{q}(t) = \bar{q}(0)\exp\left\{k\left[R''(q^0) - C''(q^0)\right]t\right\},$$

from which, given the second-order conditions for a maximum, $\lim_{t\to\infty} \bar{q}(t) = 0$, and so the equilibrium point is stable.

20.6 Elements of comparative dynamics

The concept of comparative dynamics is closely related to that of comparative statics. The main difference lies in the fact that comparative dynamics is concerned with the effects of changes in parameters, etc., on the whole *motion over time* of a dynamic economic model. According to Samuelson (1947), the changes which are the concern of comparative dynamics can be any one of the following:

(1) changes in initial condition;

(2) changes in exogenous forces (as, for example, in autonomous investment);

(3) changes in internal parameters (as, for example, in the propensity to save).

The basic method of comparative dynamics can be summarized as follows: we have a set of functional equations, whose solution gives the time path of the economic system. In this solution initial conditions, exogenous elements and internal parameters also appear. Differentiating totally the solution functions (or the functional equations themselves) with respect to the argument representing the element whose shift we are interested in, we try to determine the effect of this shift.

The comparative dynamics method has been extensively employed in growth models to analyse the effect of shifts in some exogenous elements or internal parameters on the steady-state equilibrium growth path of the model. Consider, for example, the neoclassical growth model (Part II, Chap. 13, Sect. 13.2) and ask: what is the effect of an increase in the propensity to save on the steady-state equilibrium growth path? The answer is: none, since in this path all variables grow at the rate n, which is independent of s.

More formally, consider a differential equation system

$$\mathbf{y}'(t) = \mathbf{f}[\mathbf{y}(t), \mathbf{z}(\mathbf{t}), \boldsymbol{\theta}], \tag{20.14}$$

where $\mathbf{y}$ is a vector of endogenous variables, $\mathbf{z}$ a vector of exogenous variables (given functions of time), and $\boldsymbol{\theta}$ a vector of parameters. If we assume that all the exogenous variables grow at a constant proportional rate (which may be equal or different across variables), we have

$$z_i(t) = z_i(0)e^{\gamma_i t}. \tag{20.15}$$

where $z_i(0)$ and γ_i are given. A steady-state (or balanced-growth solution) is a particular solution to system (20.14) having the form

$$y_i(t) = y_i(0)e^{\rho_i t}, \tag{20.16}$$

where the initial values $y_i(0)$ and the growth rates ρ_i are to be determined. The search for the steady state is usually performed by the principle of undetermined coefficients (see Chap. 11, Sect. 11.2.2 for the linear case).

Equations (20.15) and (20.16) are substituted into system (20.14) and the values of $y_i(0)$ and ρ_i are determined so that the system is identically satisfied. In other words, after these substitutions, the condition for (20.16) to be a solution is that the resulting system should be identically satisfied, and this will give rise to a set of equations in the unknowns $y_i(0)$ and ρ_i. It usually happens that the ρ_i are obtained by solving the equations derived from equating to zero the coefficients of the terms containing t, whereas the $y_i(0)$ are obtained by solving the equations derived from equating to zero all the other terms not containing t.

The solution will express the unknowns in terms of the data; usually we shall obtain

$$\begin{aligned} \rho_i &= h_i(\gamma, \boldsymbol{\vartheta}), && (20.17)\\ y_i(0) &= \varphi_i(\mathbf{z}(0), \gamma, \boldsymbol{\theta}), && (20.18) \end{aligned}$$

where $\boldsymbol{\vartheta} \in \boldsymbol{\theta}$, i.e. $\boldsymbol{\vartheta}$ is a vector containing just a few parameters of the full set of the model's parameters. It should also be noted that the functions h_i are usually fairly simple functions, while the functions φ_i are often very complicated functions. All comparative dynamics results are then obtained by calculating the partial derivatives of the functions h_i and φ_i with respect to the element we are interested in.

The conditions of stability of the equilibrium path may be useful in obtaining information on comparative dynamics, and this can be regarded as the analogue of the correspondence principle, where 'dynamic equilibrium path' has been substituted for 'static equilibrium point'.

It must be noted, finally, that comparative dynamics, as such, does not say anything about the transition from one equilibrium growth path to another; the study of this transition belongs to stability analysis.

20.7 An illustrative application of the correspondence principle: the IS-LM model

Let us consider the following elementary macroeconomic model:

$$\begin{aligned} S &= S(Y,R), \quad 0 < \frac{\partial S}{\partial Y} < 1, \quad \frac{\partial S}{\partial R} > 0, && (20.19)\\ I &= I(Y,R), \quad 0 < \frac{\partial I}{\partial Y} < 1, \quad \frac{\partial I}{\partial R} < 0, && (20.20)\\ I &= S, && (20.21)\\ L &= L(Y,R), \quad \frac{\partial L}{\partial Y} > 0, \quad \frac{\partial L}{\partial R} < 0, && (20.22)\\ L &= L_s^*. && (20.23) \end{aligned}$$

The price level is assumed to be rigid. The symbols are the usual ones and the equations are self-explanatory: in the order, they are the saving function, the investment function, the *ex-ante* equality between investment and saving (real equilibrium), the demand for money function, the equality between the demand for and the (exogenously given) supply of money (monetary equilibrium). The *a priori* assumptions on the various functions are condensed in the assumed signs of the various partial derivatives. Substituting from the first two equations in the third and from the fourth in the fifth, we obtain

$$\begin{aligned} I(Y,R) - S(Y,R) &= 0, \qquad (20.24)\\ L(Y,R) - L_s^* &= 0. \qquad (20.25)\end{aligned}$$

Eqs. (20.24) and (20.25) determine, in the (R,Y) plane, two curves which are the well-known *IS* and *LM* (or *LL*) curves, following the terminology suggested by Hicks (1937). The intersection of these curves determines the equilibrium point, say R_e, Y_e, in which real and monetary equilibrium obtain simultaneously.

In the system composed of Eqs. (20.24) and (20.25) a parameter L_s^* is already present. We can introduce at least three other parameters, one for each of the functions S, I, L, that is

$$\begin{aligned} I(Y,R,\alpha_1) - S(Y,R,\alpha_2) &= 0, \qquad (20.26)\\ L(Y,R,\alpha_3) - L_s^* &= 0, \qquad (20.27)\end{aligned}$$

where we conventionally establish that $\partial I/\partial\alpha_1 > 0, \partial S/\partial\alpha_2 > 0, \partial L/\partial\alpha_3 > 0$. In other words, α_1 is defined in such a way that I varies in the same direction as α_1 (α_1 is, for example, government expenditure), and similarly α_2, α_3.

Provided that the Jacobian

$$\begin{vmatrix} \dfrac{\partial I}{\partial Y} - \dfrac{\partial S}{\partial Y} & \dfrac{\partial I}{\partial R} - \dfrac{\partial S}{\partial R} \\[2ex] \dfrac{\partial L}{\partial Y} & \dfrac{\partial L}{\partial R} \end{vmatrix}$$

is non-zero at the point R_e, Y_e, there exist the functions

$$\begin{aligned} Y &= Y(\alpha_1, \alpha_2, \alpha_3, L_s^*), \qquad (20.28)\\ R &= R(\alpha_1, \alpha_2, \alpha_3, L_s^*), \qquad (20.29)\end{aligned}$$

of which we want to know the partial derivatives $\partial Y/\partial\alpha_1, \partial R/\partial\alpha_1$, etc., which give us the reaction of the equilibrium values of income and the rate of interest to a shift in the parameters. Let us note that not all the parameters appear in each of the relations (20.26) and (20.27), and this is an element which, in general, makes it easier to obtain the desired comparative statics results. We shall examine only the effects of a shift in government expenditure; the student can complete the analysis as an exercise.

Totally differentiating (20.26) and (20.27) with respect to α_1—account being taken of (20.28) and (20.29)—we have

$$\begin{aligned} \left(\frac{\partial I}{\partial Y} - \frac{\partial S}{\partial Y}\right)\frac{\partial Y}{\partial \alpha_1} + \left(\frac{\partial I}{\partial R} - \frac{\partial S}{\partial R}\right)\frac{\partial R}{\partial \alpha_1} &= -\frac{\partial I}{\partial \alpha_1}, \\ \frac{\partial L}{\partial Y}\frac{\partial Y}{\partial \alpha_1} + \frac{\partial L}{\partial R}\frac{\partial R}{\partial \alpha_1} &= 0, \end{aligned} \qquad (20.30)$$

whose solution is

$$\begin{aligned} \frac{\partial Y}{\partial \alpha_1} &= \frac{-\dfrac{\partial I}{\partial \alpha_1}\dfrac{\partial L}{\partial R}}{\Delta}, \\ \frac{\partial R}{\partial \alpha_1} &= \frac{\dfrac{\partial I}{\partial \alpha_1}\dfrac{\partial L}{\partial Y}}{\Delta}, \end{aligned} \qquad (20.31)$$

where

$$\Delta \equiv \left(\frac{\partial I}{\partial Y} - \frac{\partial S}{\partial Y}\right)\frac{\partial L}{\partial R} - \left(\frac{\partial I}{\partial R} - \frac{\partial S}{\partial R}\right)\frac{\partial L}{\partial Y}. \qquad (20.32)$$

The assumptions made at the beginning on the various partial derivatives allow us to determine the sign of the numerators of $\partial Y/\partial\alpha_1$ and $\partial R/\partial\alpha_1$, which turn out to be both positive, but do not allow us to determine the sign of Δ.

Let us then examine the problem of the stability of the equilibrium point in the model under consideration. The order of the exposition could also have been inverted. That is, one could examine first the stability problem and then the comparative statics problem.

Plausible behaviour assumptions, which are commonly used in studying the dynamics of the Keynesian model, are the following:

(1) In a context of rigid prices and less than full employment, an *ex-ante* difference between investment and saving reacts on income, which increases (decreases) if investment exceeds (falls short of) saving. Let us note, incidentally, that this assumption is the same as the following: income increases (decreases) if aggregate demand exceeds (falls short of) current output (which is the same as income). In fact, since consumption is income minus saving, the inequality $I \gtrless S$ is the same as $I \gtrless Y - C$, that is, as $C + I \gtrless Y$.

(2) The rate of interest tends to increase (decrease) if demand for money exceeds (falls short of) the supply. The rationale of this assumption is the following. Consider, for example, a positive excess demand for money. The scarcity of liquidity drives the owners of bonds to try to sell them; this brings about a fall in bond prices, that is, an increase in the interest rate, the latter being inversely proportional to the price of bonds.

The formal counterpart of these assumptions is

$$\begin{aligned}\frac{dY}{dt} &= \varphi_1 [I(Y,R) - S(Y,R)], \ \mathrm{sgn}\varphi_1[...] = \mathrm{sgn}[...], \varphi_1[0] = 0, \varphi_1'[0] > 0,\\ \frac{dR}{dt} &= \varphi_2[L(Y,R) - L_s^*], \quad \mathrm{sgn}\varphi_2[...] = \mathrm{sgn}[...], \varphi_2[0] = 0, \varphi_2'[0] > 0.\end{aligned}$$

Performing a linear approximation at the equilibrium point and denoting with a bar the deviations from equilibrium, we have

$$\begin{aligned}\frac{d\overline{Y}}{dt} &= c_1\left(\frac{\partial I}{\partial Y} - \frac{\partial S}{\partial Y}\right)\overline{Y} + c_1\left(\frac{\partial I}{\partial R} - \frac{\partial S}{\partial R}\right)\overline{R},\\ \frac{d\overline{R}}{dt} &= c_2\frac{\partial L}{\partial Y}\overline{Y} + c_2\frac{\partial L}{\partial R}\overline{R},\end{aligned} \tag{20.33}$$

where $c_1 \equiv \varphi_1'[0]$, $c_2 \equiv \varphi_2'[0]$. The characteristic equation of system (20.33) is

$$\begin{vmatrix} c_1\left(\frac{\partial I}{\partial Y} - \frac{\partial S}{\partial Y}\right) - \lambda & c_1\left(\frac{\partial I}{\partial R} - \frac{\partial S}{\partial R}\right) \\ c_2\frac{\partial L}{\partial Y} & c_2\frac{\partial L}{\partial R} - \lambda \end{vmatrix} = 0,$$

that is

$$\lambda^2 - \left[c_1\left(\frac{\partial I}{\partial Y} - \frac{\partial S}{\partial Y}\right) + c_2\frac{\partial L}{\partial R}\right]\lambda + c_1c_2\Delta, \tag{20.34}$$

where Δ is the same expression defined in (20.32). The necessary and sufficient stability conditions are

$$\begin{aligned}c_1\left(\frac{\partial I}{\partial Y} - \frac{\partial S}{\partial Y}\right) + c_2\frac{\partial L}{\partial R} &< 0,\\ \Delta &> 0.\end{aligned}$$

Using the condition $\Delta > 0$, we can establish that both $\partial Y/\partial\alpha_1$ and $\partial R/\partial\alpha_1$ are positive, which is the standard result that in a Keynesian model an increase in government expenditure causes an increase in both output and the rate of interest.

Let us note, finally, that in the 'pure' Keynesian case, that is $\partial I/\partial Y \simeq 0$, $\partial S/\partial R \simeq 0$, complete comparative statics results can be obtained without any dynamic consideration.

20.8 Exercises

1. Consider the Mundell-Fleming extension of the *IS-LM* model to an open economy under fixed exchange rates, and define the balance of payments as

$B = X_0 - M(Y) + K(R)$, X_0 being exogenous exports, M imports ($0 < \partial M/\partial Y < 1$), K net capital inflows ($\partial K/\partial R > 0$). We then have the model

$$\begin{aligned}X_0 + I(Y,R) - S(Y,R) - M(Y) &= 0,\\ L(Y,R) - L_s &= 0,\\ X_0 - M(Y) + K(R) &= 0,\end{aligned}$$

where the third equation is the equilibrium condition for the balance of payments. Note that the money supply is can no longer be considered exogenous (otherwise the system would be overdetermined) but is an endogenous variable whose dynamic behaviour is governed by the additional differential equation

$$\frac{dL_s}{dt} = c_3(B), \quad c_3 > 0,$$

according to which—in the absence of sterilization operations from the monetary authorities—the money supply increases (decreases) if there is a surplus (deficit) in the balance of payments. Examine the stability of the model and its comparative-static properties in the case of

(1.a) an exogenous increase in exports,
(1.b) an exogenous increase in imports,
(1.c) an exogenous increase in government expenditure.

2. Consider a single market in a partial equilibrium context, $D = D(p,\alpha_1)$, $S = S(p,\alpha_2)$, $D = S$, where α_1 represents a shift in tastes toward the good in question (hence $\partial D/\partial\alpha_1 > 0$) and α_2 a reduction in costs (hence $\partial S/\partial\alpha_2 > 0$). Derive the comparative statics results assuming that the equilibrium is stable. Is there a difference between the Walrasian and Marshallian cases? (see Chap.13, Sect. 13.1).

3. Consider the growth model with human capital examined in Chap. 19, Sect. 19.2, and suppose that there is a shift in investment from physical to human capital (i.e., s_k decreases, and s_h increases by the same amount). Determine the effects on the steady-state values of the variables k, h, and y.

20.9 References

Arrow, K.J. and F.H. Hahn, 1971, *General Competitive Analysis,* Chap. 1, § 6: Chap. 12, § 11.
Arrow, K.J. and M. McManus, 1958, A Note on Dynamic Stability.
Basset, L., 1968, The Solution of Qualitative Comparative Static Problems: Comment.
Basset, L., J. Maybee and J. Quirk, 1968, Qualitative Economics and the Scope of the Correspondence Principie.

Brock, W.A., 1986, A Revised Version of Samuelson's Correspondence Principle: Application of Recent Results on the Asymptotic Stability of Optimal Control to the Problem of Comparing Long-Run Equilibrium.
Debreu, G., 1974, Excess Demand Functions.
Farley, A.M. and K.-P. Lin, 1990, Qualitative Reasoning in Economics.
Gale, D. and H. Nikaidô, 1965, The Jacobian Matrix and Global Univalence of Mappings.
Hicks, J.R., 1937, Mr Keynes and the 'Classics': A Suggested Interpretation.
Hicks, J.R., 1939, *Value and Capital,* Parts I and II, and Mathematical Appendixes.
Hobson, E.W., 1957, *The Theory of Functions of a Real Variable & the Theory of Fourier's Series,* Vol. I, § 319.
Kemp, M.C., Y. Kumura and M. Tawada, 1990, The Global Correspondence Principle.
Lancaster, K., 1966, The Solution of Qualitative Comparative Statics Problems.
Lloyd, P.J., 1969, Qualitative Calculus and Comparative Static Analysis.
Mantel, R.R., 1974, On the Characterization of Aggregate Excess Demand.
Nikaidô, H., 1968, *Convex Structures and Economic Theory,* Chap. VII.
Patinkin, D., 1952, The Limitations of Samuelson's 'Correspondence Principle'.
Petit, M.L., 1990, *Control Theory and Dynamic Games in Economic Policy Analysis,* Chap. 3, Sect. 3.2.
Quirk, J., and R. Saposnick, 1968, *Introduction to General Equilibrium Theory and Welfare Economics,* Chap. 6.
Rader, T., 1972, Impossibility of Qualitative Economics: Excessively Strong Correspondence Principles in Production-Exchange Economics.
Ritschard, G., 1983, Computable Qualitative Comparative Static Techniques.
Samuelson, P.A, 1941b, The Stability of Equilibrium: Comparative Statics and Dynamics.
Samuelson, P.A, 1947, *Foundations of Economic Analysis,* Chaps. II, III, IX, X (pp. 294-6 and 301-2), XII.
Slutsky, E., 1953 (1915), On the Theory of the Budget of the Consumer.
Sonnenschein, H., 1972, Market Excess Demand Functions.

Chapter 21

Stability of Equilibrium: A General Treatment

21.1 Introduction

The central place of the concept of equilibrium in economic theory justifies the interest of economists in stability, an interest that dates back at least to Marshall (1879a,b) and Walras (1874). How many times has the reader seen an egg standing upon its end, as aptly Marshall (1890, pp. 424-5 note)[1] and Samuelson (1947, p. 5) put it. 'Positions of unstable equilibrium, even if they exist, are transient, nonpersistent states, and hence on the crudest probability calculation would be observed less frequently than stable states' (Samuelson, *ibidem*).

Economists employ the concept of stability in much the same way as mathematicians and physicists. Let the more scientific pendulum take the place of the down-to-earth egg, and 'from a physical point of view only equilibria that are stable are of interest. A pendulum balanced upright is in equilibrium, but this is very unlikely to occur; moreover, the slightest disturbance will completely alter the pendulum's behaviour. Such an equilibrium is unstable' (Hirsch and Smale, 1974, p. 145). An equilibrium which (supposing that it has been hit upon by chance) is such that the slightest disturbance starts a movement away from it—i.e., an unstable equilibrium—is usually considered unmeaningful from an economic point of view.

The concern over stability has led economists either to *assume* that equilibrium is stable (this involves the study of the *necessary* stability conditions and their possible use in comparative statics) or to search for 'plausible' conditions that *ensure* stability. These latter are usually *sufficient* stability conditions only (this is, for example, the approach taken in the study of the stability of general competitive equilibrium).

However, in some cases, the presence of *instability* is an essential feature

[1] Subsequently (since the 5th edition) in Appendix H, § 2, note 1.

as, for example, in Hicks' trade cycle model (see Chap. 8, Sect. 8.2), and, in general, in business cycle models: if the movement is intrinsically stable, one has to rely on exogenous shocks to keep the oscillation alive, and this is hardly a satisfactory explanation.

Besides stable and unstable equilibria there are also *neutral* equilibria: an equilibrium is neutral if a (small) disturbance leads to a new situation which does not change unless there is a further disturbance.

The study of stability follows two main approaches : the *qualitative (*or *topological)* and the *quantitative.* The *qualitative approach* consists in the analysis of the properties of the solutions to a differential equation (system) without actually knowing the solution itself or trying to approximate it; this is based on phase diagrams, Liapunov's second method, etcetera. On the contrary, the *quantitative approach* consists in trying to find the explicit analytical solution of the differential equation or to approximate it by using power series and other methods. It is then easy to check whether the time path of the system converges to the equilibrium point. This distinction, and the treatment that follows, apply equally well if one considers difference equations instead of differential equations.

It is important to observe that the quantitative approach remains an analytical method, and should not be confused with the *numerical* method(s) of integration, a topic that lies outside the scope of the present book.

The types of differential equations whose closed-form solution is known have been codified long ago (see the first four volumes of the monumental six-volume tratado treatise by Forsyth, 1890-1902, and the book by Ince, 1926). A few types that have had some, although limited, application in economic dynamics, will be treated in Chap. 24.

It should however be noted that in theoretical mathematics there has been a shift of emphasis from quantitative to qualitative methods. This reflects the well-known fact that in most cases, although we know that the solution exists (by the existence and uniqueness theorem), we are not able to 'find' it, since it cannot be expressed in terms of (a finite number of) known functions. Qualitative methods are even more necessary in economic dynamics, where we usually do not even know the form of the functions involved but only some of their qualitative properties (e.g., the signs of the partial derivatives). However, there is a quantitative method that is widely used in economic dynamics: the linearisation method, which is the standard tool for studying local stability. Hence we shall treat it fairly extensively.

21.2 Basic concepts and definitions

By and large, the idea of stability that economists have in mind corresponds to what mathematicians call *asymptotic stability,* which is a stronger property than simple *stability* (in the mathematical sense). Therefore some definitions

are in order to clarify the matter.

21.2.1 Stability

A system is *stable* if, when perturbed slightly from its equilibrium state, all subsequent motions remain in a correspondingly small neighbourhood of the equilibrium. If, in addition to being stable, every motion starting sufficiently near the equilibrium point *converges* to it as $t \to \infty$, then the equilibrium is *asymptotically stable.* These are *local* concepts (also called 'in the small'). If stability is independent of the distance of the initial state from the equilibrium point, we have *(asymptotic) stability in the large,* also called *global (asymptotic) stability.*

Let us now come to the formal definitions. Consider a generic non-autonomous vector-differential equation system (which may of course be a single equation)

$$\mathbf{y}' = \mathbf{f}(\mathbf{y}, t), \tag{21.1}$$

and assume that the (vector-valued) function $\mathbf{f}$ is sufficiently smooth so that the system has a unique solution

$$\mathbf{y}(t) = \phi(t; \mathbf{y}_0, t_0) \tag{21.2}$$

starting at any initial state $\mathbf{y}_0$ at any time t_0. The system is called *non autonomous* because it involves time explicitly; if t were not present as an explicit argument of $\mathbf{f}$ the system would be *autonomous.*

An *equilibrium state (*also called *rest point, fixed point, null solution)* $\mathbf{y}_e$ is a state of the dynamic system (21.1) such that

$$\mathbf{f}(\mathbf{y}_e, t) = \mathbf{0} \text{ for all } t \text{ or, equivalently, } \phi(t; \mathbf{y}_e, t_0) = \mathbf{y}_e \text{ for all } t. \tag{21.3}$$

In what follows, the notation $\|...\|$ will be used to indicate the absolute value or *norm.* We then have

Definition 21.1 (stability in the sense of Liapunov) An equilibrium state $\mathbf{y}_e$ of the dynamic system (21.1) is *stable* if for every real number $\epsilon > 0$ there exists a real number $\delta(\epsilon, t_0) > 0$ such that the inequality $\|\mathbf{y}_0 - \mathbf{y}_e\| \leq \delta$ implies

$$\|\phi(t; \mathbf{y}_0, t_0) - \mathbf{y}_e\| \leq \epsilon \quad \text{for all } t \geq t_0.$$

Figure (21.1) illustrates the definition in the simple case of a single variable. In this case $\|\mathbf{y}_0 - \mathbf{y}_e\| \leq \delta$ means any initial point lying within or on the boundary of the circle of radius δ, and $\|\phi(t; \mathbf{y}_0, t_0) - \mathbf{y}_e\| \leq \epsilon$ means that the motion remains within or on the boundary of the circle of radius ϵ. In the case of a system, the circles are replaced by spheres or hyperspheres (also called balls).

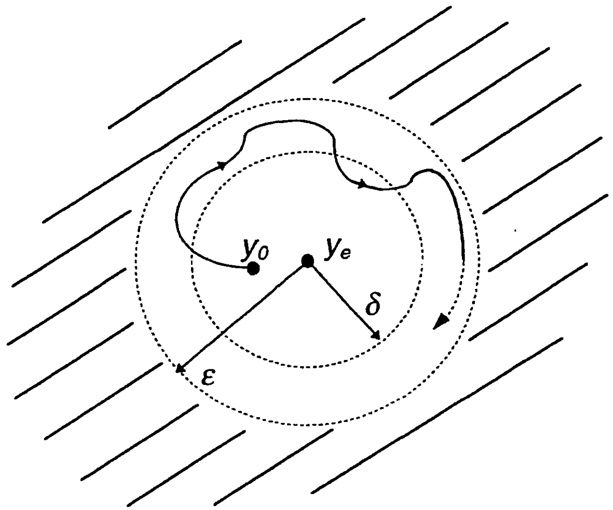


Figure 21.1: Stability

Note that stability in the sense of Liapunov is a *local* concept since it refers to behaviour near the equilibrium state. The reason is that one does not know a priori how small δ may have to be chosen.

We now pass to

Definition 21.2 (asymptotic stability in the sense of Liapunov) An equilibrium state $\mathbf{y}_e$ of the dynamic system (21.1) is *asymptotically stable* if

(i) it is stable, and

(ii) every motion starting sufficiently near $\mathbf{y}_e$ converges to $\mathbf{y}_e$ as $t \to \infty$. In other words, to every real number $\mu > 0$ there corresponds a real number $T(\mu, \mathbf{y}_0, t_0) > 0$ such that the inequality $\|\mathbf{y}_0 - \mathbf{y}_e\| \leq r(t_0)$, where $r(t_0) > 0$ is a real constant, implies '' $\delta(\varepsilon, t_0)$

$$\|\phi(t; \mathbf{y}_0, t_0) - \mathbf{y}_e\| \leq \mu \quad \text{for all } t \geq t_0 + T,$$

namely

$$\lim_{t \to \infty} \phi(t; \mathbf{y}_0, t_0) = \mathbf{y}_e.$$

Figure (21.2) illustrates the definition, again in the simple case of a single variable. The inequality $\|\mathbf{y}_0 - \mathbf{y}_e\| \leq r(t_0)$ defines a ball in the hyperplane $t = t_0$ and the inequality $\|\phi(t; \mathbf{y}_0, t_0) - \mathbf{y}_e\| \leq \mu$ determines a hypercylinder (also called tube) of radius μ around the t axis. Then by choosing the initial point in a sufficiently small ball we can force the path of the solution

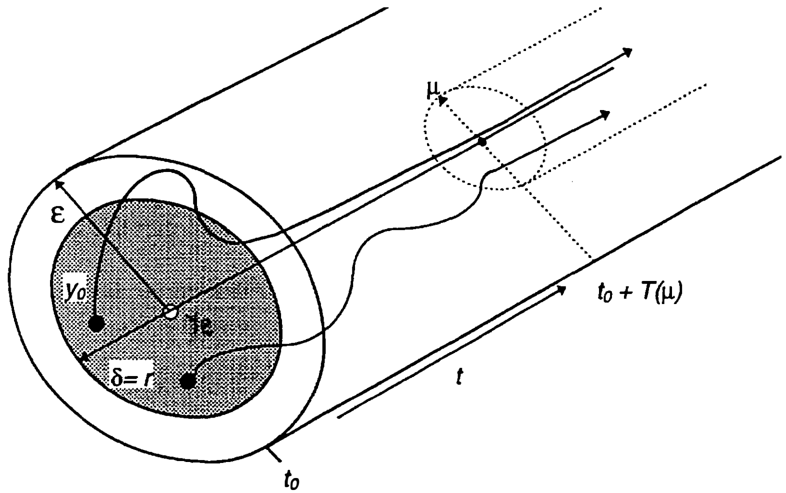


Figure 21.2: Asymptotic stability

to remain entirely inside the hypercylinder, however small, from a certain time (which depends on μ as well as on the initial state and initial time) adelante onwards. Asymptotic stability (called 'stability of the first kind in the small' by Samuelson, 1947, pp. 261-262) is also a local concept, for one does not know a priori how small $r(t_0)$ may have to be.

An asymptotically stable point is also called an *attractor* (by converse, an unstable point is called *a* repulsor repeller). Note, however, that the concept of attractor is broader that an asymptotically stable point (see Chap. 26, Sect. 26.3.1). The set of all initial points which are attracted by the stable fixed point, namely the set of points such that $\|\mathbf{y}_0 - \mathbf{y}_e\| \leq r(t_0)$ (see Definition 21.2), is called the *basin of attraction* of $\mathbf{y}_e$.

When stability is independent of the distance of the initial state from $\mathbf{y}_e$ we speak of *stability in the large* or *global* stability ('perfect stability of the first kind' in Samuelson's terminology, 1947, pp. 261-262). More precisely, we have

Definition 21.3 (asymptotic stability in the large) An equilibrium state $\mathbf{y}_e$ of the dynamic system (21.1) is *asymptotically stable in the large* if

(i) it is stable, and

(ii) every motion converges to $\mathbf{y}_e$ for $\|\mathbf{y}_0 - \mathbf{y}_e\| \leq r$, where r is fixed but arbitrarily large, as $t \to \infty$.

A somewhat puzzling feature of Definition 21.2 (and hence of Definition 21.3) is that one has to assume *explicitly* that 21.1 holds. This is due to the fact that Point (ii) of Definition 21.2 does *not* in general imply Point (i) unless we are in the case of a single equation (see exercise 4, and Kalman and Bertram, 1960, pp. 375-376). However, if all motions are continuous in $\mathbf{y}_0$, then (ii) implies (i). In economic applications the assumption that all motions are continuous in $\mathbf{y}_0$ seems reasonable, hence asymptotic stability is usually

When the various constants that appear in the definitions do not depend on the initial time, we speak of *uniform* (with respect to time) stability. More precisely, we have

Definition 21.4 An equilibrium state $\mathbf{y}_e$ of the dynamic system (21.1) is *uniformly stable* if δ in Definition 21.1 does not depend on t_0.

Definition 21.5 An equilibrium state $\mathbf{y}_e$ of the dynamic system (21.1) is *uniformly asymptotically stable* if δ, r, T that appear in Definition 21.2 do not depend on t_0.

Definition 21.6 An equilibrium state $\mathbf{y}_e$ of the dynamic system (21.1) is *uniformly asymptotically stable in the large* when it is asymptotically stable in the large, and δ, r, T do not depend on t_0.

It should be noted that in the case of uniform asymptotic stability (both in the small and in the large), the assumption of uniform stability cannot be dispensed with, as shown for example by Kalman and Bertram (1960, p. 377).

The definitions above are particularly simple to check when the differential equation system (21.1) is linear and with constant coefficients. Note that the linear form of (21.1) is a first-order system in normal form. After the extensive treatment of these systems given in Chap. 18, the following statements turn out to be obvious:

Property L.1 The null solution of a constant-coefficient linear system is asymptotically stable if, and only if, all the latent roots of the system's matrix have negative real parts (this is also called a *sink).*

Property L.2 If the null solution of a constant-coefficient linear system is asymptotically stable, then it is asymptotically stable in the large.

Property L.3 The null solution of a constant-coefficient linear system is unstable if at least one root has positive real part (the case in which all the roots have positive real parts is called a *source*).

Property L.4 If a constant-coefficient linear system has no root with positive real part but at least one root with zero real part, the null solution cannot be asymptotically stable, but it can be stable provided that all the roots with zero real part are simple roots (recall that, in the case of multiple roots, the solution will contain terms of the type $P(t)e^{(rprr)t}$, where $P(t)$ are polynomials in t, and $(rprr)$ denotes the real part of any repeated root).

21.2.2 Further definitions

It may sometimes happen (see Chap. 18, Sect. 18.2.2.3 and below, Fig, 21.3) that an equilibrium point is stable or unstable according to the initial position of the system. This is called *one-sided stability-instability* (in Samuelson's terminology, 1947, pp. 295-6), *semi-stability,* or *conditional stability* (i.e., stability conditional on the initial position). More generally, we define the *stable manifold* as the set of points for which $\lim_{t\to\infty} \phi(t; \mathbf{y}_0, t_0) = \mathbf{y}_e$, namely $\mathbf{y}(t) \to \mathbf{y}_e$ as $t \to \infty$. Where this has dimension one, it is referred to as the stable arm. In the case of a linear system (which may be the linear approximation to a non-linear system: see below, Sect. 21.4.1), the dimension of the stable manifold is equal to the number of characteristic roots with negative real part.

Thus semi-stability is sometimes considered an intermediate concept between that of full stability (i.e., stability in the sense of Liapunov, which holds for any initial point) and that of full instability (no matter where the initial point lies, the motion is away from equilibrium).

The definitions of stability are sometimes taken to imply that the equilibrium point is unique, but this is true only for global asymptotic stability (henceforth we shall omit the adjective asymptotic, on the understanding that it is implicitly present, unless its presence is necessary to avoid misunderstanding). In fact, if there are multiple equilibria none of them can be globally stable; on the contrary, there can be locally stable multiple equilibria (but usually they will be alternatively stable and unstable).

The possible occurrence of multiple equilibria in economic models has led mathematical economists to introduce the concepts of *global stability of an adjustment process* and of *quasi-stability.* An *adjustment process* is said to be *globally stable* if for every initial state there is an equilibrium point which the system converges to (this point need not be the same for all initial situations). The concept of *quasi-stability* is more complicated and involves the construction of a sequence of points (from any initial point) corresponding to a sequence of times. If any such sequence of points converges to a point when the sequence of times tends to infinity, and if all these limit points are equilibrium points, then the process is called quasi-stable. This has important applications in the theory of general competitive equilibrium.

A related problem that worries economists is that of *path-dependent* equilibrium and stability. This was noted long ago (Marshall, 1879a, Chap. II; 1879b, Chap. I, § 7; Edgeworth, 1881, pp. 19-20) and was called the problem of the indeterminateness of equilibrium by Kaldor (1934, pp. 17ff; for a modern treatment see Fisher, 1983). It is due to the fact that the movement of the system when it is out of equilibrium may change the data on which the static equations which define the equilibrium are based, so that these equations will change and determine a *different* equilibrium and so on and so forth. In other words, the (set of) equilibrium points is *not independent of the dynamic movement of the system,* i.e. this set is path-dependent. Also this problem is particularly important in general competitive equilibrium.

If path-dependence occurs, it is logically mistaken to study (the existence of) equilibrium without simultaneously studying its stability, unless the path-dependence is in some sense 'small' and/or the convergence to equilibrium is almost instantaneous or very fast so that, as it were, the equilibrium point is 'reached' before any appreciable change in the data can occur. This leads us to two considerations.

The first is that the study of the *rapidity of convergence* becomes essential. It should also be noted that this study is important in any case, since the knowledge that the system tends to equilibrium as $t \to \infty$ is not very interesting if, due to the slowness of the motion, a great deal of time is required for the system to get reasonably near to equilibrium (see, for example, Chap. 13, Sect. 13.2.4.1).

The second is the importance of the study of *structural stability,* that is yet another concept of stability, which the next subsection is devoted to.

21.2.3 Structural stability

The problem of structural stability arose in mathematical physics precisely because none of the factors taken as given can remain absolutely constant during the motion of the system, so that when one hypothesises that certain parameters are constant one is really assuming that small variations in these parameters do not significantly alter the character of the motion.

Since models are often fairly crude representations of reality, it would be nice to know that 'roughly similar' models give 'roughly similar' paths. This robustness would appear to be a reasonable requirement for any plausible model. Dynamic systems that are such as not to vary in their essential features for a small variation of the form of the differential equations are called structurally stable systems or 'coarse' systems (in the terminology of Andronov et al., 1966, Chap. VI, § 4). More precisely, consider an autonomous (i.e., in which time does not appear explicitly) 2×2 system

$$\begin{aligned} y_1' &= f_1(y_1, y_2), \\ y_2' &= f_2(y_1, y_2), \end{aligned} \tag{21.4}$$

and the modified system

$$\begin{aligned} y_1' &= f_1(y_1, y_2) + p(y_1, y_2), \\ y_2' &= f_2(y_1, y_2) + q(y_1, y_2), \end{aligned} \tag{21.5}$$

where $p(y_1, y_2), q(y_1, y_2)$ are small and analytical, and have small partial derivatives. Then we have the following definition (Andronov et al., 1966, p. 376):

Definition 21.7 (*structural stability*) System (21.4) is called 'coarse' in a region G if for any $\epsilon > 0$ there is a $\delta > 0$ such that for all analytic functions $p(y_1, y_2), q(y_1, y_2)$ that satisfy in G the inequalities

$$\begin{aligned} &p(y_1, y_2) < \delta, \quad p_{y_1}(y_1, y_2) < \delta, \quad p_{y_2}(y_1, y_2) < \delta, \\ &q(y_1, y_2) < \delta, \quad q_{y_1}(y_1, y_2) < \delta, \quad q_{y_2}(y_1, y_2) < \delta, \end{aligned}$$

there exists a topological transformation of G into itself, for which each path of system (21.4) is transformed into a path of the modified system (21.5) (and conversely), the points that correspond to each other in this transformation being found at distances less than ϵ.

Structural stability can of course be applied to single equations as well as to n-equation systems. A typical case of a structural stability problem is when the basic equations depend on a parameter, and forms the subject of bifurcation theory (see Chap. 25).

The terminology 'structural stability' is somewhat misleading, since it seems to point to some form of (Liapunov) stability as defined in Sect. 21.2. This is not true, for what matters here is the preservation of the qualitative properties of the system, whatever they are. From the mathematical point of view, a (Liapunov) unstable system that remains unstable under slight changes is structurally stable. Hence the denomination 'coarse' (as originally suggested by Andronov et al., 1966, p. 374) or 'noncritical' (as suggested by Minorsky, 1962, p. 185) might seem preferable. However, the denomination 'structural stability' is by now standard in the mathematical literature and we shall use it as well.

It goes without saying that one is mainly interested in (Liapunov) stable systems that are also structurally stable: an expression like 'stable coarse system' or 'unstable coarse system' would be clear, while 'stable structurally stable system' or 'unstable structurally stable system' are certainly confusing.

The belief that good models are those that are structurally stable was embodied in a *stability dogma,* in which structurally unstable systems were regarded as somehow suspect. 'This dogma stated that, due to measurement uncertainties, etc., a model of a physical system was valuable only if its qualitative properties did not change with perturbations.' (Guckenheimer and Holmes, 1986, p. 259). In physics, this dogma has been found faulty, and has been reformulated to state that 'the only properties of a dynamical system...which are *physically relevant* are those which are preserved under perturbations of the system. The definition of physical relevance will clearly depend on upon the specific problem. This is quite different from the original statement that the only good systems are ones with *all* of their qualitative properties preserved by perturbations' (Guckenheimer and Holmes, 1986, p. 259).

In economics (in what follows, by 'structurally stable' we refer to systems that are Liapunov asymptotically stable and structurally stable) we must distinguish between purely theoretical and applied economics. In the purely theoretical field some writers deny that structural stability is necessarily desirable (see, for example, Blatt, 1983, Chaps. 7 and 8; Gabisch and Lorenz, 1989, Chap. 5, Sect. 5.3), and point out that the works of Marx, Schumpeter, and Keynes can be reinterpreted by using the general notion of structural instability (see, for example, Vercelli, 1984). But the mainstream emphasis upon stable equilibria coupled with the fact that in economic models the values of the parameters are even more uncertain than in physics, is definitely albeit implicitly in favour of a *revised* stability dogma as that set forth by Guckenheimer and Holmes (simply replace 'physically' and 'physical' with 'economically' and 'economic').

This leads us to applied economics, since the numerical values of the parameters are not 'measured' experimentally, but estimated econometrically. Hence the study of the structural stability of the model is important, especially if one contemplates policy applications of the model itself. If slight changes in (the numerical value of) a parameter lead to a bifurcation value, i.e. to a value of the parameter at which a qualitative change in the nature of equilibrium (e.g., from stable to unstable) occurs, then one has to proceed with care in any practical application of the model. To avoid terminological confusion, we stress that we are using 'structural stability' in the sense defined above, and not in the different sense usually given to 'structural stability' in econometrics (see any econometrics textbook).

Of course, in applied economics two related problems arise: (i) how to pinpoint the parameter(s) 'responsible' for the bifurcation, and (ii) how to give a practical contents to the theoretical notion of 'slight' changes. Both problems can be solved as follows (see Gandolfo, 1992).

As regards the first problem, it can be solved by using sensitivity analysis. Consider the j-th latent root of the linearised version of the model under consideration (on linearisation see below, Sect. 21.4.1) and compute

$$d\lambda_j = \sum_{i=1}^{p} \left(\frac{\partial \lambda_j}{\partial \theta_i}\right) d\theta_i$$

or, if one wishes to consider a particular critical parameter only,

$$d\lambda_j = \frac{\partial \lambda_j}{\partial \theta_i} d\theta_i,$$

where the partial derivatives are given by the formulae discussed in Chap. 18, Sect. 18.2.2.2. Therefore, if $\lambda_j \neq 0$, by letting $\lambda_j + d\lambda_j = 0$, one can determine the corresponding $d\theta_i$,

$$d\theta_i = -\lambda_j \Big/ \frac{\partial \lambda_j}{\partial \theta_i},$$

and so the neighbourhood of the bifurcation value of the i-th parameter, $d\theta_i + \theta_i$.

As regards the second problem, since we are in the presence of *estimated* parameters, we suggested to define 'slight' those changes which lie within the confidence interval built around the parameter. The validity of this practical rule follows immediately from the fact that the 'true' value of the parameter can lie anywhere in the confidence interval (with the given probability). The model, therefore, is structurally stable (unstable) with respect to a parameter when the bifurcation value of the parameter lies outside (inside) the confidence interval for the same parameter. For more details see Gandolfo (1992).

21.3 Qualitative methods: phase diagrams

Let us recall from Sect. 21.1 that by 'qualitative' or 'topological' theory of differential equations we mean the analysis of the properties of the solution of a differential equation (or a system) without actually knowing the solution itself nor trying to approximate it by means of power series or other 'quantitative' methods[2]. The same definition applies of course to difference equations. Given this definition, Liapunov's second method for the study of global stability problems also belongs to the realm of qualitative analysis. However, the importance of Liapunov's second method is such that we have treated it in a separate chapter of its own (Chap. 23).

We shall examine single first-order equations and systems of two simultaneous first-order equations (the latter can also be considered as the transformation of a second-order equation). In any case the equations considered will be *autonomous*, that is, not involving time explicitly nor involving given function(s) of time. Of course, they involve time indirectly, through the unknown function and its derivative. With reference to (21.6) below, non-autonomous equations would be $\varphi(y, dy/dt, t) = 0$, $\varphi(y, dy/dt, u(t)) = 0$, $\varphi(y, dy/dt, u(t), t) = 0$, where $u(t)$ is a known function.

The techniques used will be mainly geometric.

[2] Among the 'quantitative' methods, Krylov and Bogoliubov's method of *equivalent linearization* may have a future in economics. This method consists in the reduction of a non-linear differential equation (or system) to a quasilinear equation which retains the essential features of the original equation and can be analyzed by linear methods; in other words, the original non-linear equation is replaced by an *equivalent linear equation* with the property that the solutions of the two equations can be made to differ from each other by an error of a certain order. For this method see, e.g. Minorsky (1962, Part II, Chap. 14, § 7, pp. 384 ff.; the student interested in a general treatment of quantitative methods may consult Part II of Minorsky's book); for an economic application to Goodwin's (1951) non-linear business cycles model see Bothwell (1952).

21.3.1 Single equations

A graphical technique widely used to analyse first-order autonomous equations, that is, equations of the type

$$\varphi\left(y, \frac{\mathrm{d}y}{\mathrm{d}t}\right) = 0, \tag{21.6}$$

is the so-called *phase diagram*. Before explaining this technique a few considerations are necessary. Let us assume that (21.6) is explicit or can be made so, that is having (or such that it can be put in) the form

$$\frac{\mathrm{d}y}{\mathrm{d}t} = f(y). \tag{21.7}$$

Equation (21.7) has separable variables, since we can write it as $\mathrm{d}y - f(y)\mathrm{d}t = 0$, that is,

$$\frac{1}{f(y)}\mathrm{d}y - \mathrm{d}t = 0, \tag{21.8}$$

which gives

$$\int \frac{1}{f(y)}\mathrm{d}y - t = A. \tag{21.9}$$

A legitimate question is now the following: since the equation is integrable, why the need to use other methods? The reasons are formal and economic. Formally we can note the following:

(1) the function $f(y)$ may be such that the integral $\int[1/f(y)]\mathrm{d}y$ cannot be expressed in terms of known functions. When this happens, we can try to expand the integrand in power series and to integrate term by term. This belongs to the domain of 'quantitative' approximations.

(2) even if difficulty (1) does not occur, so that

$$\int \frac{1}{f(y)}\mathrm{d}y = G(y),$$

where G is a known function, it may happen that the relation $t = H(y)$ (where $H(y) = G(y) - A$) is not invertible, and so we cannot obtain the explicit form of the solution, $y = y(t) \equiv H^{-1}(t)$, which is the one we are interested in to determine the time path of y.

From the economic point of view there is the usual motive: in economic theory, more often than not the form of the function $f(y)$ is not specified, but only its qualitative properties are given, so that only a qualitative analysis is possible. If we add the (formal) fact that it may happen that Eq. (21.6) cannot be put into the explicit form (21.7), we have more than enough to justify the great usefulness of the phase diagram.

Consider two orthogonal Cartesian axes; on the abscissae we measure y and on the ordinates we measure $\mathrm{d}y/\mathrm{d}t$, that is, $f(y)$ if (21.6) can be made

explicit. Then the graphical counterpart of Eq. (21.6) or (21.7) is a well-defined curve, which is the *phase diagram* of the equation. In Fig. (21.3) some of the ways in which $\varphi(y, \mathrm{d}y/\mathrm{d}t) = 0$ might qualitatively behave are represented.

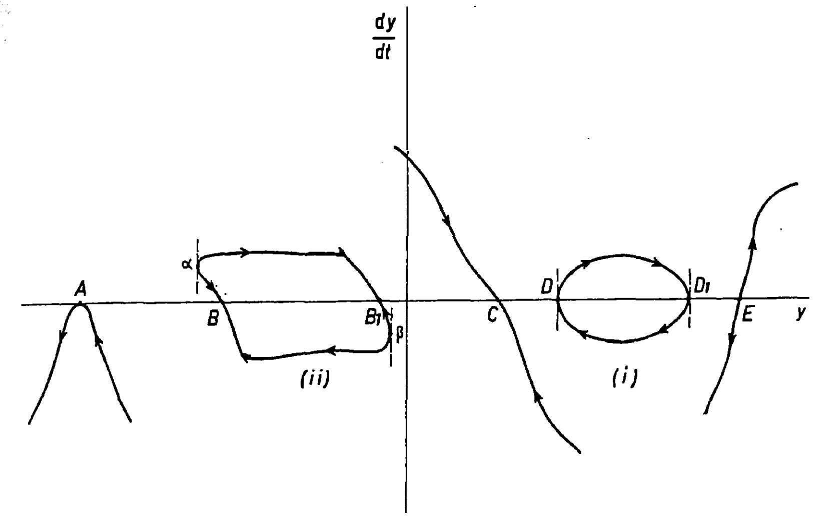


Figure 21.3: First-order non-linear differential equation: phase diagram

The basic rules for interpreting these diagrams are very simple:

i) *in all the points above the y-axis,* $\mathrm{d}y/\mathrm{d}t$ *is positive and so y is increasing;*

ii) *in all the points below the y-axis,* $\mathrm{d}y/\mathrm{d}t$ *is negative and so y is decreasing;*

iii) *in all the points falling on the y-axis,* $\mathrm{d}y/\mathrm{d}t$ is zero and so y *is stationary.*

The arrows indicate the direction of the movement: if $\mathrm{d}y/\mathrm{d}t > 0$, we move from the left to the right on the phase curve (y increases); if $\mathrm{d}y/\mathrm{d}t < 0$, we move from the right to the left along the phase curve (y decreases). An equilibrium point occurs where $\mathrm{d}y/\mathrm{d}t = 0$, that is, where the phase curve intersects (or is tangent to) the y-axis. Of course, there may be more than on such point.

It is then quite simple to analyse the stability of an equilibrium point: for example, point A is stable from the right but unstable from the left (*one-sided stability-instability*), point C is stable, point E is unstable. In other words, it is true, in general, that:

(a) if the phase curve intersects the y-axis with a positive slope, the equilibrium point (that is, the intersection point) is unstable (points like E);

(b) if the phase curve intersects the y-axis with a negative slope, the equilibrium point is stable (points like C);

(c) if the phase curve is tangential to the y-axis, remaining wholly on one side of it (that is, if the point of tangency is a maximum or a minimum of the phase curve), the equilibrium point is stable from one side and unstable from the other (points like A). If, instead, the point of tangency corresponds to a horizontal inflection point of the phase curve, then the equilibrium point is stable if the phase curve lies above (below) the y-axis to the left (right) of the inflection point and unstable in the opposite case.

More generally, global stability obtains if the phase curve lies wholly above (below) the y-axis to the left (right) of the equilibrium point. Apart from intuitive graphical considerations, this condition can be proved rigorously by means of Liapunov's second method (see Chap. 23). Let $V = \frac{1}{2}(y - y_e)^2$ be a Liapunov function. Then $\mathrm{d}V/\mathrm{d}t = (y - y_e) f(y)$, which is zero for $y = y_e$, and $\mathrm{d}V/\mathrm{d}t < 0$ if $f(y) \gtrless 0$ for $y - y_e \lessgtr 0$, which is the condition we have stated.

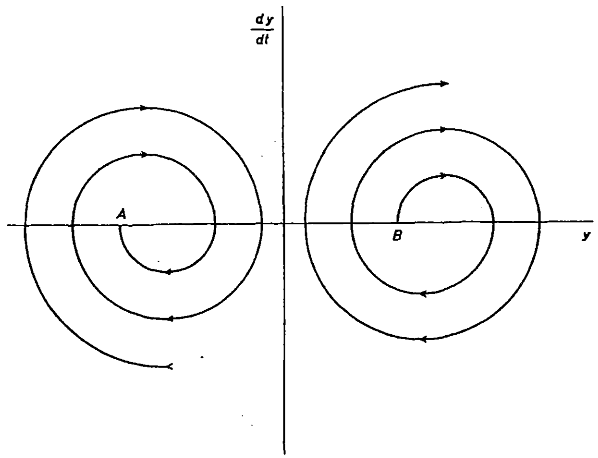


Figure 21.4: First-order non-linear differential equation: spiral-like phase diagram

An interesting case occurs when the phase curve forms a *closed loop*. Of course, this is possible only if the function $f(y)$ is multivalued. When the phase diagram has the form of a closed curve, an oscillatory movement (necessarily of constant amplitude) may occur. The conditions for the occurrence

of a cycle when the phase curve is closed are the following:

(1) that the curve lies partly above and partly below the y-axis, so that there can be a stage of increase and one of decrease in y;

(2) that the phase curve has an infinite slope (tangent parallel to the $\frac{\mathrm{d}y}{\mathrm{d}t}$-axis) at the points of intersection with the y-axis. In fact, if it were not so, such intersection points would be either stable or unstable according to the rules outlined above, under (a) and (b), so that the movement in the first case would converge monotonically to the equilibrium point and would not cross the y-axis, and in the second case it would diverge from the equilibrium point and would not cross the y-axis either.

In Fig. 21.3 the phase curve of type (i) satisfies both conditions, so that it gives rise to a constant-amplitude oscillatory movement. The phase curve of type (ii) does not satisfy condition (2), and the movement converges to point B or B_1 according to whether the initial point falls in the portion $\alpha B\beta$ or $\alpha B_1\beta$ of the phase curve.

It must be stressed that the oscillatory movements that may occur are not necessarily of constant amplitude. Oscillatory movements of the convergent or divergent kind may also occur with multivalued functions $f(y)$. Of course, the phase curve must not be a closed loop, but must have a spiral-like form, as in Fig. 21.4. In any case the same conditions (1) and (2) must hold. Spirals like A give rise to an oscillatory movement which converges to the equilibrium point A, whereas the opposite is true in the case of spirals like B.

What we have said in this section does not exhaust all possible cases. But a complete taxonomy would not be possible and anyway would not be of much interest, since once we have understood the fairly simple principles which underlie phase diagrams, we are sufficiently equipped to understand any such diagram that we may meet.

Let us note, finally, that the phase diagram can be applied also to first-order linear equations with constant coefficients (a particular case of the general first-order autonomous equation) of the type

$$\frac{\mathrm{d}y}{\mathrm{d}t} = ay + b,$$

examined in Part II, Chap. 12. The phase diagram of this equation is a straight line which intersects the y-axis at the point $-b/a$ and which has a positive (negative) slope if a is positive (negative).

21.3.2 Two-equation simultaneous systems

21.3.2.1 Introduction: phase plane and phase path

Here we shall examine autonomous system of the type

$$\begin{aligned}\frac{dy_1}{dt} &= \varphi_1(y_1, y_2),\\ \frac{dy_2}{dt} &= \varphi_2(y_1, y_2).\end{aligned} \tag{21.10}$$

The system is autonomous because it does not involve time explicitly. System (21.10) can be given as such or it may derive from a second-order equation. Consider, for example, the equation

$$\frac{d^2y_1}{dt^2} + f\left(y_1, \frac{dy_1}{dt}\right) = 0, \tag{21.11}$$

and define a new variable y_2 such that $dy_1/dt = y_2$. Then the system

$$\begin{aligned}\frac{dy_1}{dt} &= y_2,\\ \frac{dy_2}{dt} &= -f(y_1, y_2),\end{aligned} \tag{21.12}$$

is equivalent to Eq. (21.11) and is of type (21.10). The reason for transforming Eq. (21.11) into system (21.12) is that by so doing we can use the qualitative techniques that we are going to explain.

Let us now give some definitions. By *phase plane* (also called the *plane of the states*) we mean the plane (y_1, y_2). To each 'state' of the system, that is to each pair of coordinates y_1, y_2, there corresponds a point in the phase plane. Thus the motion of the system can be represented by a succession of points in the phase plane, that is a *phase path* or *trajectory* (also called *'characteristic curve'* or simply *'characteristic')*. The *phase path* must not be confused with the *phase diagram* treated in the previous section. The analogue of the phase path for a single equation would be the y-axis itself

The phase path, in other words, is mathematically described by the parametric equations

$$\begin{aligned}y_1 &= y_1(t),\\ y_2 &= y_2(t),\end{aligned} \tag{21.13}$$

where $y_1(t), y_2(t)$ is the solution of system (21.10). For the graphic construction of a phase path by means of (21.13), see Part II, Chap. 19, Sect. 19.3.

Eliminating t from (21.13) we have

$$\psi(y_1, y_2) = 0, \tag{21.14}$$

which is the implicit equation of the phase path.

The same trajectory can also be obtained in a different way. Eliminating dt from (21.10) we have the (timeless) differential equation

$$\frac{dy_2}{dy_1} = \frac{\varphi_2(y_1, y_2)}{\varphi_1(y_1, y_2)}, \tag{21.15}$$

whose solution yields the *integral curves* of system (21.10) in the phase plane, each curve corresponding to given values of the arbitrary constants. Now, such integral curves are the same[3] as the trajectory obtained by (21.13)—actually, the solution of (21.15) is the same as (21.14), apart from arbitrary constants—with the difference that from (21.13) we also obtain the *direction of the movement* of the representative point on the phase path, whereas from the integration of (21.15) we obtain a purely geometric curve, without any reference to what is happening in time. However, the 'direction of travel' along the integral curves can be obtained from Eq. (21.10): observing that $dy_i/dt \gtrless 0$ when $\varphi_i(y_1, y_2) \gtrless 0$, we can immediately tell whether y_i is increasing or decreasing. Here, as well as in the remaining treatment (unless otherwise explicitly stated), the direction of the movement is to be understood for $t \to +\infty$.

The practical reason for obtaining the integral curves is that it may happen that Eq. (21.15) is integrable[4] whereas Eqs. (21.13) are not obtainable from the solution of system (21.10), either because such a solution cannot be found (and this is a very common case with non-linear systems) or because it can be given only in the form of an integral curve. Even if Eq. (21.15) is not integrable, a graphic analysis of it to obtain the approximate shape of the integral curves may be simpler than the analysis of system (21.10).

21.3.2.2 Singular points

Any point in which the two functions $\varphi_1(y_1, y_2)$ and $\varphi_2(y_1, y_2)$ do not vanish simultaneously is called an *ordinary point* (or regular point), whereas any point in which the two functions are simultaneously zero is called a *singular point*. It is clear that singular points represent states of rest, that is, of equilibrium of the system, since in a singular point $dy_1/dt = dy_2/dt = 0$ by definition. Thus the analysis of the integral curves usually starts by trying to determine the singular points, if any, around which the integral

[3]Some authors, e.g., Andronov et al. (1966, pp. 7, 34), prefer to distinguish between *phase path* and *integral curves*, since an integral curve may consist of several phase paths. Consider, for example, Fig. 21.5a below. Each integral curve consists of three phase paths, two of which are two branches of the typical integral curve, and the third is the equilibrium point itself. It is, however, true that the integral curves are the loci of all the phase paths.

[4]By 'integrable' we mean that the integral can be explicitly expressed in terms of (a finite number of) known functions. Thus by 'not integrable' we do not mean that the integral does not exist, but that it cannot be expressed, etcetera.

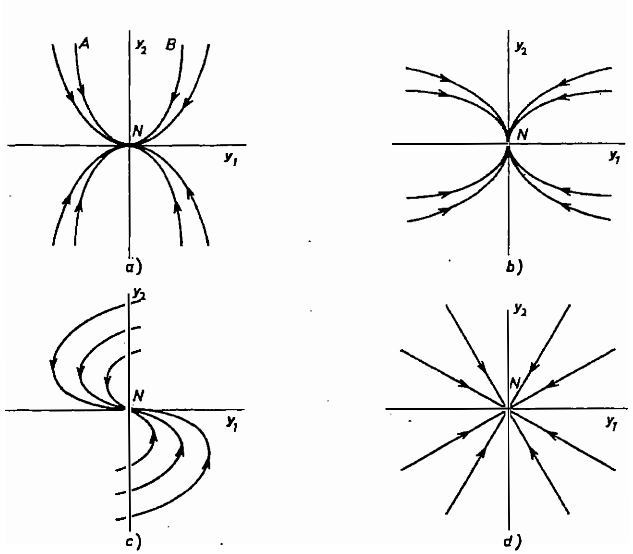


Figure 21.5: Singular points of 2×2 systems: the node

curves are then constructed. In what follows, for graphical simplicity we shall take the origin as the singular point. This does not involve any loss of generality, since if the singular point is elsewhere, we can always imagine making a transformation of coordinates so that the origin and the singular point coincide; if there are several singular points, one of them is taken as the origin.

The 'elementary' singular points are the *node*, the *saddle point* (or *col),* the *focus* (or *vortex),* and the *centre,* the classification being made according to the shape of the integral curves.

A *node* is a singular point such that all integral curves pass through it. Some examples are given in Fig. 21.5. In the particular case in which the integral curves are straight lines passing through the singular point the node is called a stellar node or *star.*

If the direction of the movement (indicated by the arrows) is towards (away from) the singular point, the latter is, of course, stable (unstable).

A *saddle point* is a singular point through which only two integral curves pass, which are asymptotes to all remaining curves. The asymptotes can be the axes themselves, as in Fig. 21.6a, or not, as in Fig. 21.6b. It can be seen that whatever the direction of the movement along an integral curve, the motion is always away from the equilibrium point *except for the one along*

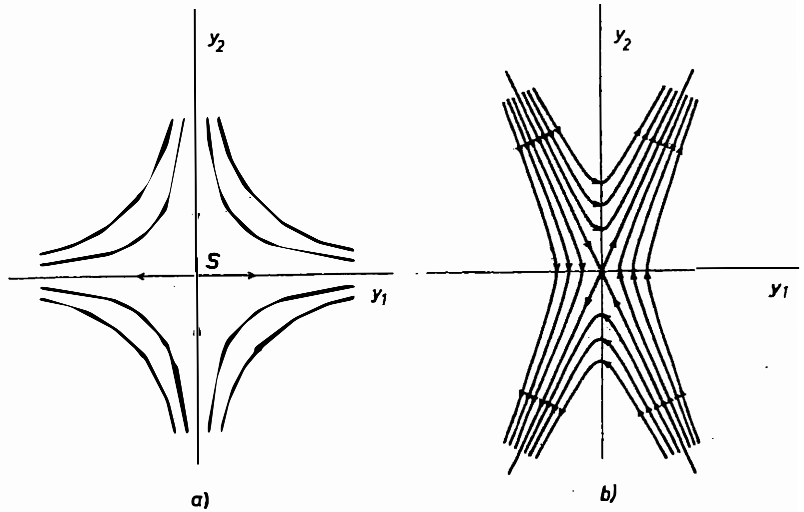


Figure 21.6: Singular points of 2×2 systems: the saddle point

one of the two asymptotes, which is called the stable arm of the saddle.

A *focus* is a singular point which is the limit point of all integral curves, which have the form of spirals enclosed in each other (Fig. 21.7). The direction of the movement can be toward the equilibrium point (stable focus or *spiral sink)* or away from it (unstable focus or *source*).

Finally, a *centre* (also called *focal point*) is a singular point through which no integral curve passes (that is, an isolated singular point) surrounded by closed integral curves (Fig. 21.8). In this case, no matter what the direction of the movement is, the variables y_1, y_2 have a periodic motion whose amplitude is neither decreasing nor increasing, whereas in the case of a focus they have an oscillatory movement whose amplitude is either decreasing or increasing.

In general, a system may have several singular points of different types. Thus, if we 'map' the whole phase plane, we obtain certain domains possessing different properties. The boundaries of these domains are certain asymptotic trajectories called *separatrices* (see e.g. Fig. 21.9). Figure 21.9 also illustrates the concept of *saddle loops* or *homoclinic orbits* Consider the origin, which is clearly a saddle point, and the two equilibrium points V_1, V_2. The two loops that appear when the trajectories leave the saddle point in either direction on the unstable arm and return to it on the stable arm, are called homoclinic orbits. More generally, a homoclinic orbit is a closed curve

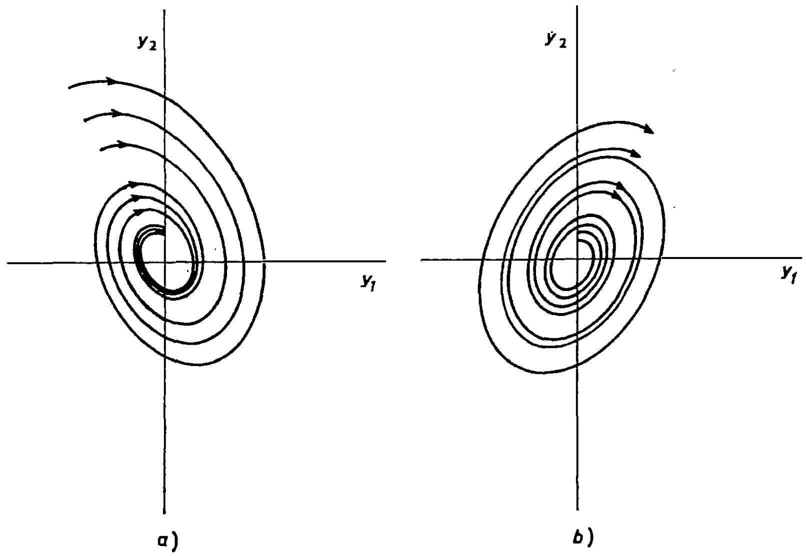


Figure 21.7: Singular points of 2×2 systems: the focus

that connects the same fixed point to itself, and, by contrast, a *heteroclinic orbit* arises when connecting distinct fixed points. An example of heteroclinic orbit is the separatrix connecting the two extreme saddle points in Fig. 21.9. If the stable and unstable manifold originating from a saddle point intersect transversely (i.e., not tangentially) at another point p, this point (not shown in the diagram) is called a transversal homoclinic point, and the backward and forward orbit of p is called a *transversal homoclinic orbit.*

It may also happen that when the system involves a parameter, the passage of the latter through a critical value causes a qualitative change in the nature of the singular point(s) and of the trajectories. The value(s) of the parameter at which such a change occurs is (are) called *bifurcation* (or *branch) value(s)*: see, e.g., Andronov et al. (1966, Chap. II, § 5); Minorsky (1962, Chap. 7). Bifurcation theory will be treated at some length in Chap. 25.

21.3.2.3 Graphical construction of the trajectories

Let us now show how to examine the trajectories of system (21.10) when the functions φ_1 and φ_2 are specified only qualitatively. First of all, the curves $\varphi_1(y_1, y_2) = 0$ and $\varphi_2(y_1, y_2) = 0$ must be studied; their intersection in the (y_1, y_2) plane, if it exists, determines a singular point of the system. These curves are sometimes called *isokines*. More generally, given a differential

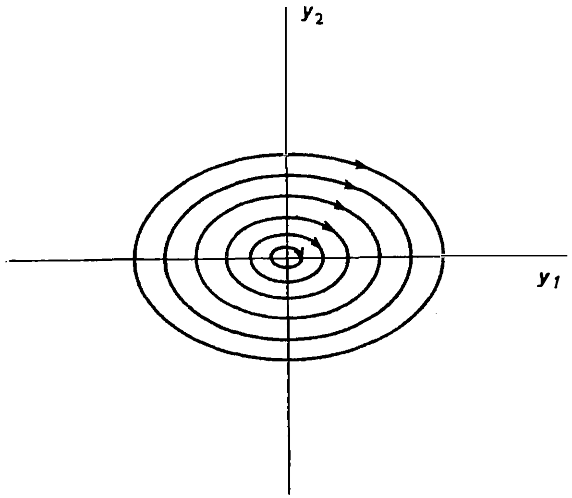


Figure 21.8: Singular points of 2×2 systems: the centre

equation $\mathrm{d}y_i/\mathrm{d}t = \varphi_i(y_1, y_2)$, an isokine is a curve in the (y_1, y_2) plane that yields the same value of $\mathrm{d}y_i/\mathrm{d}t$, i.e. a contour of the surface $z_i = \varphi_i(y_1, y_2)$ for z_i a given constant.

Of course, more than one intersection may occur; for simplicity we assume that only one exists. It is then possible to determine whether these curves are increasing or decreasing by computing $(\mathrm{d}y_2/\mathrm{d}y_1)_{\varphi_i=0}$. This is easy if y_2 can be expressed as an explicit function of y_1 (or vice versa) from $\varphi_i = 0$. If not, we can use the theorem on the differentiation of implicit functions, which gives

$$\left(\frac{\mathrm{d}y_2}{\mathrm{d}y_1}\right)_{\varphi_1=0} = -\frac{\dfrac{\partial\varphi_1}{\partial y_1}}{\dfrac{\partial\varphi_1}{\partial y_2}},$$

$$\left(\frac{\mathrm{d}y_2}{\mathrm{d}y_1}\right)_{\varphi_2=0} = -\frac{\dfrac{\partial\varphi_2}{\partial y_1}}{\dfrac{\partial\varphi_2}{\partial y_2}}.$$

The signs of the partial derivatives $\partial\varphi_i/\partial y_j$ are usually known from the underlying economic assumptions, therefore the increasing or decreasing behaviour of the two curves can be ascertained immediately.

The next step is to examine the nature of the points not lying on the

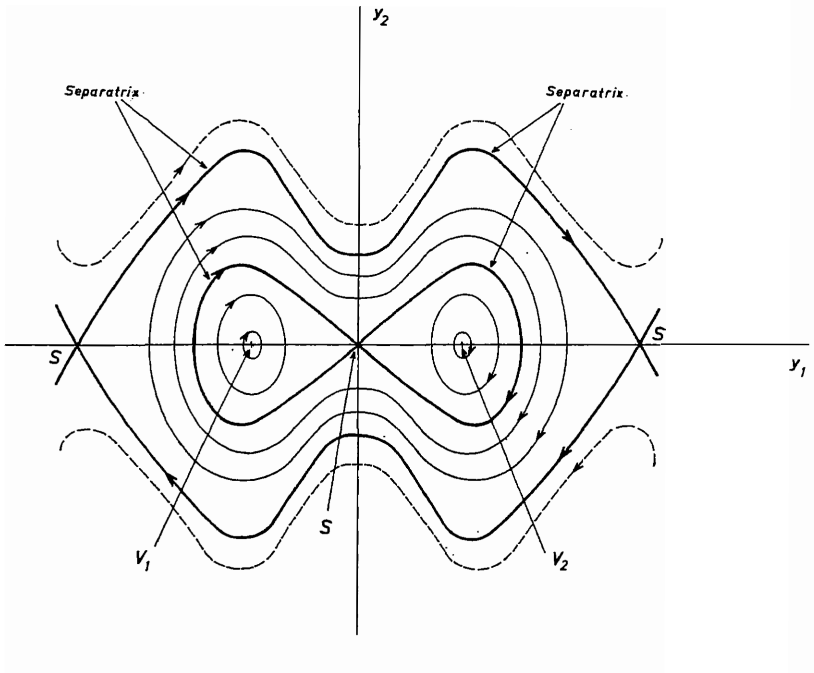


Figure 21.9: Multiple singular points of 2×2 systems

curve $\varphi_1 = 0$. This is important because, given system (21.10), $\mathrm{d}y/\mathrm{d}t \gtrless 0$ as $\varphi_1 \gtrless 0$, i.e. y_1 is increasing (decreasing) in the points where φ_1 is positive (negative). Similarly for the points not lying on the curve $\varphi_2 = 0$.

Such an examination can be made using the knowledge of the signs of the partial derivatives $\partial\varphi_i/\partial y_j$.

To illustrate the procedure, suppose that $\partial\varphi_1/\partial y_1 < 0, \partial\varphi_1/\partial y_2 < 0$, so that $\varphi_1 = 0$ is a decreasing curve (see Fig. 21.10). Take an arbitrary point P on the $\varphi_1 = 0$ curve, and consider a point P' lying vertically above P in Fig. 21.10. In P', y_1 is the same as in P, and y_2 is greater than in P, so that $\mathrm{d}\varphi_1 = (\partial\varphi_1/\partial y_2)\mathrm{d}y_2 < 0$; it follows that the value of φ_1 in P' is smaller than in P. Therefore, since $\varphi_1 = 0$ in P, it must be $\varphi_1 < 0$ in P'. Since the reasoning can be repeated for all points above the curve $\varphi_1 = 0$, we have that $\varphi_1 < 0$ in all such points. Similarly it can be shown that $\varphi_1 > 0$ in all the points below the curve $\varphi_1 = 0$.

Note that instead of keeping y_1 constant and varying y_2 we could have kept y_2 constant and varied y_1 (point P''): as $\mathrm{d}\varphi_1 = (\partial\varphi_1/\partial y_1)\mathrm{d}y_1 < 0$, in P'' we have $\varphi_1 < 0$ and the same conclusion as before is reached ($\varphi_1 \lessgtr 0$

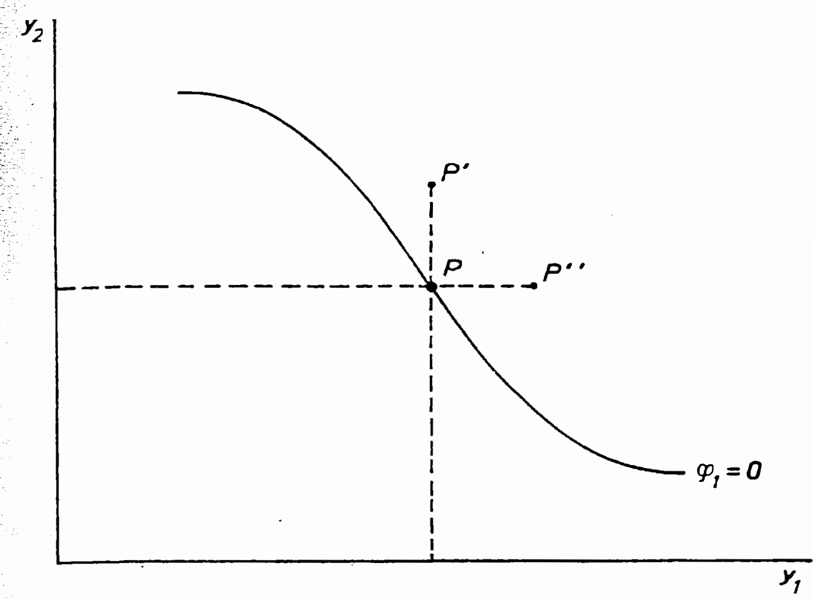


Figure 21.10: Nature of the points not lying on $\varphi_1 = 0$

respectively above and below the curve $\varphi_1 = 0$). A similar procedure can be followed when the signs of $\partial\varphi_1/\partial y_1$ and $\partial\varphi_1/\partial y_2$ are different from those assumed above.

In a similar way we can analyse the $\varphi_2(y_1, y_2) = 0$ curve and establish the regions of the plane where $\varphi_2 \gtrless 0$.

Let us now consider an arbitrary point different from P, such as A in Fig. 21.11. It is then possible to draw two 'arrows' which indicate the forces acting on the system: as A lies below the $\varphi_1 = 0$ curve, we have $\varphi_1 > 0$ and so $\mathrm{d}y_1/\mathrm{d}t > 0$, i.e. y_1 tends to increase (horizontal arrow pointing rightwards). Assuming that $\partial\varphi_2/\partial y_1 < 0, \partial\varphi_2/\partial y_2 > 0$, the $\varphi_2 = 0$ curve is increasing and $\varphi_2 \lessgtr 0$ respectively below (above) it. Therefore in A we have $\varphi_2 > 0$ and so y_2 tends to increase (vertical arrow pointing upwards). Thus, point A will tend to move in a direction included between the two arrows, describing the trajectory of the system. In a similar way we can proceed for any other point different from A. Now, both arrows from A 'point' toward equilibrium, so that it would be tempting to draw the conclusion that the singular point E is stable (a stable node, as indicated by the phase path from A to E). But this would not be legitimate in the absence of further investigations: in fact, point E could also happen to be a saddle point (and so unstable), as

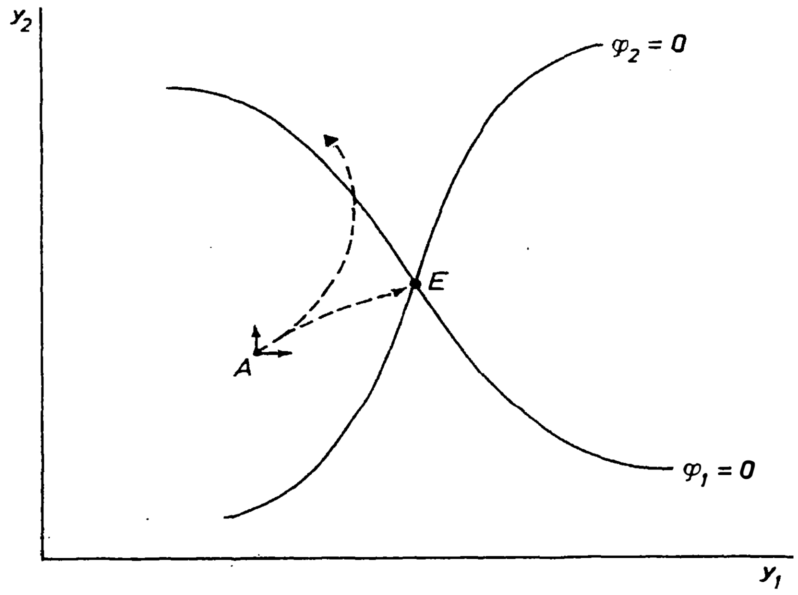


Figure 21.11: Graphical determination of the nature of a singular point

indicated by the other phase path starting from A.

If the four quadrants in which the $\varphi_1 = 0$ and $\varphi_2 = 0$ curve divide the plane are numbered from I to IV, e.g. clockwise and starting from the one where A lies, it can be checked graphically that the stable asymptote passes through E increasing from quadrant I to quadrant III whereas the unstable asymptote passes through E decreasing from quadrant II to quadrant IV. Actually, it is easy to prove that point E is a saddle point in the case under examination (see exercise 2).

The following theorem provides useful *sufficient* conditions for the global stability of point E (Olech, 1963).

Olech's theorem. Consider the autonomous system

$$\frac{dy_i}{dt} = \varphi_i(y_1, y_2), \qquad i = 1, 2,$$

where the functions φ_i have continuous first-order partial derivatives, and

$\varphi_i = 0$ for $y_i = y_i^e$, $i = 1, 2$. Suppose that

$$\text{(i)} \quad \frac{\partial \varphi_1}{\partial y_1} + \frac{\partial \varphi_2}{\partial y_2} < 0 \qquad \text{everywhere,}$$

and

$$\text{(ii)} \quad \frac{\partial \varphi_1}{\partial y_1}\frac{\partial \varphi_2}{\partial y_2} - \frac{\partial \varphi_1}{\partial y_2}\frac{\partial \varphi_2}{\partial y_1} > 0 \quad \text{everywhere.}$$

Furthermore, assume that either

$$\text{(iii)} \quad \frac{\partial \varphi_1}{\partial y_1}\frac{\partial \varphi_2}{\partial y_2} \neq 0 \quad \text{everywhere,}$$

or

$$\text{(iii')} \quad \frac{\partial \varphi_1}{\partial y_2}\frac{\partial \varphi_2}{\partial y_1} \neq 0 \quad \text{everywhere.}$$

Then the equilibrium state (y_1^e, y_2^e) is asymptotically stable in the large.

We have seen above that the fact that the arrows point towards equilibrium may be misleading. Also the opposite case may occur, in the sense that when the arrows point away from equilibrium, this does not necessarily imply that equilibrium is unstable, because it might be a centre (see Fig. 21.12). All this confirms what we said elsewhere (Part II, Chap. 19, Sect. 19.3) about the insufficiency of the arrows *alone* to analyse a stability problem.

The notion of centre defined above leads us to investigate the problem of the existence and stability of *limit cycles* (Fig. 21.13).

A limit cycle is an isolated closed integral curve to which all nearby paths approach from both sides in a spiral fashion. If the direction of the movement along the nearby path is towards (away from) the limit cycle, the latter is called *orbitally stable (unstable).*

Intermediate cases may also occur in which all paths on one side of the limit cycle approach it, while on the other side they move away from it; in such cases the limit cycle can be called *semi-stable.* The (local) stability of a limit cycle can be investigated by means of the following theorem:

Orbital stability theorem. Let $y_1 = y_1(t), y_2 = y_2(t)$ be a periodic motion (limit cycle) of system (21.10). This limit cycle is locally stable (unstable) if its *characteristic exponent*

$$h = \frac{1}{T}\int_0^T \left\{ \frac{\partial \varphi_1 [y_1(t), y_2(t)]}{\partial y_1} + \frac{\partial \varphi_2 [y_1(t), y_2(t)]}{\partial y_2} \right\} dt,$$

where T is the period of the oscillation, is respectively negative (positive). For a proof the reader is referred, for example, to Andronov et al., 1966, pp.

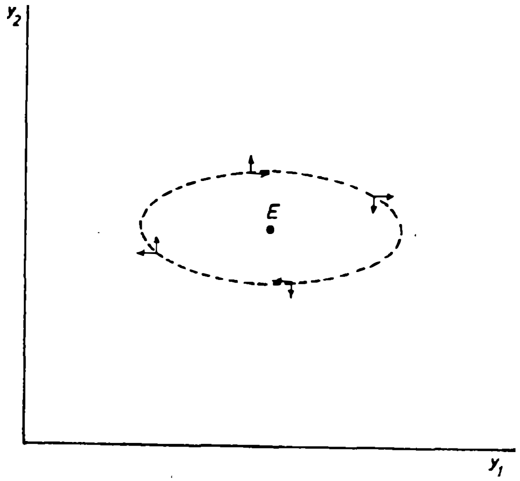


Figure 21.12: Arrows and centres

289-90 and 296-300. Further study of limit cycles is contained in Chap. 24, Sect. 24.3.

21.3.2.4 Linear systems

The graphical technique of phase diagrams and integral curves can be applied also to linear systems with constant coefficients. This application, though interesting in itself, does not add anything to our knowledge of the solution of such systems, since the latter can always be obtained explicitly. Moreover, it may happen that the difficulty of the problem is increased. For example, consider the system

$$\begin{aligned} \frac{dy_1}{dt} &= y_2, \\ \frac{dy_2}{dt} &= -\omega^2 y_1 - 2by_2, \end{aligned}$$

which is easily integrable by the standard procedure explained in Part II, Chap. 18 (the characteristic equation is $\lambda^2 + 2b\lambda + \omega^2 = 0$, etc.). Eliminating

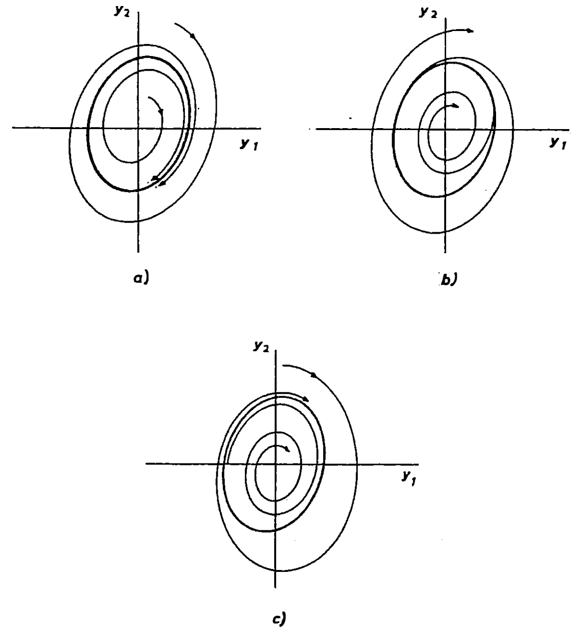


Figure 21.13: Limit cycles

dt we obtain the differential equation of the integral curves

$$\frac{dy_2}{dy_1} = -\frac{2by_2 + \omega^2 y_1}{y_2},$$

which is rather more difficult to integrate than the original system. Thus the technique of the integral curves is most useful in the analysis of non-linear systems which cannot be integrated or whose solution cannot be put in explicit form, as we said above (an illustration will be given in the next section). However, in economic applications this technique is being increasingly used also in the case of linear system with constant coefficients. In fact, a common procedure is the following: given a system of the type of (21.10), after calculating its singular point(s) a linear approximation is performed, and the resulting linear system is used to construct the integral curves. Although this procedure may be misleading if one does not keep in mind that the integral curves so obtained have a *local* validity only, we shall state here the relationship between the characteristic equation of a 2 × 2 linear differential system with constant coefficients and the singular points of the same system

Table 21.1: Singular points of the linear system $y_1' = a_{11}y_1 + a_{12}y_2$, $y_2' = a_{21}y_1 + a_{22}y_2$
$[\Delta = (a_{11} + a_{22})^2 - 4(a_{11}a_{22} - a_{12}a_{21})]$

Δ	Determinant, or trace	Roots	Singular point	Stability	Other
$\Delta > 0$	$a_{11}a_{22} - a_{12}a_{21} > 0$	real, same sign	*node*	stable if $a_{11} + a_{22} < 0$; unstable if $a_{11} + a_{22} > 0$	
	$a_{11}a_{22} - a_{12}a_{21} < 0$	real, opposite sign	*saddle point*	conditional stability (depends on initial position)	
$\Delta < 0$	$a_{11} + a_{22} \neq 0$	complex conjugate	*focus*	stable if $a_{11} + a_{22} < 0$; unstable if $a_{11} + a_{22} > 0$	
	$a_{11} + a_{22} = 0$	pure imaginary	*centre*		
$\Delta = 0$		real and equal	*node*	stable if $a_{11} + a_{22} < 0$; unstable if $a_{11} + a_{22} > 0$	*star* only if $a_{11} = a_{22}$ and $a_{12} = a_{21} = 0$

(for proofs see, e.g., Sansone and Conti, 1964, Chap. II). Given the system

$$\frac{dy_1}{dt} = a_{11}y_1 + a_{12}y_2,$$

$$\frac{dy_2}{dt} = a_{21}y_1 + a_{22}y_2,$$

the characteristic equation of which is $\lambda^2 - (a_{11} + a_{22})\lambda + (a_{11}a_{22} - a_{12}a_{21}) = 0$, we have the various cases laid out in Table 21.1.

In the particular case in which $a_{11}a_{22} - a_{12}a_{21} = 0$, the system reduces to $dy_2/dt = k(dy_1/dt)$, i.e. $dy_2/dy_1 = k$, and the integral curves are straight lines which no longer possess a singularity at the origin.

Let us note that, in the case of a saddle point, the stable arm of the saddle corresponds to the negative latent root, say λ_1, along which the movement will be given by

$$y_1(t) = A_1 e^{\lambda_1 t}, \quad y_2(t) = A_1 \alpha_2^{(1)} e^{\lambda_1 t}. \tag{21.16}$$

This immediately yields the equation of the stable arm in the (y_1, y_2) phase plane, namely

$$y_2 = \alpha_2^{(1)} y_1, \quad y_1 = y_2/\alpha_2^{(1)}, \tag{21.17}$$

where $\alpha_2^{(1)} = (\lambda_1 - a_{11})/a_{12} = a_{21}/(\lambda_1 - a_{22})$ as shown in Chap. 18, Sect. 18.1.2.

If, at the initial time, the initial state of the system happens to be on the stable arm, then the system will move along it toward equilibrium. This is why we have classified the saddle point as conditionally stable in Table 21.1. Such an initial state of the system must satisfy Eq. (21.17), namely

$$y_2(0) = \alpha_2^{(1)} y_1(0), \tag{21.18}$$

where $y_1(0), y_2(0)$ are the initial conditions. For arbitrarily given initial conditions, Eq. (21.18) can be satisfied only by chance. If we accept the classical definition of probability, this event, although not impossible, has probability approaching zero, as the number of favourable cases is given by the number of points lying on the line (21.18) (this number has the dimension ∞), while the possible cases are all the points lying in the y_1, y_2 plane (∞^2).

However, if we are somehow free to *choose* one of the initial conditions—it doesn't matter whether $y_1(0)$ or $y_2(0)$—we can always put the system on the stable arm of the saddle by choosing the 'free' initial condition so as to satisfy Eq. (21.18). The problem of making stable a not-wholly-unstable system has been dealt with in general in Chap. 18, Sect. 18.2.2.3, and its economic applications will be the subject of Chap. 22.

21.4 Quantitative methods

Although there are many types of non-linear differential equations whose closed-form solution is known, only a very limited number of them have had some application in economics. These types will be treated in Chap. 24. Actually, in economic theory we are willing to assume a precise functional form (such a Cobb-Douglas production function) just in a few cases, which then give rise to precise types of differential equations (the typical differential equation arising from the Cobb-Douglas functional form is a Bernoulli equation). Usually, however, we do not know the functional form of $\mathbf{f}(\mathbf{y},t)$ but at most its qualitative properties, such as the signs of the partial derivatives. Hence either we use qualitative methods or, if we are content with local stability, we perform a linear approximation.

Thus the paramount quantitative tool used in the study of the stability of economic models remains linearisation. Hence the precise conditions for its proper use should be part of every economist's tool-kit: in fact, it is regrettable that this (apparently simple) method is often used rather sloppishly.

21.4.1 Linearisation

The usual procedure to examine the local stability of system (21.1) is to linearise, namely to expand the functions $\mathbf{f}$ in Taylor's series about the equilibrium solution $\mathbf{y} = \mathbf{y}_e$ and neglect all higher-order terms. Examination of the stability properties of the resulting linear system is then extended to the local stability properties of the original non-linear system. This is the sloppy approach. The rigorous approach is the following. Consider the generic non-linear, non-autonomous system (21.1), that we reproduce here for the reader's convenience,

$$\mathbf{y}' = \mathbf{f}(\mathbf{y},t),$$

and let $\mathbf{y}^*(t)$ be a solution of the system. To study the properties of the system in the neighbourhood of the solution $\mathbf{y}^*(t)$ we introduce the variables $\overline{y}_i(t) = y_i(t) - y_i^*(t)$, that is

$$\overline{\mathbf{y}}(t) = \mathbf{y}(t) - \mathbf{y}^*(t), \tag{21.19}$$

and call $\overline{\mathbf{y}}(t) = \mathbf{0}$, or $\mathbf{y}(t) = \mathbf{y}^*(t)$, the *unperturbed motion* or null solution (Krasovskii, 1963, p. 1). Since the variables $y_i(t)$ satisfy system (21.1), the new variables $\overline{y}_i(t)$ satisfy the system

$$\overline{\mathbf{y}}' = \mathbf{g}(\overline{\mathbf{y}},t), \tag{21.20}$$

where

$$\mathbf{g}(\overline{\mathbf{y}},t) \equiv \mathbf{f}(\overline{\mathbf{y}}+\mathbf{y}^*,t) - \mathbf{f}(\mathbf{y}^*,t). \tag{21.21}$$

The validity of Eqs. (21.20)—called the *equations of perturbed motion*—can be checked by direct substitutions, remembering that $\mathbf{y}^*$ is a solution to system (21.1) and so satisfies it by definition. It should be noted that in the case in which $\mathbf{y}^*(t)$ is the equilibrium state of the system, i.e. $\mathbf{y}^*(t) = \mathbf{y}_e(t)$, Eqs. (21.21), given Eq. (21.3), reduce to

$$\mathbf{g}(\overline{\mathbf{y}},t) \equiv \mathbf{f}(\overline{\mathbf{y}}+\mathbf{y}_e,t). \tag{21.22}$$

The next step to examine the local stability of system (21.1) is to expand the right-hand side of system (21.20) in Taylor's series about the null solution. We thus obtain the system

$$\overline{\mathbf{y}}' = \mathbf{A}(t)\overline{\mathbf{y}} + \mathbf{h}(\overline{\mathbf{y}},t), \tag{21.23}$$

where $\mathbf{A}(t)$ is the Jacobian matrix of the functions $\mathbf{g}$ evaluated at $\overline{\mathbf{y}} = \mathbf{0}$ [which coincides with the Jacobian of $\mathbf{f}$ when $\mathbf{y}^*(t) = \mathbf{y}_e(t)$] and $\mathbf{h}(\overline{\mathbf{y}},t)$ represents all higher-order terms in Taylor's series.

Consider now the linear part of system (21.23), namely

$$\overline{\mathbf{y}}' = \mathbf{A}(t)\overline{\mathbf{y}}. \tag{21.24}$$

System (21.24) is called a *uniformly good approximation* to the original non-linear system if

$$\|\mathbf{h}(\overline{\mathbf{y}},t)\| \,/\, \|\overline{\mathbf{y}}\| \to 0 \quad \text{uniformly in } t \text{ with } \|\overline{\mathbf{y}}\| \to 0. \tag{21.25}$$

(the notion of uniform convergence has been defined above, see Definition 21.4).

We now come to the gist of the matter, which is the following

Poincaré-Liapunov-Perron Theorem. If the linearised system (21.24) near equilibrium is uniformly asymptotically stable and is a uniformly good approximation to the original non-linear system near $\mathbf{y}_e$, then the original non-linear system is locally uniformly asymptotically stable at $\mathbf{y}_e$.

This fundamental theorem on the validity of the linearisation procedure is attributed to Liapunov and Poincaré (for example by Bellman, 1953, p. 93; and by Sansone and Conti, 1964, Chap. IX, § 3.1) and to Perron (for example by Coddington and Levinson, 1955, p. 314). A joint attribution seems preferable because it seems that all three authors have proven the theorem independently.

Let us now come to practical matters. It is easy to see that *if the original non-linear system is autonomous, then the linearised system certainly is a uniformly good approximation, and, moreover, the matrix* $\mathbf{A}$ *is a matrix of constants.*

In fact, if the original non-linear system is autonomous, then:

(i) the function $\mathbf{h}$ in its Taylor's series expansion (21.23) does not contain t, so that the linearised system (21.24) is always a uniformly good approximation. This is due to the fact that, when $\mathbf{h}$ does not contain time explicitly, the problem of *uniform* convergence does not arise, and it is certain that $\|\mathbf{h}(\bar{\mathbf{y}})\| / \|\bar{\mathbf{y}}\|$ converges to zero as $\|\bar{\mathbf{y}}\|$ tends to zero because $\|\mathbf{h}(\bar{\mathbf{y}})\|$ represents all higher-order terms, which are of a lower order of smallness than $\|\bar{\mathbf{y}}\|$ as $\|\bar{\mathbf{y}}\|$ tends to zero.

(ii) the matrix $\mathbf{A}$ is obviously independent of t, since the Jacobian of a vector-valued function depends on, and only on, the independent variables present in the function.

The local stability analysis of non-linear autonomous systems by means of the usual linearisation is then always correct, but when the system is non-autonomous we must be careful to check that the conditions of the Poincaré-Liapunov-Perron theorem are satisfied.

We know that the behaviour of linear system with constant coefficients depends on the latent roots of the system's matrix (see above, Sect. 21.2, Properties L.1-L.4). When there are roots with zero real part, however, problems may arise. More precisely, the **Hartman-Grobman Theorem** states that the local properties of the linearised system carry over to the non-linear system if the linear system's matrix has no root with zero real part (zero real root or pure imaginary root) (Guckenheimer and Holmes, 1986, p. 13; see also below, Exercise 1). An equilibrium point such that the matrix of the linearised system has no root with zero real part is called a *nondegenerate* (or *hyperbolic)* fixed point.

This theorem is sometimes misinterpreted as stating that the system's matrix must not have any root with zero real part if we want to ascertain the properties of the non-linear system from those of the linear approximation. This is certainly true as regards local asymptotic *stability,* but not as regards *instability*. The presence of at least one root with positive real part is sufficient to state that the non-linear system is locally unstable, even if there are other roots with zero real part.

Thus we can sum up our treatment in the following (Braun, 1986, Chap. 4, Sect. 4.3):

Local stability theorem If the linear approximating system (21.24) is a uniformly good approximation to the original non-linear system (21.1), and the linear system's matrix $\mathbf{A}$ is constant, then:

(a) the equilibrium solution is locally asymptotically stable if all the latent roots of $\mathbf{A}$ have negative real part;

(b) the equilibrium solution is unstable if at least one latent root of $\mathbf{A}$ has positive real part;

(c) the stability of the equilibrium solution cannot be determined from the properties of the linear approximating system if all the latent roots of $\mathbf{A}$ have real part ≤ 0 but at least one root of $\mathbf{A}$ has zero real part.

It should again be stressed that these are *local* properties, and it would be completely unwarranted to extend these properties to the global properties of the original non-linear system. In particular, this latter might be locally unstable but converge, for example, to a closed orbit (orbital stability), or be stable for small displacements but unstable for displacements greater than a critical value, etcetera.

We conclude by observing that the above theorems Poincaré-Liapunov-Perron theorem can equally well be applied to non-linear systems that do not need linearisation, namely to systems composed of a linear plus a non-linear part. These systems are in the form (21.23) from the beginning.

21.5 Elements of the qualitative theory of difference equations

In the previous sections we have examined solely differential equations. We now turn to difference equations.

The concept of 'qualitative' or 'topological' theory of difference equations is the same as for differential equations. We shall only add that bifurcation theory for difference equations will be examined in Chap. 24, together with that for differential equations.

21.5.1 Single difference equations

In this section we shall examine the first-order non-linear autonomous difference equation, that is the equation of the type

$$\varphi\left(y_{t+1}, y_t\right) = 0. \tag{21.26}$$

The graphical analysis of Eq. (21.26) can be made by means of a diagram in which y_t is measured on the horizontal and y_{t+1} on the vertical axis. We shall call it the *phase diagram* by analogy with the similar diagram used for differential equations (Sect. 21.3.1). The analogy, however, must be taken with caution. For example, the analogue of dy/dt is Δy_t and not y_{t+1}. If Δy_t instead of y_{t+1} were used, the horizontal axis would become what is now 45° line. In other words, the 45° line plays the role of the horizontal axis used in differential equations (points above the 45° line represent points of increasing y, etc.).

It may happen that Eq. (21.26) can be made explicit as

$$y_{t+1} = f\left(y_t\right), \tag{21.27}$$

but, if this is not possible, the graphical analysis can be made equally well, using (21.26). In Fig. 21.14 we have a possible form of the phase line. The 45° line is useful given its property of being the locus of all points such that the abscissa and the ordinate are equal. We consider only the first

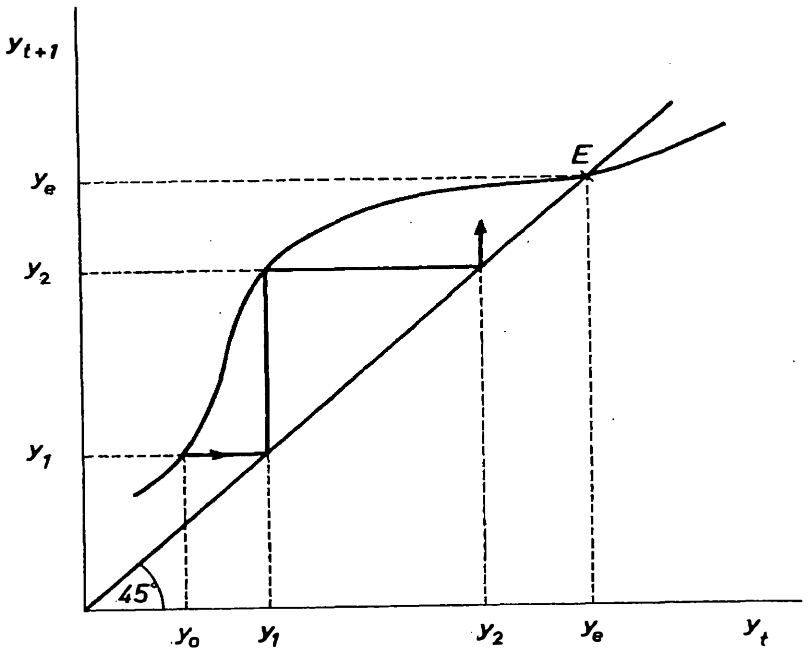


Figure 21.14: Phase diagram for a first-order difference equation: convergent monotonic movement

quadrant, since the variables which appear in economic applications must be non-negative.

Suppose that the arbitrary initial point is y_0. The successive value y_1 is obtained finding the point on the phase line having the abscissa y_0. By means of the 45° line we transfer graphically y_1 onto the horizontal axis and then we obtain y_2 as the ordinate of the point on the phase line corresponding to the abscissa y_1; we transfer y_2 onto the horizontal axis by means of the 45° line, and so on. The time path of y is given by the succession of the values $y_0, y_1, y_2, \ldots$ so obtained. Let us note, incidentally, that the path is monotonic and convergent towards the equilibrium point E (the same result would be obtained if the initial point were to the right of E). An equilibrium point is by definition a point such that y is stationary over time, that is, a point such that $y_{t+1} = y_t$. Therefore *the equilibrium points are the points of contact of the phase line with the* 45° *line*. In other words, given (21.27), equilibrium points are the fixed points of the mapping $y \rightarrow f(y)$.

We shall now expound some general properties of the phase diagrams

under consideration.

Case (1): monotonic movements

Monotonic movements arise when the phase line is a monotonically increasing function; if it is everywhere or in some stretches decreasing, oscillating movements occur (see the next case (2)). The movement may be convergent (as in Fig. 21.14) or divergent (as in Fig. 21.15). It can be noted

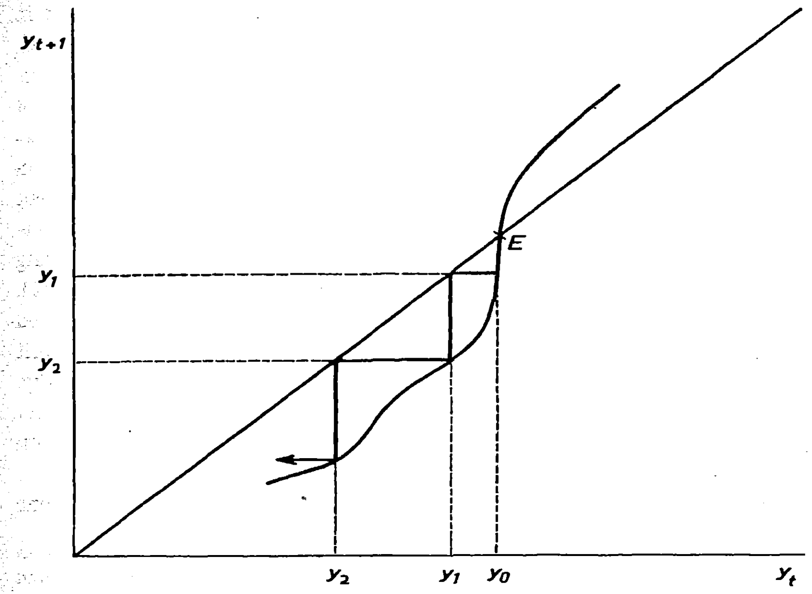


Figure 21.15: Phase diagram for a first-order difference equation: divergent monotonic movement

that in the first case the phase line is such that lies *above* the 45° line to the *left* of, and *below* the 45° line to the *right* of, the equilibrium point, whereas in the second case exactly the opposite is true. Thus it is the *position* of the phase line with respect to the 45° line that determines the stability of the equilibrium point[5]. The reader can check as an exercise that: (a) when the

[5]Baumol (1970, pp. 259-61) states that the time path will converge to the equilibrium point if $0 < \frac{df}{dy_t} < 1$ and diverge if $\frac{df}{dy_t} > 1$. These conditions, however, are necessary and sufficient only for *local* stability (or instability), whereas in general they are unduly restrictive, being only sufficient but not necessary. In fact, it can easily be checked that $\frac{df}{dy_t} > 1$ in some stretches of the phase line in Fig. 21.14 and that $0 < \frac{df}{dy_t} < 1$ in some stretches of the phase line in Fig. 21.15; and this is enough to show the lack of necessity. Given that $f(y_t)$ is an increasing function, the general necessary and sufficient

phase line is tangent to the 45° line and lies wholly on one side of it, the point of tangency is semi-stable, that is, stable or unstable according to whether the initial point is situated on one side or on the other of the equilibrium point; (b) when there are multiple equilibria they are alternatively stable (or semi-stable) and unstable (or semi-stable); (c) when there are no equilibrium points the movement is in any case divergent to $+\infty$ or $-\infty$.

Case (2): oscillations

Oscillatory movements occur when the phase line has a negative slope at least in some stretches. For simplicity, we shall examine only the case in which it has a negative slope everywhere. The oscillations may be damped, explosive or with a constant amplitude, as in Fig. 21.16 (a), (b) and (c), respectively. It is interesting to note that the presence of oscillations does not require the phase line to form a closed loop. Closed-loop phase lines, however, allow constant-amplitude oscillations which are not alternations [an alternation is an oscillation such that y reverses the direction of movement each period, as in diagrams 21.16 (a), (b) and (c)], that is, periodic movements such that the direction of movement is reversed ever n periods, n being greater than 1. For a closed-loop phase diagram, see Baumol (1970, p. 264). Also a *limit cycle* may occur, that is, a constant-amplitude cycle surrounded by other cycles whose amplitude is either decreasing or increasing. The limit cycle is stable if all the nearby cycles tend to it from both sides, as in Fig. 21.16 (d), and unstable in the opposite case. It can also be semi-stable, that is, when nearby cycles approach it from one sides and move away from it on the other side. Of course, multiple limit cycles may also occur.

We do not think that further exemplification is necessary. Once one

conditions are those given in the text in terms of the position of the phase line with respect to the 45° line. Apart from intuitive graphical considerations, they can be proved rigorously. The sufficiency can be proved using Liapunov's second method (see Chap. 23, Sect. 23.2, Theorem 23.2). Consider the equation $y_{t+1} = f(y_t)$. As a Liapunov function, take $V_t = |y_t - y_e|$. Then

$$\Delta V_t = |y_{t+1} - y_e| - |y_t - y_e| = |f(y_t) - y_e| - |y_t - y_e| .$$

Suppose now that $y_t < y_e$ and that $f(y_t) > y_t$ (that is, $f(y_t)$ lies above the 45° line to the left of the equilibrium point). Since f is an increasing function, it will also be $f(y_t) < y_e$. Thus we have

$$y_t < f(y_t) < y_e \quad \text{for} \quad y_t < y_e.$$

From this inequality we obtain $y_t - y_e < f(y_t) - y_e < 0$, so that, taking absolute values,

$$|y_t - y_e| > |f(y_t) - y_e| ,$$

and so $\Delta V_t < 0$. Similar reasoning shows that $\Delta V_t < 0$ also for $y_t > y_e$ if $f(y_t) < y_t$ (that is, if $f(y_t)$ lies below the 45° line to the right of the equilibrium point). Therefore $\Delta V_t <$ everywhere except for $y_t = y_e$, and the equilibrium point is (globally) stable. The necessity can be proved by showing a case in which the condition is not satisfied and the equilibrium is unstable, and this case is, for example, the one in Fig. 21.15 This completes the proof.

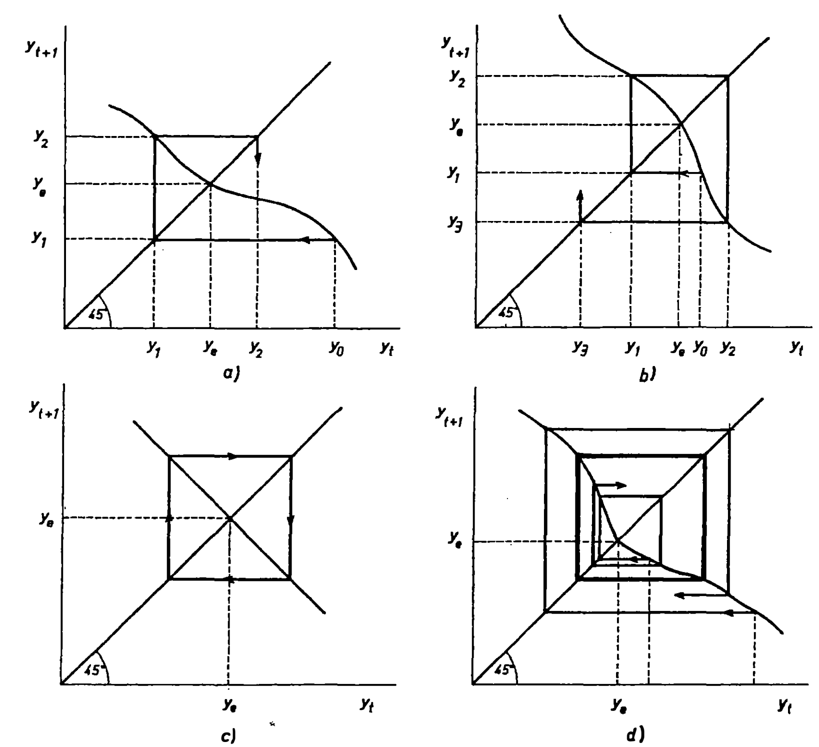


Figure 21.16: Phase diagram for a first-order difference equation: various oscillatory movements

has grasped the simple rules underlying the phase diagram method, one can continue alone to draw and analyse a great variety of such diagrams. A case of particular interest arises when the function f is not a monotonic curve as in the diagrams presented here, but a unimodal curve. In this case, in fact, apparently erratic movements may arise, that will be treated in the chapter devoted to chaotic dynamics (Chap. 26).

Let us note, finally, that the phase diagram can also be applied to the first-order linear equation with constant coefficients

$$y_{t+1} = ay_t + b,$$

which gives rise to a straight line in the phase diagram; an economic application *ante litteram* has been made in Chap. 4, Sect. 4.2, to which we refer the reader.

21.5.2 Two simultaneous difference equations

Let us consider the simultaneous 2×2 autonomous system

$$\begin{aligned} y_{1t+1} &= \phi_1(y_{1t}, y_{2t}), \\ y_{1t+2} &= \phi_2(y_{1t}, y_{2t}), \end{aligned} \tag{21.28}$$

and rewrite it as

$$\begin{aligned} \Delta y_{1t} &= \varphi_1(y_{1t}, y_{2t}), \\ \Delta y_{2t} &= \varphi_2(y_{1t}, y_{2t}), \end{aligned} \tag{21.29}$$

where $\Delta y_{it} \equiv y_{it+1} - y_{it}, \varphi_i \equiv \phi_i - y_{it}, i = 1, 2$.

We can then proceed to find the singular points, namely the values of y_{it} that make $\Delta y_{it} = 0$, and to draw the trajectories graphically by means of arrows as shown in Sect. 21.3.2.3 in relation to differential equation systems. However, the utmost care should be taken in this procedure. In fact, the system will not move continuously along one of the trajectories. Since we are in discrete time, the system will jump from point to point on that trajectory. Hence the system, though staying on the path that converges to equilibrium, may oscillate back and forth with ever increasing 'improper' oscillations or alternations (a negative real root with modulus greater than unity). Thus we must in any case compute the latent roots of the linear approximation to check the type of movement.

For example, the discrete-time analogue of the *monotonic* movement along the stable arm of a saddle point in continuous time, is the movement given by two positive real roots, one smaller and the other greater than one. But in discrete time we may also have an alternating movement along a unique stable trajectory (the stable real root is negative with modulus less than unity). There is no simple graphical possibility of distinguishing this latter case from the case of an alternating unstable movement along a similar path.

Thus we advise against the use of phase diagrams for checking the stability of a 2×2 non-linear difference system.

21.6 Economic applications

Stability analysis permeates all the economic applications treated in this book, hence the reader can take her pick. We only add that economic applications involving global stability will be treated in Chap. 23 and that economic applications involving the so called saddle-path stability will be treated in Chap. 22.

21.7 Exercises

1. Show that the linear system

$$\begin{bmatrix} \overline{y}_1' \\ \overline{y}_2' \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} \overline{y}_1 \\ \overline{y}_2 \end{bmatrix}$$

arises from the linearisation of both

$$\begin{aligned} y_1' &= y_2 + y_1(y_1^2 + y_2^2), \\ y_2' &= -y_1 + y_2(y_1^2 + y_2^2), \end{aligned}$$

and

$$\begin{aligned} y_1' &= y_2 - y_1(y_1^2 + y_2^2), \\ y_2' &= -y_1 - y_2(y_1^2 + y_2^2), \end{aligned}$$

at the equilibrium point (0,0). Show however that the first (non-linear) system is unstable while the second one is locally stable (Hint: to analyse the behaviour of the non-linear systems, multiply the first equation by y_1, the second by y_2, and add. Then show that the result comes from

$$y_1^2(t) + y_2^2(t) = \frac{c}{1 \pm 2ct}$$

where $c = y_1^2(0) + y_2^2(0)$, and the ± sign in the denominator respectively corresponds to the second and first non-linear system. Then...).

2. Consider the following non-linear system

$$\begin{aligned} \frac{dy_1}{dt} &= \varphi_1(y_1, y_2), \quad \frac{\partial \varphi_1}{\partial y_1} < 0, \frac{\partial \varphi_1}{\partial y_2} < 0, \\ \frac{dy_2}{dt} &= \varphi_2(y_1, y_2), \quad \frac{\partial \varphi_2}{\partial y_1} < 0, \frac{\partial \varphi_2}{\partial y_2} > 0. \end{aligned}$$

Performing a linear approximation (barred symbols denote deviations from equilibrium) we obtain

$$\begin{aligned} \frac{d\overline{y}_1}{dt} &= a_{11}\overline{y}_1 + a_{12}\overline{y}_2, \\ \frac{d\overline{y}_2}{dt} &= a_{21}\overline{y}_1 + a_{22}\overline{y}_2, \end{aligned} \qquad a_{ij} \equiv \left(\frac{\partial \varphi_i}{\partial y_j}\right)_E .$$

Given the assumptions made on the signs of $\partial\varphi_i/\partial y_j$, show that the equilibrium point is a saddle (Hint: see Sect. 21.3.2.4).

3. Consider the system $\mathbf{y}' = \mathbf{A}\mathbf{y} + \mathbf{h}(\mathbf{y})$, where

$$\mathbf{A} \equiv \begin{bmatrix} -2 & 1 & 3 \\ 0 & -6 & -5 \\ 0 & 0 & -1 \end{bmatrix}, \quad \mathbf{h}(\mathbf{y}) \equiv \begin{bmatrix} 9y_2^2 \\ 7y_3^5 \\ y_1^2 + y_2^2 \end{bmatrix}.$$

Check that this system is locally asymptotically stable (Hint: apply the Poincaré-Liapunov-Perron theorem).

4. Consider the system in polar coordinates $\mathbf{y} = (r, \theta), 0 \le r < \infty, 0 \le \theta < 2\pi$:

$$\begin{aligned} r' &= [g'(\theta,t)/g(\theta,t)]r, \\ \theta' &= 0, \end{aligned}$$

where

$$g(\theta,t) = \frac{\sin^2\theta}{\sin^4\theta + (1 - t\sin^2\theta)^2} + \frac{1}{1+t^2}.$$

The system's solution is

$$\begin{aligned} r(t; r_0, \theta_0, t_0) &= [g(\theta_0,t)/g(\theta_0,t_0)]r_0, \\ \theta(t; r_0, \theta_0, t_0) &= \theta_0. \end{aligned}$$

Show that this solution tends to the null solution as $t \to \infty$ but the system is not stable (Hint: take a $t_1 = \sin^{-2}\theta_0$ and show that $g(\theta_0, t_1) > \sin^{-2}\theta_0$, which can be made arbitrarily large as $\theta_0 \to \pm\pi$. Here the function $T(\mu, \mathbf{y}_0, t_0)$ is not continuous in $\mathbf{y}_0$ along the ray $\theta_0 = \pm\pi$).

21.8 References

Andronov, A.A., A.A. Vitt and S.E. Khaikin, 1966, *Theory of Oscillators*, Chaps. I-V.
Baumol, W.J., 1958, Topology of Second Order Linear Difference Equations with Constant Coefficients.
Baumol, W.J., 1970 *Economic Dynamics*, Chap. 13; Chap. 14, Sect. 1.
Bellman, R., 1953, *Stability Theory of Differential Equations.*
Blatt, J.M., 1983, *Dynamic Economic Systems,* Chaps. 7 and 8.
Boole, G., 1960, *A Treatise on the Calculus of Finite Differences.*
Bothwell, F.E., 1952, The Method of Equivalent Linearization.
Braun, M., 1986, *Differential Equations and their Applications,* Chap. 4, Sects. 4.1-4.4, 4.7.
Coddington, E.A. and N. Levinson, 1955, *Theory of Differential Equations,* pp. 314 ff.
Edgeworth, F.Y., 1881, *Mathematical Psychics,* pp. 19-20.

Fisher, F.M., 1983, *Disequilibrium Foundations of Equilibrium Economics.*
Forsyth, A.R., 1890-1902, *Theory of Differential Equations*, Vols. I-IV.
Gabisch, G. and H.-W. Lorenz, 1989, *Business Cycle Theory: A Survey of Methods and Concepts.*
Gandolfo, G., 1992, Sensitivity Analysis in Continuous Time Econometric Models.
Guckenheimer, J. and P. Holmes, 1986, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields,* Chap. 1, Sect. 1.7, Chap. 5, pp. 258-259.
Hirsch, M.W. and S. Smale, 1974, *Differential Equations, Dynamical Systems and Linear Algebra.*
Ince, E.L., 1956 (1926), *Ordinary Differential Equations.*
Kaldor, N., 1934, A Classificatory Note on the Determinateness of Equilibrium.
Kalman, R.E. and J.E. Bertram, 1960, Control System Analysis and Design Via the 'Second Method' of Lyapunov, I (Continuous-Time Systems).
Krasovskiĭ, N.N., 1963, *Stability of Motion,* Introduction.
Marshall, A., 1879a, *The Pure Theory of Foreign Trade.*
Marshall, A., 1879b, *The Pure Theory of Domestic Values.*
Marshall, A., 1890, *Principles of Economics.*
Milne-Thomson, L.M., 1933, *The Calculus of Finite Differences.*
Minorsky, N., 1962, *Nonlinear Oscillations*, Chaps. 1-4.
Olech, C., 1963, On the Global Stability of an Autonomous System on the Plane.
Samuelson, P.A., 1947, *Foundations of Economic Analysis,* p. 5.
Sansone, G., and R. Conti, 1964, *Non-Linear Differential Equations,* Chaps. I, II, IV, VI.
Vercelli, A., 1984, Fluctuations and Growth: Keynes, Schumpeter, Marx, and the Structural Instability of Capitalism.
Walras, L., 1954 (1874), *Elements of Pure Economics.*

Chapter 22

Saddle Points and Economic Dynamics

A recurrent singular point in economic dynamics is the saddle point (see Chap. 21, Sect. 21.3.2). The reason is that two widely used dynamic tools, utility maximisation through time in optimal growth models, and rational expectations, often give rise to this type of singular point.

It is also customary in economics to speak of 'saddle-path (or saddle-point) stability'. Although a saddle point is unstable according to the basic mathematical definition of stability (see Chap. 21, Sect. 21.2.1), it is conditionally stable (see Chap. 21, Sect. 21.2.2). The economic denomination finds its justification in the fact that in the problems under consideration the initial conditions can usually be chosen so as to put the system on the stable arm of the saddle.

More precisely, when the dynamic equations of the problem show that the singular point is a saddle point, we are in the context of 'conditional stability' described in Chap. 18, Sect. 18.2.2.3 and Chap. 21, Sect. 21.2.2. If, given the initial value of one of the two variables, we are free to choose the initial value of the other one, then it will always be possible to put the system on the stable arm of the saddle, as shown in Sects. 18.2.2.3 and 21.3.2.4. This 'freedom' is exactly what we have in optimal control problems (thanks to the presence of a control variable) and in rational expectations models (thanks to perfect information and foresight on the part of economic agents).

Here we shall show how these cases can be handled rigorously. For this purpose, it will be necessary briefly to recall a few notions from optimal control theory and from rational expectations theory. In Sect. 22.1.1 the general problem of optimal control is introduced, and in Sect. 22.1.2 one of the methods currently used to solve it (the maximum principle) is stated without proofs. In Sect. 22.3.1 the properties of rational expectations are stated, together with the distinction between 'jump' and 'historical' variables.

22.1 Saddle points in optimal control problems

22.1.1 Introduction

The difference between control (see Chap 17, Sect. 17.1) and *optimal* control is simply that in the latter case an optimization process through time is involved. More precisely, given a dynamic system (containing n *state variables* $\mathbf{x}(t)$ and r *control variables* $\mathbf{u}(t)$, $n \gtreqless r$) and its equations of motion, and given also an explicit index of the 'performance' of the system through time, the control variables are determined so as to optimise this index, called the *objective functional.* The control variables are constrained to take on values belonging to a compact and convex set Ω, namely $\mathbf{u}(t) \in \Omega$. Hence a control trajectory is feasible if its values belong to Ω. The set of all feasible control trajectories forms the control set U, in which we must choose the optimal trajectory as stated above.

Optimal control is an important tool in economic analysis, but its treatment would require a book on its own. Here optimal control (also called *dynamic optimization*) comes into consideration simply because in economic applications one of the methods currently used to deal with it (the maximum principle) often gives rise (in the 2×2 case) to differential equation systems whose phase diagram shows a typical saddle point. Hence we shall simply state the problem and recall the few basic theorems that are required for the purpose at hand. For proofs and an in-depth treatment see, for example, Intriligator (1971) and Kamien and Schwartz (1991).

The typical layout of a general optimal control problem is given in Eq. (22.1), where the equations are self-explanatory. Note that the final function is not necessarily present, and that when the terminal time is ∞ we speak of an *infinite horizon* optimal control problem. This of course has a sense only if the infinite integral converges, which is often accomplished by the introduction of discounting. Also note that the classical calculus of variations can be considered as a particular case of the general control problem, obtained by putting $\mathbf{x}' = \mathbf{u}$.

The solution of the optimal control problem will give the optimal trajectory $\mathbf{u}^*$ of the control variables and hence, through the equations of motion, of the state variables. As regards the control variables, the solution may give rise to two kinds of optimal trajectories. One is *open loop control,* i.e. $\mathbf{u}^*= \mathbf{u}^*(t)$, which means that the optimal trajectory is a function of time only and is completely determined at the initial time. The other one is *closed loop control,* $\mathbf{u}^*= \mathbf{u}^*(\mathbf{x}(t), t)$, namely the optimal trajectory is not only a function of time, but also of the current (at time t) values of the state variables, which means that the control actions can be revised in the light of new information arising from the current state of the system. A Friedman monetary rule

in which money supply grows at a constant and predetermined rate is an example of open loop control. A Keynesian monetary rule in which money supply growth is determined according to the current state of the economy is an example of closed loop control.

$$\max_{\underbrace{\{\mathbf{u}(t)\}}_{\text{control vector}}} J = \underbrace{\int_{t_0}^{t_1} \underbrace{I(\mathbf{x},\mathbf{u},t)\mathrm{d}t}_{\textit{intermediate function}} + \underbrace{F(\mathbf{x}_1,t_1)}_{\textit{final function}}}_{\textit{objective functional}}, \qquad t_1 \text{ possibly } = \infty$$

subject to:

$\mathbf{x}' = \mathbf{f}(\mathbf{x},\mathbf{u},t)$	equations of motion of the state variables,
t_0 and $\mathbf{x}(t_0) = \mathbf{x}_0$ given	initial conditions,
t_1 (and hence $\mathbf{x}(t_1)$) given, or, alternatively $(\mathbf{x}(t),t) \in T$ for $t = t_1$, i.e. $\mathbf{T}(\mathbf{x},t) = \mathbf{0}$ for $t = t_1$	terminal time given, or terminal surface T given,
$\{\mathbf{u}(t)\} \in U$ control vector belonging to U	control set U given.

(22.1)

Apart from the calculus of variations when applicable, there are two methods to deal with the general control problem: the *maximum principle* by Pontryagin and associates, and Bellman's *dynamic programming.* Here we shall deal with the maximum principle for the reason stated above.

22.1.2 The maximum principle

The maximum principle involves as a first step the construction of the *Hamiltonian*

$$H(\mathbf{x}, \mathbf{u}, \boldsymbol{\mu},t) = I(\mathbf{x},\mathbf{u},t) + \boldsymbol{\mu}\mathbf{f}(\mathbf{x},\mathbf{u},t), \tag{22.2}$$

where $\boldsymbol{\mu}$ is a vector of *costate variables* (also called dynamic Lagrange multipliers, auxiliary variables, adjoint variables, dual variables). The second step consists in determining the optimal trajectories for $\mathbf{u}(\mathbf{t}), \mathbf{x}(\mathbf{t}), \boldsymbol{\mu}(t)$, which are those that satisfy the following conditions (in the case of a problem with given terminal time)

$$\max_{\{\mathbf{u}\in\Omega\}} H(\mathbf{x}, \mathbf{u}, \boldsymbol{\mu},t) \qquad \text{for any } t,\ t_0 \le t \le t_1 \tag{22.3}$$

and

$$\begin{aligned} \mathbf{x}' &= \frac{\partial H}{\partial \boldsymbol{\mu}} = \mathbf{f}(\mathbf{x},\mathbf{u},t), \quad \mathbf{x}(t_0) = \mathbf{x}_0, \\ \boldsymbol{\mu}' &= -\frac{\partial H}{\partial \mathbf{x}}, \quad \boldsymbol{\mu}(t_1) = \frac{\partial F}{\partial \mathbf{x}_1}. \end{aligned} \tag{22.4}$$

Equations (22.3) involve the maximisation of the Hamiltonian with respect to the control variables. Note that when the control variables are not constrained, then the maximisation reduces to applying the traditional conditions $\partial H/\partial \mathbf{u} = \mathbf{0}$.

Equations (22.4) are called the *canonical equations* (sometimes called Euler equations for their similarity with equations appearing in the classical calculus of variations, due to Euler), and their solution involves a *two-point boundary value problem.* The name derives from the fact that the $2n$ side conditions involve two *different* points of time, the initial time (the n differential equations for the state variables) and the terminal time (the n differential equations for the costate variables; note that when the final function is not present, we simply have $\boldsymbol{\mu}(t_1) = \mathbf{0}$, or $\lim_{t\to\infty} \boldsymbol{\mu}(t) = \mathbf{0}$ in the case of an infinite terminal time). Two-point boundary value problems are rather more difficult to solve than the standard initial value problem. However, in economic applications it often (fortunately) happens that the maximisation of the Hamiltonian gives the form of the solution of the optimal control problem without having to solve the canonical equations.

Conditions (22.3) and (22.4) are *necessary* conditions. They turn out to be also sufficient if the Hamiltonian is linear in the control variables or if the maximised Hamiltonian is a concave function of the state variables.

When instead of the terminal time we are given the *terminal surface,* i.e. $\mathbf{T}(\mathbf{x},t) = \mathbf{0}$ for $t = t_1$, conditions (22.3) and (22.4) must be supplemented with the *transversality condition*

$$\left(H + \frac{\partial F}{\partial t_1}\right) + \left(\frac{\partial F}{\partial \mathbf{x}_1} - \boldsymbol{\mu}\right)\left(\frac{d\mathbf{x}}{dt}\right)_{\mathbf{T}(\ldots)=\mathbf{0}} = 0, \tag{22.5}$$

where all the variables are evaluated at the terminal time t_1.

When the final function F is not present, we have in any case the transversality condition

$$\boldsymbol{\mu}(t_1)\mathbf{x}(t_1) = 0, \tag{22.6}$$

which takes the form

$$\lim_{t_1\to\infty} \boldsymbol{\mu}(t_1)\mathbf{x}(t_1) = 0, \tag{22.7}$$

in the case of infinite terminal time.

By and large, the maximum principle can be considered as an extension to dynamic optimization of the method of Lagrange multipliers used in static optimization problems. The (static) Lagrange multipliers, as is well known, yield useful information on the effects of parameter changes on the optimal solution, hence it seems natural to inquire whether the costate variables can serve the same purpose in dynamic optimization problems. The answer is yes, and denoting by J^* the optimal value of the objective functional we have the following results:

$$\begin{aligned} \frac{\partial J^*}{\partial t_0} &= -[I(\mathbf{x}^*, \mathbf{u}^*, t)]_{t_0}, \\ \frac{\partial J^*}{\partial t_1} &= [I(\mathbf{x}^*, \mathbf{u}^*, t)]_{t_1} + \frac{\partial F}{\partial \mathbf{x}^*(t_1)} \frac{d\mathbf{x}^*(t_1)}{dt_1} + \frac{\partial F}{\partial t_1}, \\ \frac{\partial J^*}{\partial \mathbf{x}(t_0)} &= \boldsymbol{\mu}^*(t_0). \end{aligned} \tag{22.8}$$

Let us note, to conclude this brief summary, that when the objective functional has the dimension price×quantity and the state variables have the dimension quantity, then the costate variable has the dimension price, i.e. it is a *shadow price.* Thus a dual problem of valuation through time (determination of the trajectories of the costate variables) corresponds to a dynamic allocation problem. This is clearly the dynamic analogue of the interpretation of Lagrange multipliers in static optimization problems.

Before going on to economic applications, it should be stressed that the saddle-point nature of the differential equation systems arising from the solution to optimal control problems is by no means a mathematical property. In other words, there is nothing in the mathematical structure of optimal control that makes the emergence of a saddle point necessary. In general, even in 2×2 systems, other singular points are equally possible, as shown for example by Kamien and Schwartz (1991, Part II, Section 9).

It is then the *economic* nature of the problem that usually gives rise to saddle-point stability (Samuelson, 1972). But usually does not mean always, and in fact it can be shown that also oscillatory movements may arise in optimal growth problems. These cases require some more advanced mathematical tools, and will be treated in Sect. 25.2.5.

22.2 Optimal economic growth

Optimal growth theory dates back at least to Ramsey's (1928) famous model, but it was only in the 1960s that it became a hot topic concomitantly with the flow of research on neoclassical growth theory (see Chap. 13, Sect. 13.2). And it disappeared when the interest in growth theory dwindled away. Formal growth theory has now come back into fashion and optimal growth theory with it. The point of departure remains the Ramsey model, as amended by the growth theorists of the 1960s to take account of discounting.

22.2.1 Optimal growth: traditional

In the neoclassical growth model (see Chap. 13, Sect. 13.2) the saving rate is exogenous and constant. The brilliant idea of Ramsey (1928) was to determine the saving rate endogenously, through a dynamic maximisation process. Ramsey's original treatment—that, by the way, antedates any formal treatment of growth theory—is examined in Exercise 1; here we illustrate the modern treatment of the problem.

22.2.1.1 The setting of the problem

Let us then consider the basic Solow-Swan model and put it in a slightly different form. Assuming radioactive depreciation, gross investment $I(t)$ serves partly to replace the fraction δ of the existing capital stock and partly to increase the capital stock (net investment), namely

$$I(t) = \delta K(t) + K'(t),$$

and, using lower-case letters to denote variables in per capita terms, we have

$$i(t) = \frac{K'(t)}{L(t)} + \delta k(t) = k'(t) + (n+\delta)k(t), \tag{22.9}$$

where n is the rate of growth of the labour force (possibly including labour-augmenting technical progress), and $k'(t) = K'(t)/L(t) - kn$ by definition. To prevent notational confusion, we point out that in Sect. 13.2 we used r to denote capital per head, in conformity with Solow's notation. The notation that we use here is standard in current growth theory.

If we now recall the macroeconomic equilibrium condition $Y(t) = C(t) + I(t)$ and put it in per capita terms, we have $y(t) = c(t) + i(t)$. From this and the production function $Y(t) = F(K(t), L(t))$—which gives $y(t) = F(k(t), 1) = f(k(t))$—we can rewrite Eq. (22.9) as

$$f(k(t)) = c(t) + (n+\delta)k(t) + k'(t), \tag{22.10}$$

which states that output per head is allocated to three uses. These are, in the order: consumption per head, maintainment of the level of capital per head (both by equipping new workers with the same amount of capital per head as the existing ones, and by replacing worn-out capital), increase in capital per head. Let us note for future reference that there is a stock of capital per head such that the entire output per head is just enough to maintain it, with no consumption nor net investment. This stock is obtained by setting $c(t) = k'(t) = 0$ in Eq. (22.10) and solving the resulting equation

$$f(k(t)) = (n+\delta)k(t) \Rightarrow k = \tilde{k}. \tag{22.11}$$

Since the production function is assumed to satisfy the Inada conditions (see Sect. 13.2.1), this equation has a unique positive solution $\tilde{k}$. In fact, the solution(s) of Eq. (22.11) can be considered as the point(s) of intersection of the curve $f(k)$ with the straight line $(n+\delta)k$. Excluding the trivial solution $k = 0$, recall that $f(k)$ is a monotonically increasing function whose slope $f'(k)$ tends to infinity as $k \to 0$ and to zero as $k \to \infty$. Hence the curve $f(k)$ certainly lies above the straight line at the origin, and will cut it once, and only once, at a positive finite value of k.

The basic equation of the Solow-Swan model is easily derived from Eq. (22.10) by assuming a given and constant propensity to consume b, whence $c(t) = by(t)$. Substituting in Eq. (22.10), remembering that $(1-b) = s$ and rearranging terms, we obtain $k' = sf(k) - (n+\delta)k$, which coincides with the fundamental dynamic equation (13.45).

The problem is now to *determine* the time path of $c(t)$ so as to maximize social utility through time. Thus $c(t)$ is the control variable in the terminology of optimal control theory. The control set is clearly $0 \le c(t) \le f(k)$, since current consumption per head cannot be negative nor greater than current output per head. If we allow for the possibility of 'eating' also the existing stock of capital—a possibility consistent with our one-good economy—the upper limit to $c(t)$ would be $f[k(t)] + k(t)$, but we shall ignore this complication.

The maximisation process can be put into being by households (considered as a set of immortal extended families) or by a benevolent social planner: in both cases the result is the same (see Barro and Sala-i-Martin, 1995, Chap. 2, Sect. 2.4). The traditional assumption is that society's utility is a cardinal function of consumption per head, with positive but decreasing marginal utility, that is

$$U(t) = U(c(t)), \quad U_c(c) = \frac{dU}{dc} > 0, \; U_{cc}(c) = \frac{d^2U}{dc^2} < 0, \tag{22.12}$$

with the additional properties

$$\lim_{c\to 0} U_c(c) = \infty, \quad \lim_{c\to\infty} U_c(c) = 0. \tag{22.13}$$

If we assume that all households have the same cardinal utility function $u(c)$, then we can write

$$U(t) = u(c(t))e^{nt}, \tag{22.14}$$

where $e^{nt} = L(t)$ after normalising the initial population to unity.

Let us also define for future reference the *elasticity of marginal utility*

$$\sigma(c) = -c\frac{U_{cc}}{U_c}, \tag{22.15}$$

which is positive for any $c > 0$. It is often assumed that the utility function is such that the elasticity of marginal utility is constant, but this is not required for the results that we are going to derive.

The utility function U (also called the felicity function) gives the overall utility corresponding to per capita consumption at a given point of time, but the problem is that of choosing a trajectory of per capita consumption, which implies the valuation of utility levels at different points of time. We then assume that utility levels at different points of time are independent, namely that utility at any given point of time only depends on consumption per head at that point. This assumption of a *time separable* utility function is essential for simplifying the analysis. We also assume that future utility levels are *discounted* at a given and constant rate ρ, which represents the rate of time preference. Discounting means giving a greater weight to nearer (in time) generations than to more distant generations, and has been subjected to much criticism (Ramsey, for example, argued on ethical grounds that there should be no discounting). We shall however keep with current practice and allow for discounting.

The second preliminary question to be discussed is the time horizon, which is usually taken from now to infinity (the case with finite terminal time will be examined in exercise 2). Thus we have to analyse the convergence of the infinite integral

$$W = \int_{t_0}^{\infty} e^{-\rho(t-t_0)} U(c(t))\mathrm{d}t. \tag{22.16}$$

A sufficient convergence condition is that the initial capital per head, k_0, is smaller than the maximum sustainable level $\bar{k}$ (see above). In fact, since $f(k) \leq f(\bar{k})$ for $k \leq \bar{k}$, and $c(t) \leq f(k)$, we have $c(t) \leq f(\bar{k})$ and consequently

$$\begin{aligned}\int_{t_0}^{\infty} e^{-\rho(t-t_0)} U(c(t))\mathrm{d}t &\leq \int_{t_0}^{\infty} e^{-\rho(t-t_0)} U(f(\bar{k}))\mathrm{d}t = U(f(\bar{k})) \int_{t_0}^{\infty} e^{-\rho(t-t_0)}\mathrm{d}t \\ &= U(f(\bar{k}))\left[-\frac{1}{\rho}e^{-\rho(t-t_0)}\right]_{t_0}^{\infty} = \frac{U(f(\bar{k}))}{\rho},\end{aligned} \tag{22.17}$$

which shows that W has a finite upper bound.

If we use the alternative formulation (22.14), then a sufficient convergence condition that applies when c is constant is $\rho > n$.

We now have all the information required to set up our optimal control

problem, which is

$$\max_{\{c(t)\}} W = \int_{t_0}^{\infty} e^{-\rho(t-t_0)} U(c(t))\mathrm{d}t\ ,$$

subject to:

$$\begin{aligned}
&k' = f(k) - (n+\delta)k - c, && \text{equation of motion of the state variable}\\
&k(t_0) = k_0 < \bar{k}, && \text{initial condition}\\
&0 \leq c(t) \leq f(k). && \text{control set}
\end{aligned} \tag{22.18}$$

This is a fairly simple optimal control problem, that can easily be dealt with by the maximum principle.

22.2.1.2 The optimality conditions in the basic neoclassical model

Let us form the Hamiltonian

$$H = e^{-\rho(t-t_0)} U(c) + \mu[f(k) - (n+\delta)k - c],$$

that is, by introducing the new costate variable

$$q = \mu e^{\rho(t-t_0)}, \tag{22.19}$$

$$H = e^{-\rho(t-t_0)} \{U(c) + q[f(k) - (n+\delta)k - c]\}. \tag{22.20}$$

The reader should note that this change of variable is a mathematically convenient way of dealing with problems involving discounting.

We now look for the possible existence of an interior maximum to the Hamiltonian by setting

$$\frac{\partial H}{\partial c} = e^{-\rho(t-t_0)}\{U_c(c) - q\} = 0,$$

from which

$$q = U_c(c). \tag{22.21}$$

Since the utility function is concave (i.e. $U_{cc}(c) < 0$), $U_c(c)$ is monotonically decreasing from ∞ to zero by properties (22.13). Hence a unique solution always exists. Moreover, since $U_{cc}(c) < 0$, this solution is indeed a maximum. Equation (22.21) states that the shadow price of per capita capital accumulation along the optimal path equals the marginal utility of per capita consumption.

Let us now come to the canonical equation for the (original) costate variable, which is

$$\mu' = -\frac{\partial H}{\partial k} = -\mu[f_k(k) - (n+\delta)]. \tag{22.22}$$

To transform (22.22) into a differential equation for the costate variable q (which also represents a shadow price) observe that $\mu = qe^{-\rho(t-t_0)}$, hence $\mu' = q'e^{-\rho(t-t_0)} - \rho q e^{-\rho(t-t_0)}$. By substituting into Eq. (22.22) and collecting terms we have

$$q' = -[f_k(k) - (n+\delta) - \rho]q, \tag{22.23}$$

from which

$$f_k(k) + \frac{q'}{q} - \delta - n - \rho = 0. \tag{22.24}$$

We can interpret this equation as the condition of *zero net profit rate*, where the net profit rate is defined as the gross profit rate (the marginal productivity of capital) plus capital gains (q'/q) minus the losses due to depreciation (δ) minus the 'dilution' due to population growth (n) minus the intertemporal cost of waiting, represented by the subjective discount rate (ρ).

Note that Eqs. (22.21) and (22.24) already give us a lot of information on the optimal growth path even if we do not yet know it.

Let us now observe that by logarithmic differentiation of the optimality condition (22.21) with respect to time we obtain

$$\frac{q'}{q} = \frac{U_{cc}(c)}{U_c(c)}c' = -\sigma(c)\frac{c'}{c},$$

which can be substituted in the canonical equation (22.24) to obtain

$$c' = \frac{1}{\sigma(c)}[f_k(k) - (n+\delta+\rho)]c.$$

In this way we have reduced the differential equation for the costate variable to a differential equation for the control variable. Thus the canonical equations of our optimal control problem are

$$\begin{aligned} c' &= \frac{1}{\sigma(c)}[f_k(k) - (n+\delta+\rho)]c, \\ k' &= f(k) - (n+\delta)k - c, \end{aligned} \tag{22.25}$$

which are a set of two *autonomous* non-linear differential equations.

Let us note that, if we used the alternative formulation of the utility function given in Eq. (22.14), then we would obtain the system

$$\begin{aligned} c' &= \frac{1}{\sigma(c)}[f_k(k) - (\delta+\rho)]c, \\ k' &= f(k) - (n+\delta)k - c, \end{aligned} \tag{22.26}$$

where $\sigma(c) = -u_{cc}(c)/u_c(c)$. This system is formally equivalent to (22.25) except for the obvious disappearance of the parameter n in the differential equation concerning c, and can be analysed in the same way. Let us then consider system (22.25) and analyse it by means of a phase diagram.

We first determine the singular point of the system by setting $c' = k' = 0$. This gives

$$\begin{aligned} f_k(k) &= n+\delta+\rho, && \text{which determines } k^*, \\ c^* &= f(k^*) - (n+\delta)k^*, && \text{which determines } c^*. \end{aligned} \tag{22.27}$$

The solution can be simply illustrated in a diagram (see Fig. 22.1).

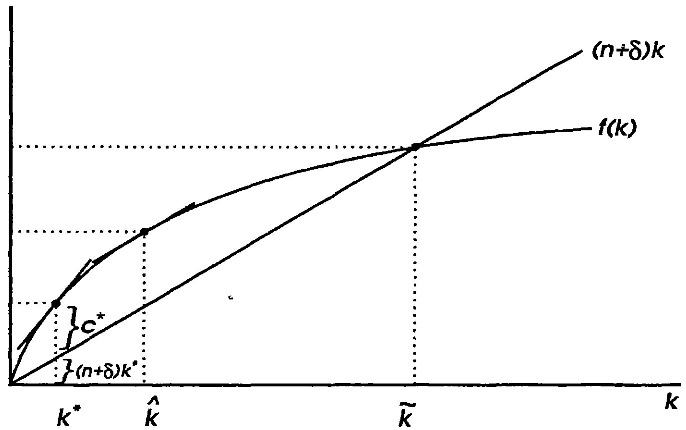


Figure 22.1: Equilibrium point of the neoclassical optimal growth model

The diagram shows $\tilde{k}$ as defined in Eq. (22.11) and k^* as determined by equating the slope of $f(k)$ to $(n+\delta+\rho)$, which is what the first equation in (22.27) says. Then c^* is determined by subtracting the ordinate of the straight line, which is $(n+\delta)k^*$, from the ordinate of the curve, which is $f(k^*)$. It can be immediately checked that $0 < c^* < f(k^*)$, so that the constraints on the control variable are satisfied.

For reference we also have drawn $\hat{k}$, which is the value at which $f_k(k) = n+\delta$. This is what the *golden rule* (see Sect. 13.2.3.2) states. For analogy, the condition $f_k(k) = n + \delta + \rho$ holding here is called the *modified golden rule*. It is obvious that for $\rho = 0$ we are back to the old golden rule.

The solution k^*, c^* is called the *balanced growth path* because in it both capital per head and consumption per head are constant, which means that the capital stock (and hence output) and consumption grow at the same rate as the population.

Let us now observe that, since we are in a case in which the final function is absent, the boundary condition for μ is $\lim_{t\to\infty}\mu(t) = 0$. This is satisfied because integrating the differential equation (22.22) we have $\mu = M\exp\left(-\int[f_k(k) - (n+\delta)]dt\right)$, where M is an arbitrary constant. Since $f_k(k) - (n+\delta) > 0$ because at the optimum we have $f_k(k) = (n + \delta + \rho)$ as shown by Eqs. (22.27), it follows that μ converges to zero as t goes to infinity.

22.2.1.3 Saddle-point transitional dynamics in the basic neoclassical model

We start from the phase diagram proper, which serves to examine the nature of the equilibrium point or, as it were, the *transitional dynamics* of the model.

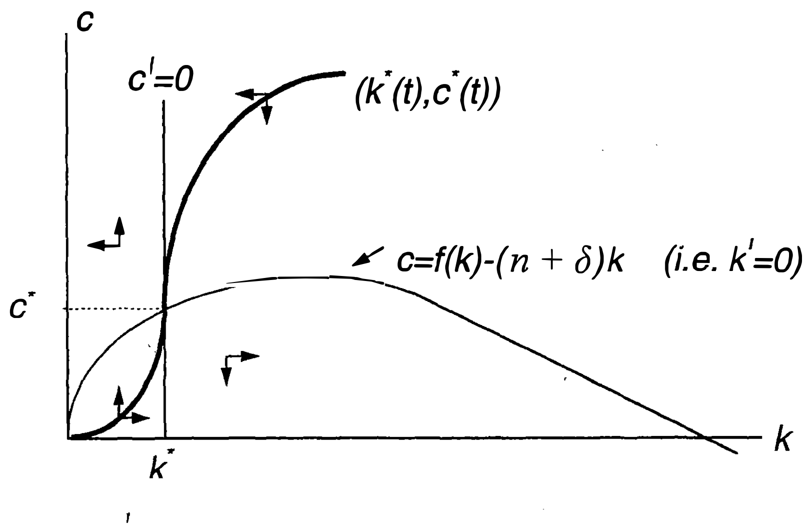


Figure 22.2: The neoclassical optimal growth model: transitional dynamics

The relation $f_k(k) - (n + \delta + \rho) = 0$ determines $k = k^*$ and hence in Fig. 22.2 we have the curve $c' = 0$. Since $f_k(k)$ is a monotonically decreasing function of k, we have

$$c' \gtreqless 0 \quad \text{according as } f_k(k) \gtreqless (n + \delta + \rho) \text{ i.e. according as } k \lesseqgtr k^*.$$

Therefore, at all points to the *left* (*right*) of k^*, c' is positive (negative) and hence c *increases* (*decreases*). This is denoted by the arrows respectively pointing upwards (downwards).

From the differential equation of the costate variable q we immediately have

$$k' \gtreqless 0 \quad \text{according as } c \lesseqgtr f(k) - (n+\delta)k.$$

Therefore, at all points *below* (*above*) the curve $c = f(k) - (n+\delta)k, k'$ is positive (negative) and hence k *increases* (*decreases*), as shown by the arrows respectively pointing rightwards (leftwards).

A careful examination of the arrows suggests that the singular point c^*, k^* is a *saddle point*, whose stable arm will necessarily pass through the zones where both arrows point toward equilibrium. This stable arm will then have the aspect drawn in Fig. 22.2 and denoted $(k^*(t), c^*(t))$. The saddle point property is a typical result in optimal growth models.

To confirm that the singular point is indeed a saddle point we examine its local properties by means of a linearisation at (k^*, c^*). Since the differential system (22.25) is autonomous the conditions of the Poincaré-Liapunov-Perron theorem are satisfied. The first-order terms of the Taylor series expansion yield

$$\begin{aligned} \bar{c}' &= \frac{c^* f_{kk}(k^*)}{\sigma(c^*)}\bar{k}, \\ \bar{k}' &= -\bar{c} + [f_k(k^*) - (n+\delta)]\bar{k} = -\bar{c} + \rho\bar{k}, \end{aligned} \tag{22.28}$$

where the overbar denotes deviations from equilibrium as usual. Note that $\bar{c}$ does not appear in the first equation because the partial derivative with respect to c is zero at the equilibrium point; in the second equation we have used the fact that $f_k(k^*) - (n+\delta) = \rho$ at the equilibrium point.

The characteristic equation of system (22.28) is

$$\begin{vmatrix} -\lambda & \dfrac{c^* f_{kk}(k^*)}{\sigma(c^*)} \\ -1 & \rho - \lambda \end{vmatrix} = \lambda^2 - \rho\lambda + \frac{c^* f_{kk}(k^*)}{\sigma(c^*)} = 0. \tag{22.29}$$

The succession of the signs of the coefficients is $+ - -$, hence there will be one positive and one negative real root, which denotes a saddle point (see above, Sect. 21.3.2.4).

Since the initial per capita stock of capital, k_0, is given, it is essential that the initial value of the control variable c_0 be chosen as the corresponding ordinate of the stable arm, namely as the ordinate corresponding to k_0 on the stable arm. Any other choice would give rise to a path that eventually violates the necessary conditions for an optimum, either because it would imply non-feasible points in the upper left part of the figure (in that part c grows without bounds and violates the constraint $c \leq f(k)$), or because it would imply inferior point in the lower right part of the figure (inferior because there is less c available for any given k).

From the mathematical point of view, choosing c_0 in correspondence to k_0 on the stable arm means choosing the initial conditions in such a way that the arbitrary constant appearing in the term containing the unstable root turns out to be zero, as shown in Chap. 18, Sect 18.2.2.3.

Let us observe, in conclusion, that—since the movement along the stable arm is monotonic—we shall observe a monotonic convergence to the balanced growth path, with values of both (c, k) that are ever increasing or ever decreasing according as the initial per capita stock of capital, k_0, is to the left or to the right of k^*. The speed of convergence is examined in exercise 3.

22.2.2 Optimal growth: endogenous

22.2.2.1 A model of optimal endogenous growth

In parallel with endogenous growth models (see Sect. 13.2.4.3), models of optimal endogenous growth have been developed. One of the crucial properties of models of endogenous growth is the absence of diminishing returns to capital. This apparently unrealistic assumption can be justified if one thinks of capital in the broad sense, namely as including human capital in addition to physical capital. The simplest production function with these properties is the AK production function, namely

$$Y = AK, \tag{22.30}$$

where A is a positive constant that reflects the technological level. The AK model, however, does not exhibit the convergence property (see Chap. 13, Sect. 13.2.4.2; Barro and Sala-i-Martin, 1995, Sect. 1.3.1). It is however possible to obtain both results (endogenous growth and convergence) if one reintroduces diminishing returns to capital combined with a positive (rather than zero) lower limit to capital's marginal productivity as the capital stock tends to infinity. This latter property amounts to allowing the violation of one of the Inada conditions (see Chap. 13, Sect. 13.2.1).

The simplest way of combining both properties is to use the technology considered by Jones and Manuelli (1990)

$$Y = F(K, L) = AK + \Omega(K, L), \tag{22.31}$$

which is a combination of the AK production function and of $\Omega(K, L)$, that is a well-behaved neoclassical production function satisfying the Inada conditions. Since $\partial Y/\partial K = A + \partial\Omega/\partial K$, it is easy to see that $\lim_{K\to\infty}[\partial Y/\partial K] = A > 0$.

The idea underlying (22.31) is that the AK part of the production function will provide endogenous growth, whereas the $\Omega(K, L)$ part will give rise to convergence. We shall now put these ideas in an optimizing framework, following Barro and Sala-i-Martin (1995, Chap. 4, Sect. 4.5). To keep matters simple we shall assume a Cobb-Douglas functional form for the $\Omega(K, L)$ part. Thus we start from

$$Y = F(K, L) = AK + BK^{\alpha}L^{1-\alpha}, \quad 0 < \alpha < 1, \tag{22.32}$$

where capital is assumed to depreciate at the constant rate δ and L grows at the constant rate n (as usual, we could introduce labour-augmenting technical progress and measure L in efficiency units). In per-capita terms we have

$$y = f(k) = Ak + Bk^{\alpha}. \tag{22.33}$$

We also assume that $A > \delta + n$, so that there is no finite stock of capital per head such that the entire output per head is just enough to maintain it, with no consumption nor net investment. As we know from Sect. 22.1.3.1, such a stock would be obtained by solving the equation

$$f(k) = (n + \delta)k,$$

that has a unique positive solution when the Inada conditions are satisfied. In our case the condition $\lim_{k\to 0} f_k(k) = \infty$ is still satisfied, but the condition $\lim_{k\to\infty} f_k(k) = 0$ is not, because $\lim_{k\to\infty} f_k(k) = A > n + \delta$. Hence the slope of the curve $f(k)$ remains always greater than the slope of the straight line $(n + \delta)k$ as $k \to \infty$, and the equation $f(k) = (n + \delta)k$ has no positive finite solution.

We finally assume that the discount factor is greater than the rate of growth of the population, namely $\rho > n$, and that $A > \delta + \rho$. This last assumption means that the production function is sufficiently productive to ensure a steady state positive growth of consumption per head, which is the essence of endogenous growth in this model. Also note that $A > \delta + \rho$ together with $\rho > n$ imply $A > n + \delta$.

Apart from the specification of the production function, the setting of the problem is the same as in the previous section. Here we adopt the per-capita specification of the utility function, and we normalize the initial time to zero, so that we have

$$\max_{\{c(t)\}} W = \int_0^\infty e^{-(\rho-n)t} u(c(t))dt\ ,$$

subject to:

$k' = f(k) - (n+\delta)k - c,$	equation of motion of the state variable
$k(t_0) = k_0,$	initial condition
$0 \le c(t) \le f(k).$	control set

(22.34)

22.2.2.2 The conditions for optimal endogenous growth

We start from the Hamiltonian

$$H = e^{-(\rho-n)t} u(c) + \mu[f(k) - (n+\delta)k - c], \tag{22.35}$$

and follow the same procedure adopted in the model of traditional optimal growth. Since the various steps have been detailed there, here we can proceed more synthetically. Thus we introduce the new costate variable

$$q = \mu e^{(\rho-n)t}, \tag{22.36}$$

and look for an interior maximum of the Hamiltonian

$$H = e^{-(\rho-n)t}\{u(c) + q[f(k) - (n+\delta)k - c]\}, \tag{22.37}$$

which turns out to exist when

$$q = u_c(c), \tag{22.38}$$

namely when the shadow price of per-capita capital accumulation equals the individual marginal utility of consumption.

The canonical equation for the original costate variable is

$$\mu' = -\frac{\partial H}{\partial k} = -\mu[f_k(k) - (n+\delta)]. \tag{22.39}$$

Since we are in a case in which the final function is absent, the boundary condition for μ is $\lim_{t\to\infty} \mu(t) = 0$. This is satisfied because integrating the differential equation (22.39) we have

$$\mu = M \exp\left(-\int [f_k(k) - (n+\delta)]dt\right), \tag{22.40}$$

where M is an arbitrary constant. This expression converges to zero since $f_k(k) > (n+\delta)$ as assumed above.

To transform the canonical equation for μ into a differential equation for q we observe that $\mu = qe^{-(\rho-n)t}$, hence $\mu' = q'e^{-(\rho-n)t} - (\rho-n)qe^{-(\rho-n)t}$. By substituting into Eq. (22.39) and rearranging terms we have

$$q' = -q[f_k(k) - (\rho+\delta)]. \tag{22.41}$$

Observing that logarithmic differentiation of the optimality condition (22.38) yields

$$q'/q = [u_{cc}(c)/u_c(c)]\, c' = -\sigma(c)c'/c, \tag{22.42}$$

we can rewrite Eq. (22.41) in terms of the control variable c. Thus we end up with the dynamic equations

$$\begin{aligned} \frac{k'}{k} &= \frac{f(k)}{k} - \frac{c}{k} - (n+\delta) = A + Bk^{\alpha-1} - \frac{c}{k} - (n+\delta), \\ \frac{c'}{c} &= \frac{1}{\cancel{\sigma(c)}}[f_k(k) - \delta - \rho] = \frac{1}{\theta}[A + B\alpha k^{\alpha-1} - \delta - \rho], \end{aligned} \tag{22.43}$$

where we have inserted the functional form (22.32) and assumed a constant elasticity of marginal utility ($\sigma(c) = \theta$). This latter assumption is inessential but simplifies the algebra.

Since we have assumed $A > \delta + \rho$, there is no positive value of k that yields a singular point of the second equation. Actually, the model generates endogenous growth thanks to this assumption. Both c and k grow without bounds, but their steady-state *growth rate* turns out to be $(1/\theta)(A - \delta - \rho)$. In fact, as $k \to \infty$, $k^{\alpha-1} \to 0$, and so, letting γ_c^* denote the steady-state growth rate of consumption, from the second equation of (22.43) we have

$$\gamma_c^* = (1/\theta)(A - \delta - \rho),$$

which implies

$$c(t) = c_0 e^{(1/\theta)(A-\delta-\rho)t}. \tag{22.44}$$

The finding that steady-state consumption per head is not a constant, but grows at a constant rate, is an essential result of optimal endogenous growth theory.

Substituting (22.44) into the first equation of (22.43), we have

$$k' = (A - n - \delta)k - c_0 e^{(1/\theta)(A-\delta-\rho)t}.$$

This is a simple first-order linear non-homogeneous differential equation, whose solution (see Chap 12, Sects. 12.1 and 12.2.2) is

$$k(t) = Ne^{(A-\delta-n)t} + [c_0/\varphi]e^{(1/\theta)(A-\delta-\rho)t}, \tag{22.45}$$

where N is an arbitrary constant, and

$$\varphi \equiv (A-\delta)(\theta-1)/\theta + \rho/\theta - n. \tag{22.46}$$

If we now apply the transversality condition (22.7) we get

$$\lim_{t\to\infty} k(t)\mu(t) = 0,$$

and observing that $\mu(t)$ is given by Eq. (22.40), where $f_k(k) \to A$ as t, and hence k, tend to infinity, we have

$$\lim_{t\to\infty}\left[k(t)Me^{-(A-\delta-n)t}\right] = M\lim_{t\to\infty}\left[k(t)e^{-(A-\delta-n)t}\right] = 0,$$

from which we obtain, account being taken of Eq. (22.45),

$$\lim_{t\to\infty}\left\{N + [c_0/\varphi]e^{[(1/\theta)(A-\delta-\rho)-(A-\delta-n)]t}\right\} = \lim_{t\to\infty}\left\{N+[c_0/\varphi]e^{-\varphi t}\right\} = 0. \tag{22.47}$$

This clearly requires $\varphi > 0$, a condition which is satisfied if the utility functional converges (see exercise 4). Equation (22.47) also requires $N = 0$. Hence Eq. (22.45) becomes

$$k(t) = [c_0/\varphi]e^{(1/\theta)(A-\delta-\rho)t}, \tag{22.48}$$

which shows that k grows at the same steady-state rate as c, and $c(t)/k(t) = \varphi$. Since in the limit we have $y = Ak$, also output per head grows at this same steady-state rate, and we have balanced growth.

From the mathematical point of view, all this shows that system (22.43) cannot be analysed in the (k, c) space, where it has no singular point. To analyse the transitional dynamics we must perform a transformation to variables that are constant in the steady state.

22.2.2.3 Optimal endogenous growth: saddle-point transitional dynamics

Let us consider the dynamic system (22.43) and perform the following transformation of variables

$$\begin{aligned} z &= \frac{f(k)}{k}, \\ \chi &= \frac{c}{k}. \end{aligned} \tag{22.49}$$

Note that z, the average productivity of capital, is a state-like variable as k, and χ, the ratio of consumption to capital, is a control-like variable as c.

If we substitute these definitions of z, χ into Eqs. (22.43) we have

$$\begin{aligned} \frac{k'}{k} &= z - \chi - n - \delta, \\ \frac{c'}{c} &= \frac{1}{\theta}(f_k(k) - \delta - \rho). \end{aligned} \tag{22.50}$$

From (22.49) we have, upon differentiation,

$$\begin{aligned} z' &= \frac{f_k(k)k'k - f(k)k'}{k^2} = f_k\frac{k'}{k} - \frac{f(k)}{k}\frac{k'}{k} = \left(f_k(k) - \frac{f(k)}{k}\right)\frac{k'}{k} \\ &= (f_k(k) - z)\frac{k'}{k}, \\ \chi' &= \frac{c'k - k'c}{k^2} = \frac{c'}{k} - \frac{k'}{k}\frac{c}{k} = \frac{c'}{c}\frac{c}{k} - \frac{k'}{k}\frac{c}{k} = \left(\frac{c'}{c} - \frac{k'}{k}\right)\chi, \end{aligned}$$

from which we immediately derive

$$\begin{aligned} \frac{k'}{k} &= (f_k(k) - z)^{-1}z', \\ \frac{c'}{c} &= \frac{\chi'}{\chi} + \frac{k'}{k}. \end{aligned} \tag{22.51}$$

If we substitute Eqs. (22.49) and (22.51) into (22.50) we obtain

$$\begin{aligned} z' &= (f_k(k) - z)(z - \chi - n - \delta), \\ \chi' &= \chi\left[-\frac{k'}{k} + \frac{1}{\theta}(f_k(k) - \delta - \rho)\right] = \chi\left[-z + \chi + n + \delta + \frac{1}{\theta}f_k(k) - \frac{\delta}{\theta} - \frac{\rho}{\theta}\right] \end{aligned} \tag{22.52}$$

We can now use the specification (22.33) of the production function and observe that

$$\begin{aligned} f_k(k) - z &= (A + B\alpha k^{\alpha-1}) - (A + Bk^{\alpha-1}) = -(1-\alpha)Bk^{\alpha-1} \\ &= -(1-\alpha)(z - A), \\ f_k(k) &= (f_k(k) - z) + z = A + \alpha(z - A). \end{aligned} \tag{22.53}$$

Substituting from Eqs. (22.53) into Eqs. (22.52) we have

$$
\begin{aligned}
z' &= -(1-\alpha)(z-A)(z-\chi-n-\delta),\\
\chi' &= \chi\left[-z+\chi+n+\delta+\frac{1}{\theta}A+\frac{\alpha}{\theta}(z-A)-\frac{\delta}{\theta}-\frac{\rho}{\theta}\right]\\
&= \text{(by adding and subtracting } A).\\
&= \chi\left[-(z-A)-(A-\delta)+\chi+n+\frac{1}{\theta}A+\frac{\alpha}{\theta}(z-A)-\frac{\delta}{\theta}-\frac{\rho}{\theta}\right]\\
&= \chi\left\{\chi-\left[(A-\delta)-\frac{1}{\theta}(A-\delta)+\frac{\rho}{\theta}-n\right]-(1-\frac{\alpha}{\theta})(z-A)\right\}\\
&= \chi\left\{\chi-\left[(A-\delta)\frac{\theta-1}{\theta}+\frac{\rho}{\theta}-n\right]-(1-\frac{\alpha}{\theta})(z-A)\right\},\\
&\text{hence}\\
\chi' &= \chi\left\{(\chi-\varphi)-\frac{\theta-\alpha}{\theta}(z-A)\right\},
\end{aligned}
\tag{22.54}
$$

where in the last passage we have used the definition of φ given above, Eq. (22.46).

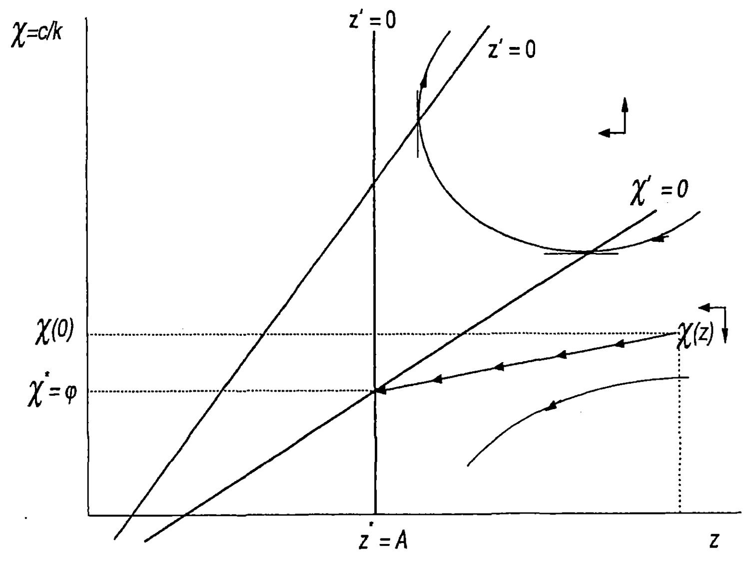


Figure 22.3: An optimal endogenous growth model: transitional dynamics

System (22.54) can be analysed in the usual way. From the first equation

we have $z' = 0$ for $z^* = A$ (the vertical line A in Fig. 22.3, taken from Barro and Sala-i-Martin) and for $\chi = z-(n+\delta)$, which is a straight line with slope 1 and negative intercept.

From the second equation of (22.54) we find that, apart from the trivial case $\chi = 0$, $\chi' = 0$ for $\chi = [(\theta-\alpha)/\theta]\,z + [\varphi - A(\theta-\alpha)/\theta]$, which is a straight line with positive slope (as shown) if $\theta > \alpha$. The opposite case is ruled out because (as Barro and Sala-i-Martin observe) it would require an unrealistically high degree of intertemporal substitution, since θ would have to be significantly below unity. Assuming then $\theta > \alpha$, the positive slope of the $\chi' = 0$ line is also smaller than unity.

We now observe that, when $z = A$, the $\chi' = 0$ line takes on the value $\chi^* = \varphi$, whereas the straight line $\chi = z-(n+\delta)$ (corresponding to $z' = 0$) takes on the value $A-(n+\delta)$. It is easy to check that $A-(n+\delta) > \varphi$ given that $A > \rho+\delta$, as assumed from the beginning. Hence at $z = A$ the $z' = 0$ line lies above the $\chi' = 0$ line, as drawn in the figure.

Let us also note that the only relevant region of the diagram is where $z \geq A$. In fact, $z = A + Bk^{\alpha-1} \geq A$.

Given these preliminaries, the only relevant singular point of the system is $z^* = A$, $\chi^* = \varphi$. The arrows can be drawn in the usual way. Consider, for example, point $(z(0), \chi(0))$. It is to the right of $z^* = A$, hence—by the first equation of system (22.54)—$z' < 0$ and z is decreasing. It is also to the right of the straight line $\chi = [(\theta-\alpha)/\theta]\,z + [\varphi - A(\theta-\alpha)/\theta]$, hence $\chi' < 0$ and χ is decreasing as well.

The construction of the arrows hints at the usual saddle point, which can be proven rigorously by performing a linearisation of system (22.54) at the equilibrium point. We rewrite this system here for the reader's convenience

$$
\begin{aligned}
z' &= -(1-\alpha)(z-A)(z-\chi-n-\delta) &&= \Phi_1(z,\chi),\\
\chi' &= \chi\left\{(\chi-\varphi)-\frac{\theta-\alpha}{\theta}(z-A)\right\} &&= \Phi_2(z,\chi).
\end{aligned}
$$

It is then easy to calculate

$$
\frac{\partial\Phi_1}{\partial z} = -(1-\alpha)\left[(z-\chi-n-\delta)+(z-A)\right], \quad \frac{\partial\Phi_1}{\partial\chi} = (1-\alpha)(z-A),
$$

$$
\frac{\partial\Phi_2}{\partial z} = -\chi\frac{\theta-\alpha}{\theta}, \quad \frac{\partial\Phi_2}{\partial\chi} = \left[(\chi-\varphi)-\frac{\theta-\alpha}{\theta}(z-A)\right]+\chi.
$$

By evaluating these partial derivatives at the equilibrium point $z = A, \chi = \varphi$ we obtain the linearised system

$$
\begin{aligned}
\overline{z}' &= -(1-\alpha)\left\{[A-(n+\delta)]-\varphi\right\}\overline{z},\\
\overline{\chi}' &= -\varphi\frac{\theta-\alpha}{\theta}\overline{z}+\varphi\overline{\chi}.
\end{aligned}
\tag{22.55}
$$

We have already shown above that $\theta - \alpha > 0$ and that $A - (n + \delta) > \varphi$. Hence the succession of signs in the coefficients of the characteristic equation of system (22.55) is $+?-$, which means two real roots with opposite sign. This shows that the singular point is indeed a saddle point (see Chap. 21, Table 21.1).

For any given initial average product of capital, $z(0)$, an appropriate initial value of the ratio of consumption (the control variable) to capital, $\chi(0)$, has to be chosen on the stable arm of the saddle, as shown in the diagram. Again we note that, from the mathematical point of view, choosing $\chi(0)$ in correspondence to $z(0)$ on the stable arm means choosing the initial conditions in such a way that the arbitrary constant appearing in the term containing the unstable root turns out to be zero, as shown in Chap. 18, Sect 18.2.2.3.

Along the stable path both the average and the marginal productivity of capital decline towards their steady state value. This generates falling growth rates of per capita variables, so that the model exhibits the convergence property of the traditional neoclassical model together with endogenous growth.

In this simple model physical and human capital have been aggregated. They can of course be kept distinct, but if an identical production function for both of them is assumed, there is not much to be gained by this distinction, that on the contrary becomes crucial when different production functions are introduced (see Barro and Sala-i-Martin, 1995, Chap. 5).

22.3 Rational expectations and saddle points

22.3.1 Introduction

Let us consider an agent who forms expectations on the future value of a variable according to the rational expectations hypothesis (REH). As is well known, rational expectations, introduced by Muth (1960, 1961), mean that the forecasting agent uses all available information, including the knowledge of the "true" model that determines the evolution of the variable(s) concerned. We recall that expectations formed according to REH have the following properties:

a) There are no *systematic* forecast errors, as these errors are stochastically distributed with mean zero (*unbiasedness).* If $E_\tau[r(t)]$ denotes the expected value of the variable r (expectations are formed at time $\tau < t$ with reference to time t), then we have $E_\tau[r(t)] - r(t) = \epsilon(t)$, where $E[(\epsilon(t)] = 0$ for all t.

b) Forecast errors are *uncorrelated* over time, i.e. $E[\epsilon(t)\epsilon(s)] = 0$ for any $s \neq t$, where s denotes another point in time and has the same domain as t.

c) Forecast errors are not correlated with the past history of the variable being forecasted nor with the other variables contained in the information set, I_τ, available at the time of forecast (*orthogonality*, i.e. $E(\epsilon_t \mid I_\tau) = 0$). This property implies that RE are statistically *efficient,* namely the variance of the forecast error is lower than that of forecasts obtained with any other method (minimum variance property).

We have already met rational expectations in this book—see Chap. 6, Sect. 6.2 and Sect. 6.3, exercise 6.3(5.e). In a deterministic context, such as the one we are dealing with here, rational expectations imply perfect foresight, namely $E_\tau[r(t)] - r(t) = 0$, or $E_\tau[r(t)] = r(t)$.

A natural question that arises in dynamic models containing deterministic rational expectations is why possible disequilibrium situations are not immediately corrected by an appropriate behaviour of the relevant economic agents who all know where the equilibrium point is. The main reason lies in the distinction between *jump (*or *jumping)* variables and other variables, also called *predetermined* or *historical* variables.

At any given point in time, there are certain variables whose value cannot be changed instantaneously by any action of the economic agents (hence they are predetermined from the agents' point of view). This impossibility may derive from various reasons, the most obvious being the intrinsic nature of the variable. In a closed economy the capital stock, for example, can be increased only through net investment, hence at time t the amount of capital available is $K(t)$, and even if we know that the equilibrium value of the capital stock is $K^* > K(t)$, we cannot 'jump' from $K(t)$ to K^*.

On the contrary, there are certain variables that can instantaneously take on any possible value in their definition set. The most obvious case is that of asset prices: in a free market economy, such prices can instantaneously drop to zero or increase by a huge amount. A jump variable (sometimes also called costate variable, by analogy with optimal control problems) is not continuous at the point in time where it jumps, but this is not important provided that the variable is *right-continuous*, and that it possesses a *right-hand-side* derivative.

A generic variable expressed as a function of time $y(t)$ is said to be *right-continuous at point* $\bar{t}$ when $y(t)$ approaches $y(\bar{t})$ as t approaches $\bar{t}$ from above, i.e. from the future (this is one of the reasons why jump variables are often called *forward-looking* variables). More formally,

$$\lim_{t \to \bar{t}, t > \bar{t}} y(t) = y(\bar{t}). \tag{22.56}$$

An equivalent notation is

$$\lim_{t \downarrow \bar{t}} y(t) = y(\bar{t}).$$

When $y(t)$ is right-continuous at any point in its interval of definition, then it is called a right-continuous function of time. This means that, although the variable y can jump at any point of time, it possesses property (22.56) at the jump.

Similarly we define the *right-hand-side derivative* of a right-continuous function $y(t)$ as

$$\lim_{t \to \bar{t}, t > \bar{t}} \frac{y(t) - y(\bar{t})}{t - \bar{t}}. \qquad (22.57)$$

The reason for being content with right-continuity and right-hand-side differentiability is that we are interested in the dynamics of the system *after* the jump, namely from the jump onwards, hence we are not disturbed by the fact that the jump variable is not differentiable nor continuous with respect to time at the instant of the jump.

By and large, predetermined variables have the dimension of stocks that can be changed only through flows, whereas jump variables have the dimension of prices in the wide sense (i.e., including interest rates, exchange rates, asset prices, etc.).

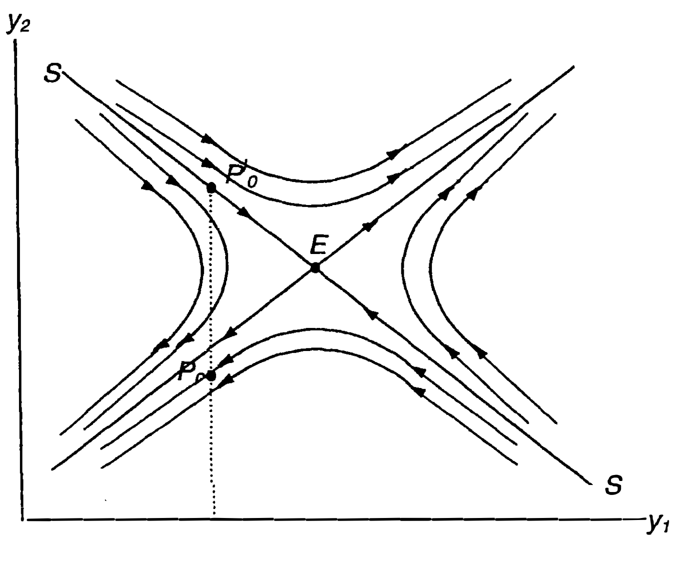


Figure 22.4: Rational expectations and saddle path dynamics: the overshooting phenomenon

The presence of both jump and predetermined variables explains why also rational expectations models need a transitional dynamics, which—in the 2×2 case with one predetermined and one jump variable—usually turns out to be of the saddle path type. This additionally explains the *overshooting*

phenomenon, that can occur when the situation is like that depicted in Fig. 22.4. Suppose that y_2 is a jump variable and y_1 a predetermined variable. Given the saddle point nature of the equilibrium point E, if the stable arm is SS and the initial point P_0, the only rational course of action is to let y_2 jump onto the stable arm, after which the economic system will continue to move, through its intrinsic dynamics, from P_0' to E on the stable arm. This jump is exactly what agents holding rational expectations will cause through their behaviour on the market, as they know that any other course of action would put the system on an unstable path.

By so doing, the jump variable, starting from a value *smaller* than its equilibrium value, will initially jump to a value *higher* than equilibrium (i.e., will *overshoot* it), after which it will gradually decrease toward equilibrium along the stable arm. It goes without saying that the same reasoning can lead to an *undershooting* if the initial point P_0 lies on the opposite side of the equilibrium point.

Since a saddle point exists in the case of one positive and one negative real root (see Table 21.1) it follows that *from the mathematical point of view, the jump from P_0 to P_0' is equivalent to selecting the initial conditions so that the constant associated with the unstable root turns out to be zero,* as explained in Sect. 18.2.2.3.

We shall now illustrate these points by means of a well-known model of exchange-rate overshooting, after which we shall treat the generalisation to $n > 2$ variables of the saddle point property of rational expectations models.

22.3.2 Rational expectations, saddle points, and overshooting

Let us consider, following Dornbusch (1976, 1980), a small open economy under flexible exchange rates with perfect capital mobility and flexible prices but with given full-employment output. We note at the outset that the exchange rate is a typical jump variable, since it is free to make jumps in response to 'news'. On the contrary, commodity prices are assumed to be predetermined variables, namely they are assumed to adjust slowly to their long-run equilibrium value (hence the denomination of 'sticky price' monetary model of exchange rate determination).

Perfect capital mobility coupled with perfect asset substitutability implies that the domestic (r) and foreign (r^*) nominal interest rates are related by *uncovered interest rate parity* (UIP) (see Gandolfo, 1995, Chap. 10, Sect. 10.7), namely

$$r = r^* + x, \qquad (22.58)$$

where x is the expected rate of depreciation of the exchange rate.

Since we are in a small open economy, the foreign interest rate is taken as given. As regards expectations, we observe that in our deterministic context rational expectations imply perfect foresight, so that the expected and actual rate of depreciation of the exchange rate coincide. Letting e denote the logarithm of the spot exchange rate E, we have

$$e' = \mathrm{d}\ln E/\mathrm{d}t = E'/E, \tag{22.59}$$

hence e' is the *actual* rate of depreciation of the nominal exchange rate E. By setting $x = e'$ in Eq. (22.58) we have

$$r = r^* + e', \tag{22.60}$$

where of course e' has to be interpreted as a right-hand-side time derivative, given the nature of jump variable of the exchange rate.

Let us now consider money market equilibrium,

$$\frac{M}{P} = \mathrm{e}^{-\lambda r} Y^{\phi},$$

where M is the money supply, P the price level, λ the semi-elasticity of money demand with respect to the interest rate, and ϕ the elasticity of money demand with respect to income Y. Taking logs (lower-case letters denote the logs of the corresponding upper-case-letter variables) we have

$$-\lambda r + \phi y = m - p. \tag{22.61}$$

Combining Eq. (22.61) with the UIP condition incorporating rational expectations—Eq. (22.60)—we obtain the condition for *asset market equilibrium*

$$p - m = -\phi\overline{y} + \lambda r^* + \lambda e', \tag{22.62}$$

where output has been taken as given at its full-employment level $\overline{y}$.

In long-run equilibrium with a stationary money supply, we have

$$\overline{p} - m = -\phi\overline{y} + \lambda r^*, \tag{22.63}$$

since actual and expected depreciation are assumed to be zero in long-run equilibrium. Equation (22.63) determines the long-run equilibrium price level.

Subtracting Eq. (22.63) from Eq. (22.62) we obtain:

$$p - \overline{p} = \lambda e',$$

or

$$e' = \frac{1}{\lambda}(p - \overline{p}). \tag{22.64}$$

This is one of the key equations of the model, as it expresses the dynamics of the current spot exchange rate in terms of the deviations of the current price level from its long-run equilibrium level.

We now turn to the goods market, that—given the assumption of a constant level of output—will serve to determine the price level.

Since output is given, excess demand for goods will cause an increase in prices. We first assume that aggregate demand for domestic output depends on the relative price of domestic goods with respect to foreign goods, $[(e + p^*) - p]$, the interest rate, and real income. Thus we have

$$\ln D \equiv d = u + \delta(e - p) + \gamma y - \sigma r, \tag{22.65}$$

where u is a shift parameter and the given foreign price level P^* has been normalised to unity (hence $p^* = \ln P^* = 0$). The usual dynamic (Walrasian) assumption that prices move in response to excess demand yields, account being taken of the logarithmic setting,

$$p' = \pi(d - y), \quad \pi > 0, \tag{22.66}$$

from which

$$p' = \pi[u + \delta(e - p) + (\gamma - 1)y - \sigma r]. \tag{22.67}$$

In long-run equilibrium, $p' = 0$ and $p = \overline{p}$. With these settings in Eq. (22.67) we finally obtain the long-run equilibrium exchange rate

$$\overline{e} = \overline{p} + (1/\delta)[\sigma r^* + (1 - \gamma)y - u], \tag{22.68}$$

where $\overline{p}$ is given by Eq. (22.63) and $r = r^*$ since the long-run equilibrium exchange rate is expected to remain constant.

Let us now turn back to the disequilibrium dynamics of Eq. (22.67). Solving Eq. (22.61) for the interest rate and substituting into Eq. (22.67) we have

$$\begin{aligned} p' &= \pi[u + \delta(e - p) + \frac{\sigma}{\lambda}(m - p) - \rho\overline{y}], \\ \rho &\equiv \phi\sigma/\lambda + 1 - \gamma. \end{aligned} \tag{22.69}$$

We now subtract from the r.h.s. of Eq. (22.69) its long-run equilibrium value, which is zero, namely $0 = \pi[u + \delta(\overline{e} - \overline{p}) + \frac{\sigma}{\lambda}(m - \overline{p}) - \rho\overline{y}]$. Since $p' = (p - \overline{p})'$ for constant $\overline{p}$, we can thus express price dynamics in terms of deviations from the long-run equilibrium values of the exchange rate and the price level, that is

$$p' = -\pi(\delta + \frac{\sigma}{\lambda})(p - \overline{p}) + \pi\delta(e - \overline{e}). \tag{22.70}$$

Equations (22.70) and (22.64) govern the dynamics of our system, which is already in (log)linear form. The characteristic equation of this system is

$$\begin{vmatrix} -\pi(\delta+\frac{\sigma}{\lambda})-\mu & \pi\delta \\ \frac{1}{\lambda} & 0-\mu \end{vmatrix} = \mu^2+\pi(\delta+\frac{\sigma}{\lambda})\mu-\frac{\pi\delta}{\lambda}=0, \qquad (22.71)$$

where μ denotes the latent roots. Since the succession of the signs of the coefficients is $++-$, there will be two real roots, one negative and the other one positive, which means a saddle point (see Chap. 21, Table 21.1).

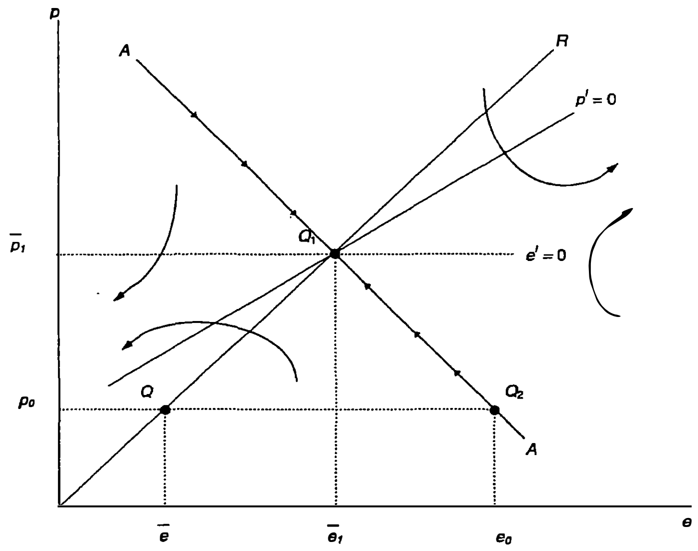


Figure 22.5: Rational expectations and exchange-rate overshooting

The phenomenon of exchange-rate overshooting in response to 'news', i.e. to unanticipated events such as a monetary shock, can be examined by means of Fig. 22.5. Suppose that the economic system is at its long-run equilibrium (point Q). By an appropriate choice of units we can set $\overline{p}=\overline{e}$, so that OR is a 45° line. Suppose now that the nominal money stock increases permanently. Economic agents immediately recognize that the long-run equilibrium price level and exchange rate will increase in the same proportion, as in the long-run money is neutral. In terms of Fig. 22.5 this means that economic agents recognize that the economy will move from Q to Q_1.

The economy, however, cannot instantaneously jump from Q to Q_1 because prices have been assumed to move only gradually. Hence the dynamic behaviour of the economic system will be governed by the differential equations (22.70) and (22.64). The $e'=0$ and $p'=0$ loci have been drawn with reference to the new long-run equilibrium point. Note that the slope of the $p'=0$ line is smaller than one because, given the normalization $\overline{p}=\overline{e}$ (and, of course, $\overline{p}_1=\overline{e}_1$), setting $p'=0$ in Eq. (22.70) and computing the slope of the resulting expression we have

$$(\mathrm{d}p/\mathrm{d}e)_{p'=0}=\frac{\pi\delta}{\pi\delta+\pi\sigma/\lambda}<1. \qquad (22.72)$$

We know that the (new) long-run equilibrium point is a saddle point. The stable arm of the saddle is the straight line AA. It is downward sloping because—as shown in Chap. 21, Sect. 21.3.2.4, Eq. (21.17)—its equation is

$$(e-\overline{e})=\frac{\mu_1-(\pi\delta+\pi\sigma/\lambda)}{\pi\delta}(p-\overline{p}), \qquad (22.73)$$

where μ_1 is the stable root of the characteristic equation (22.71). Since the sum of the roots is minus the trace, we have $\mu_1+\mu_2=(\pi\delta+\pi\sigma/\lambda)$, hence $\mu_1-(\pi\delta+\pi\sigma/\lambda)=-\mu_2<0$ since μ_2 is the positive root.

We have seen that the economy cannot jump from Q to Q_1. But the exchange rate is a jump variable, hence the economy can jump from Q to Q_2 on the stable arm of the saddle. From the economic point of view, the current exchange rate will depreciate because an increase in the money supply, given the stickiness of prices in the short run, will cause an increase in the real money supply and hence a fall in the domestic nominal interest rate. Since the UIP condition is assumed to hold instantaneously, and the nominal foreign interest rate is given, the exchange rate immediately depreciates by more than the increase in the long-run equilibrium value to create the expectation of an appreciation. This is required from the UIP condition that the interest-rate differential equals the expected rate of appreciation: in fact, given $r=r^*$ in the initial long-run equilibrium, the sudden decrease in r requires $x<0$—see Eq. (22.58)—namely an anticipated appreciation.

Thus the exchange rate initially overshoots its (new) long-run equilibrium level ($e_0>\overline{e}_1$), after which it will gradually appreciate alongside with the increase in the price level following the path from Q_2 to Q_1 on the AA line. Overshooting results from the requirement that the system possesses 'saddle-path' stability.

22.3.3 Rational expectations and saddle points: the general case

In the previous section we have dealt with the simple 2×2 case, with one predetermined and one jump variable. It is however possible to generalise the treatment to the case of $n > 2$ variables. In the general case the saddle point property will continue to hold *provided that there are as many predetermined variables as stable roots.*

We shall give a heuristic proof which is fairly simple (thanks to the fact that it builds on material previously treated at length in this book), referring the more mathematically sophisticated reader to Buiter (1984).

When we are in the presence of a model with rational expectations, we can always reduce it to a model in terms of actual values of the variables and their time derivatives by using the two basic assumptions of (deterministic) rational expectations: perfect information (including the knowledge of the model at hand) and perfect foresight. Knowledge of the model under consideration means that any (long-run) equilibrium condition implied by the (static part of the) model can be used in the relevant equations containing expectations. Perfect foresight means that expected values can be set equal to actual observed values. The 2×2 model examined in the previous section exemplifies both points quite well.

After doing this, we are left with a dynamic system that (after linearisation and reduction to first-order normal form if the case) will be of the type

$$\mathbf{y}' = \mathbf{A}\mathbf{y}, \tag{22.74}$$

where $\mathbf{y}$ is a vector of *observed endogenous variables*, usually expressed as deviations from the (long-run) equilibrium values. If the initial system also contains exogenous variables, we have

$$\mathbf{y}' = \mathbf{A}\mathbf{y} + \mathbf{g}(t), \quad \mathbf{g}(t) \equiv \mathbf{B}\mathbf{z}(t), \tag{22.75}$$

where $\mathbf{B}$ is a matrix of constants and $\mathbf{z}(t)$ a vector of exogenous variables, possibly containing expected values. We then assume perfect foresight as regards the exogenous variables as well, and consider the deviations from the particular solution of the non-homogeneous system, which amounts to examining a system like (22.74).

The solution of systems like (22.74) and (22.75) has been treated at length in Chap. 18, Sect. 18.2, to which we refer the reader. Note in particular that in the case of the non-homogeneous system we shall have to use the method of variation of parameters (Chap. 18, Sect. 18.2.3.1), since the analytic form of the $z_i(t)$ is usually unknown.

Suppose now that the matrix $\mathbf{A}$ has n distinct latent roots, partly stable and partly unstable. Let $0 < k < n$ be the number of stable roots. We can then use the procedure explained in Chap. 18, Sect. 18.2.2.3, to eliminate the terms containing the unstable roots from the general solution of the system under consideration. As shown there, this procedure will be successful if we are free to choose as many initial conditions as there are unstable roots, i.e. $(n - k)$.

Since the initial conditions that we are free to choose are those concerning the jump variables, we conclude that *the system will exhibit saddle-path stability if there are as many jump variables as unstable roots* or, equivalently, if there are as many predetermined variables as stable roots, which proves our initial statement.

If there are *fewer* stable roots than predetermined variables, it is not possible to make the system stable for arbitrary initial values of the predetermined variables. On the contrary, if there are *more* stable roots than predetermined variables, the convergent solution will no longer be unique. We can of course ensure uniqueness provided that we have a sufficient number of additional conditions, for example linear restrictions on the initial conditions. For details see Buiter (1984).

22.4 Exercises

1. Ramsey (1928) argued, on the basis of ethical considerations, that the time horizon should be infinity and that there should be no discounting, that is, the utility of future generations should have the same importance as the utility of the current generation. In this case, however, the infinite integral

$$\int_{t_0}^{\infty} U(c)\mathrm{d}t$$

clearly does not converge. Hence Ramsey devised the following way out: there is a fixed finite upper limit to utility, that he called *bliss*:

$$B = \max U(c) = U(c_B),$$

where c_B, the bliss per capita consumption, is finite. Hence the problem can be cast in terms of minimizing a *regret* function expressing the deviations of utility from the bliss level:

$$\min R = \int_{t_0}^{\infty} [B - U(c)]\mathrm{d}t.$$

By the way, this is an extraordinarily modern approach: modern optimal control theory as applied for example to policy optimization is typically concerned with the minimisation of an objective functional expressing the *deviations* of the targets from some given 'first best' level (see, for example, Petit, 1990, Chaps. 5 and 6).

Now observe that $\min R = \max(-R)$, and solve the Ramsey optimal control problem by the maximum principle. Show that the solution gives the Ramsey (or Keynes-Ramsey) rule: 'rate of saving multiplied by marginal utility of consumption should always equal bliss minus actual rate of utility enjoyed'. Compare the results with those obtained in the model with discounting.

2. Consider the optimal growth problem with finite terminal time

$$\max_{\{c(t)\}} W = \int_{t_0}^{t_1} e^{-\rho(t-t_0)} U(c(t))\mathrm{d}t,$$

with the additional terminal condition $k(t_1) \geq k_1 > 0$ (this is required to take into account that society will consume also after t_1, so that a positive terminal capital must be left over).

(2.a) Show that the solution of the problem yields the same canonical equations as in the infinite horizon case, but there is the following terminal condition on the costate variable:

$$e^{-\rho(t_1-t_0)}\mu(t_1)[k(t_1) - k_1] = 0,$$

namely either the condition on terminal capital is satisfied as an equality or the terminal shadow price of capital is zero.

(2.b) Show that the optimal growth path satisfies the so-called *turnpike property*: for t_1 that becomes sufficiently great, the optimal paths of per capita capital and per capita consumption spend an arbitrarily large fraction of the time near the balanced optimal growth path. In particular, per capita capital starts from k_0, approaches $k^*$and remains there, then in the final part of its 'journey' leaves the neighbourhood of k^* to move towards k_1.

3. Let us consider the stable arm of the linear approximation (22.28), which corresponds to the negative root of the characteristic equation (22.29), that we name $-\beta$ according to the notation of Barro and Sala-i-Martin. Then the solution for $\overline{k}$ will be

$$\overline{k} = k - k^* = (k_0 - k^*)e^{-\beta t}.$$

Dividing through by k^* we have

$$\frac{k-k^*}{k^*} = \frac{k_0-k^*}{k^*}e^{-\beta t}.$$

Now observe that, according to elementary approximation formulae, $\frac{k-k^*}{k^*} \simeq \ln\left(\frac{k}{k^*}\right)$, $\frac{k_0-k^*}{k^*} \simeq \ln\left(\frac{k_0}{k^*}\right)$. The approximation is fairly good when $(k - k^*)$ and $(k_0 - k^*)$ are small, which is indeed the case since we considering a neighbourhood of the equilibrium point. Thus we have

$$\ln\left(\frac{k}{k^*}\right) = \ln\left(\frac{k_0}{k^*}\right)e^{-\beta t},$$

from which

$$\ln k = (1 - e^{-\beta t})\ln k^* + e^{-\beta t}\ln k_0.$$

Derive the log-linearised solution for $\ln y$ and examine the speed of convergence problem along the lines of Sect. 13.2.4.1.

4. Consider the endogenous growth model of Sect.22.1.3.2 and let $c(t)$ be given by Eq. (22.44). Show that the utility functional

$$\int_0^\infty e^{-(\rho-n)t}u(c(t))\mathrm{d}t$$

converges if $\rho + \delta > [(1-\theta)/\theta](A - \delta - \rho) + n + \delta$. (Hint: with a constant elasticity of marginal utility, we have $u(c) = [c^{(1-\theta)} - 1]/(1-\theta)$). Also show that this inequality ensures that $\varphi > 0$.

5. Consider the following model (a simplification of the model by Buiter and Miller, 1981)

$$\begin{array}{rll} m & = k(y+R) - \lambda r + p & \text{LM schedule} \\ y - \overline{y} & = \delta(e-p) + \psi R_\infty & \text{IS schedule} \\ p' & = \theta(y - \overline{y}) + \pi & \text{Phillips curve} \\ \pi & = m' & m' = \text{constant} \\ e' & = r - r^* & \text{UIP} \end{array}$$

where $k, \lambda, \delta, \psi, \theta$ are positive constants, R is oil production expressed as a fraction of real non-oil income (y), $R_\infty < R$ is the permanent income equivalent of R, π is the trend or 'core' inflation rate taken as equal to the rate of growth of the money supply, e is the exchange rate. The variables m, y, p, π, e are expressed as logarithms. Introducing the transformed variables $u = m - p$ (the real money supply) and $c = e - p$ (the real exchange rate) show that the resulting system exhibits saddle point stability.

6. In the model of exercise 5, suppose that initially there is no oil ($R_\infty = R = 0$) and the economy is in equilibrium. Suppose now that there is an unanticipated oil discovery, which is assumed undepletable and yielding a constant output per period, $R = R_\infty$. The price level is sluggish but the exchange rate is a jump variable. Determine the time path of the system and examine whether there is an 'overshooting' phenomenon.

22.5 References

Barro, R.J. and X. Sala-i-Martin, 1995, *Economic Growth,* New York, McGraw-Hill.

Buiter, W.H., 1984, Saddlepoint Problems in Continuous-Time Rational Expectations Models: A General Method and some Macroeconomic Examples.

Buiter, W.H. and M. Miller, 1981, Monetary Policy and International Adjustment: The Problems of Adjustment.
Dornbusch, R., 1976, Expectations and Exchange Rate Dynamics.
Dornbusch, R., 1980, *Open Economy Macroeconomics*, Chap. 11.
Gandolfo, G., 1995, *International Economics II: International Monetary Economics and Open-Economy Macroeconomics.*
Intriligator, M.D., 1971, *Mathematical Optimization and Economic Theory.*
Jones, L.E. and R.E. Manuelli, 1990, A Convex Model of Equilibrium Growth: Theory and Policy Implications.
Kamien, M.I. and N.L. Schwartz, 1991, *Dynamic Optimization.*
Muth, J.F., 1960, Optimal Properties of Exponentially Weighted Forecasts.
Muth, J.F., 1961, Rational Expectations and the Theory of Price Movements.
Petit, M.L., 1990, *Control Theory and Dynamic Games in Economic Policy Analysis.*
Ramsey, F.P., 1928, A Mathematical Theory of Saving.
Samuelson, P.A., 1972, The General Saddle Point Property of Optimal Control Motions.

Chapter 23

Liapunov's Second Method

23.1 General concepts

We have examined the concept of stability in Chap. 21 at some length. The notion of stability in which we are interested here, is uniform asymptotic stability in the large, or *global* stability in brief (see Chap. 21, Definitions 21.3 and 21.6).

In general, Liapunov's second method, also called the 'direct' method, serves to answer questions of stability of differential or difference systems (Liapounoff, 1907)[1] *without solving the system.* On the other hand, his 'first method' (also called the 'indirect' method) consists of finding the closed-form solution of the system and utilizing it to check whether the motion of the system does converge to the equilibrium point whilst respecting the various conditions of the definition of stability adopted.

The advantage of the direct method over the indirect one lies in the fact that we do not need to know the solution of the system. Indeed, only in some special cases—among which the linear and constant coefficients case is the easiest to handle—are we able to find the solution of a differential or difference system. When the system is non-linear, it is often impossible to find its solution, and to know that the solution exists because the existence theorem tells us so is a meagre solace. It is true, of course, that we may always try a numerical integration of the system (computers will do the job), but this possibility is not of great help to the economic theorist.

In economic theory one seeks general answers, independent of numerical

[1]Liapunov's original memoir was punlished in Russian in 1893; a French translation (Problème général de la stabilité du mouvement) appeared in 1907 in *Annales de la Faculté des Sciences de Toulouse,* and has been reprinted in facsimile in *Annals of Mathematics Studies* No. 17, Princeton University Press, 1949; see especially pp. 255-67. An English translation appeared in 1966.
It must be noted that in Liapunov's original memoir only systems of differential equations are treated. The extension of the 'second method' to systems of difference equations has been made successively.

analyses which hold only for the single case, and so one works with functions which are specified only *qualitatively* (e.g., one assumes that $y = f(x)$, where $f'(x)$ (and perhaps $f''(x)$ too) has a certain sign, and that is all). Moreover, sometimes the problems are such that even if the analyst were ready to use numerically specified functions, it would be impossible to find them empirically (think of the problem of the stability of general equilibrium).

When we are confronted with an insoluble non-linear system, then, if we want to examine its stability by the 'first' method we must make a linear approximation at the equilibrium point (see Chap. 21, Sect. 21.4.1), and so the stability results that we obtain will hold only in a sufficiently small neighbourhood of that point, i.e. they will be only local stability results. If we want global stability results we cannot use the linear approximation, and it is here that the 'second' method proves particularly useful. It is, however, advisable to inform the reader now that, as far as the second method is concerned, only some general theorems exist. There are no hard and fast rules according to which we are sure to find, if it exists, a 'Liapunov function' on which the method is based. It is perhaps safe to say that much depends on the ingenuity of the user.

Let us note, incidentally, that in Parts I and II we have always used the first method, either because the models treated there were such as to give rise to systems of linear difference or differential equations with constant coefficients or because, when the models were non-linear, we used the linear approximation method.

23.2 The fundamental theorems

We shall examine only *autonomous* systems, i.e. systems of differential or difference equations of the type

$$\frac{dy_i}{dt} = f_i(y_1, y_2, ..., y_n), \quad i = 1, 2, ..., n, \tag{23.1}$$

$$y_{i_{t+1}} = f_i(y_{1_t}, y_{2_t}, ..., y_{n_t}), \quad i = 1, 2, ...n, \tag{23.2}$$

and not the more general systems ($u_i(t)$ = known functions, also called forcing functions)

$$\frac{dy_i}{dt} = f_i(y_1, y_2, ..., y_n, t, u_i(t)), \quad i = 1, 2, ..., n, \tag{23.3}$$

$$y_{i_{t+1}} = f_i(y_{1_t}, y_{2_t}, ..., y_{n_t}, t, u_i(t)), \quad i = 1, 2, ..., n. \tag{23.4}$$

Autonomous systems are those mainly used in the economic applications of the second method. For a treatment of non-autonomous systems, see the works cited in footnote 3.

The functions f_i have the following properties[2], respectively in the case of system (23.1) and of system (23.2):

$$f_i(y_1, y_2, ..., y_n) = 0 \quad \text{for} \quad y_i = y_i^e, \quad i = 1, 2, ..., n. \tag{23.5}$$

$$f_i(y_{1_t}, y_{2_t}, ...y_{n_t}) = y_i^e \quad \text{for} \quad y_{i_t} = y_i^e, \quad i = 1, 2, ..., n. \tag{23.6}$$

Sometimes the system is built in such a way that the variables y_i represent the *deviations* of other variables from an equilibrium point, i.e. $y_i = x_i - x_i^e$, all i. In this case it is obvious that $y_i^e = 0$ (the so-called 'null solution').

Before going any further, a review of the meaning of '*distance*' in mathematics is necessary.

The *distance* from the origin of a point in an n-dimensional metric space (or, if we prefer, the *norm* of the vector $\mathbf{y} = (y_1, y_2, ..., y_n)$, i.e. of the vector whose elements are the coordinates of the point) is defined as any scalar function of the variables $y_1, y_2, ..., y_n$ which has some specific properties. Such properties are (in what follows the distance will be indicated as $D(\mathbf{y})$. Equivalent notation is $\|\mathbf{y}\|$):

(1) $D(\mathbf{y}) > 0$ if $\mathbf{y} \neq \mathbf{0}$, i.e. if at least one of the numbers $y_1, y_2, ..., y_n$ is not zero;

(2) $D(\mathbf{y}) = 0$ if, and only if, $\mathbf{y} = \mathbf{0}$, i.e. if, and only if, $y_i = 0$ for all i;

(3) $D(\mu\mathbf{y}) = |\mu| D(\mathbf{y})$ for any constant μ;

(4) $D(\mathbf{y}' + \mathbf{y}'') \leq D(\mathbf{y}') + D(\mathbf{y}'')$, where $\mathbf{y}' = (y_1', y_2', ..., y_n')$ and $\mathbf{y}'' = (y_1'', y_2'', ..., y_n'')$ are any two points (vectors).

The concept of distance, and its properties, does not change if instead of the origin we refer to any other point, say $\mathbf{P}^e$. In such a case, in the place of the elements $(y_1, y_2, ..., y_n)$ we can substitute the differences $(y_1 - y_1^e, y_2 - y_2^e, ..., y_n - y_n^e)$ and reinterpret properties (1) to (4) in terms of these differences.

The functions which satisfy the stated properties are theoretically infinite; a list of the more usual ones is given below:

(a) the Euclidean distance, $D(\mathbf{y}) = +(y_1^2 + y_2^2, + ... + y_n^2)^{1/2}$. This is perhaps the first function which everyone immediately thinks of, since it measures the length of the straight line segment joining the point with the origin (the 'length' or 'modulus' of the vector $\mathbf{y}$);

(b) the modified Euclidean distance, $D(\mathbf{y}) = +(a_1 y_1^2 + a_2 y_2^2 + ... a_n y_n^2)^{1/2}$, where the a_i are given positive constants;

(c) the absolute value distance, $D(\mathbf{y}) = \sum_{i=1}^{n} h_i |y_i|$, where the h_i are given positive constants;

[2] We assume also that the f_i have all the properties needed for the solution of system (23.1) or (23.2) to exist and be unique.

(d) the 'maximum' distance, $D(y) = \max_i c_i |y_i|$, where the c_i are given positive constants.

The reader may check as a simple exercise that all the stated functions satisfy properties (1) to (4) above.

We now state the fundamental theorems.

Theorem 23.1. Consider the autonomous system

$$\frac{dy_i}{dt} = f_i(y_1, y_2, ..., y_n) \ , \qquad i = 1, 2, ..., n \ ,$$

where

$$f_i = 0 \qquad \text{for} \qquad y_i = y_i^e \ , \qquad i = 1, 2, ..., n \ .$$

Suppose there exists a scalar function

$$V(y_1 - y_1^e, y_2 - y_2^e, ..., y_n - y_n^e)$$

with continuous first partial derivatives with respect to $y_i - y_i^e$, all i, and such that:

(i) V is positive definite, i.e. $V > 0$ if at least one of the quantities $y_1 - y_1^e$, $y_2 - y_2^e, ..., y_n - y_n^e$ is different from zero; and $V = 0$ if, and only if, $y_i - y_i^e = 0$ for all i;

(ii) $V \to +\infty$ as $\|\mathbf{y} - \mathbf{y}^e\| \to +\infty$;

(iii) $\frac{dV}{dt} = \sum_{i=1}^{n} \frac{\partial V}{\partial (y_i - y_i^e)} \frac{d(y_i - y_i^e)}{dt}$ is negative if at least one of the quantities $y_1 - y_1^e$, $y_2 - y_2^e, ..., y_n - y_n^e$ is different from zero; and $dV/dt = 0$ if, and only if, $y_i - y_i^e = 0$ for all i.

Then the equilibrium state $(y_1^e, y_2^e, ..., y_n^e)$ is globally stable (uniformly asymptotically stable in the large).

A heuristic proof of the theorem is straightforward[3]. The existence of the *'Liapunov function'* V implies that the point whose coordinates are $y_1(t), y_2(t), ..., y_n(t)$, where the $y_i(t)$ are determined by the solution of system (23.1), approaches more and more, as t increases, the point $(y_1^e, y_2^e, ..., y_n^e)$. The latter is therefore globally stable. The interesting thing is that, as we have already noted, properties (i) to (iii) can be checked without solving the system, since the explicit knowledge of the functions $y_i(t)$ is not required.

It must be noted that the converse of Theorem 23.1 is also true, i.e. *if* the equilibrium state is globally stable, *then* there exists a 'Liapunov function'

[3]For rigorous formal proofs, see any of the following: Hahn (1967), Kalman and Bertram (1960), Krasovskiĭ (1963), LaSalle and Lefschetz (1961), Sansone and Conti (1964).

which satisfies all the conditions of Theorem 23.1. Thus the existence of a Liapunov function as required by Theorem 23.1 is a necessary and sufficient condition for global stability.

The second method may serve also to prove *instability*, since the theorem is true that, *if* there exists a function V having the same properties (i) and (ii) of Theorem 23.1 and dV/dt is always *positive* (being zero if, and only if, $y_i - y_i^e = 0$, for all i), *then* the equilibrium state is globally unstable. This theorem, too, is intuitive, since $dV/dt > 0$ implies that the point whose coordinates are $y_1(t), y_2(t), ..., y_n(t)$ moves farther and farther away from the equilibrium point.

In the intuitive reasoning above we have used the expressions 'to approach' and 'to move away from' the equilibrium point, which imply that the 'distance' between the point whose position is given by the solution of system (23.1) and the equilibrium point decreases or increases with time. These considerations give us a hint: *the simplest way to apply Liapunov's second method*—and this is the only practical suggestion that we can give—*is to try, in the search for a Liapunov function, the functions which define the 'distance'*.

Such functions, in fact, by their very definition satisfy conditions (i) and (ii) of Theorem 23.1, and so it only remains to check (but this is the difficult part!) that $dV/dt < 0$ (or $dV/dt > 0$, if we are interested in proving instability). For this check a word of caution is in order: some 'distance' may not be everywhere differentiable (the maximum norm is one of them). In such cases, provided that the distance is a continuous function, requirement (iii) may be replaced, at those points where the function is not differentiable, by the proof that the function is strictly decreasing there.

If the 'distance' fails to give the desired results, then we shall have to look for more general functions satisfying the requirements of the theorem. Actually, most economic results obtained applying Liapunov's second method have been reached using the 'distance' functions.

We have so far treated systems of differential equations: let us now state a theorem for systems of difference equations.

Theorem 23.2. Consider the autonomous system

$$y_{i_{t+1}} = f_i(y_{1_t}, y_{2_t}, ..., y_{n_t}) \ , \qquad i = 1, 2, ..., n \ ,$$

where $f_i = y_i^e$ for $y_{i_t} = y_i^e$, all i.

Suppose there exists a continuous scalar function

$$V(y_1 - y_1^e, y_2 - y_2^e, ..., y_n - y_n^e) \ ,$$

such that:

(i) V is positive definite, i.e. $V > 0$ if at least one of the quantities $y_1 - y_1^e, y_2 - y_2^e, ..., y_n - y_n^e$ is different from zero; and $V = 0$ if, and only if, $y_i - y_i^e = 0$ for all i;

(ii) $V \to +\infty$ as $\|\mathbf{y} - \mathbf{y}^e\| \to +\infty$;

(iii) $\Delta V \equiv V_{t+1} - V_t$ is negative if at least one of the quantities $y_1 - y_1^e, y_2 - y_2^e, ..., y_n - y_n^e$ is different from zero; and $\Delta V = 0$ if, and only if, $y_i - y_i^e = 0$, all i.

Then the equilibrium state $(y_1^e, y_2^e, ..., y_n^e)$ is globally stable (uniformly asymptotically stable in the large).

A heuristic proof of this theorem can be given as for Theorem 23.1, and we need not repeat it here.

Let us note that, when the functions f_i are such that

$$\|\mathbf{f}\| < \|\mathbf{y}\| \qquad (\mathbf{f}(0) = 0) \tag{23.7}$$

for *some* norm, they are called a *contraction*. In this case the equilibrium is stable, since choosing $V = \|\mathbf{y}_t\|$ we have

$$\Delta V = \|\mathbf{y}_{t+1}\| - \|\mathbf{y}_t\| = \|\mathbf{f}\| - \|\mathbf{y}\| < 0. \tag{23.8}$$

There is also a similar global stability condition involving the Jacobian of the functions f_i (which must then be continuously differentiable). Such a condition (Wu and Brown, 1989) is that, for some norm,

$$\|\mathbf{f}'(\mathbf{y})\| < 1 \quad \text{for all } \mathbf{y} \neq \mathbf{0}, \tag{23.9}$$

where $\mathbf{f}'(\mathbf{y}) \equiv J(\mathbf{y})$ is the Jacobian of the functions f_i. This result generalizes a previous one that required condition (23.9) to hold also for $\mathbf{y} = \mathbf{0}$. Further results along this line are given by Brock and Scheinkman (1975), Szidarovszky and Okuguchi (1988), Szidarovszky (1990).

A final remark: Theorems 23.1 and 23.2 are applicable only to first-order systems in 'normal' form, i.e. to systems having the form (23.1) or (23.2). If the system that we have to analyse is not in this form, it must be transformed into a 'normal' first-order system before we can apply the 'second method'. The transformation into a first-order system is in principle always possible by introducing new appropriately defined variables. Suppose that $\mathrm{d}^2y_i/\mathrm{d}t^2$ appears in a differential system; then define a new variable z_t such that $\mathrm{d}y_i/\mathrm{d}t = z_i$ and substitute $\mathrm{d}z_i/\mathrm{d}t$ for $\mathrm{d}^2y_i/\mathrm{d}t^2$. If $y_{i_{t+2}}$ appears in a difference system, define z_{i_t} such that $y_{i_{t+1}} = z_{i_t}$ and substitute $z_{i_{t+1}}$ for $y_{i_{t+2}}$. Derivatives (or lags) higher than the second can be eliminated in a similar way (e.g., if $\mathrm{d}^3y_i/\mathrm{d}t^3$ also appears, define a new variable w_i such that $\mathrm{d}z_i/\mathrm{d}t = w_i$ and substitute $\mathrm{d}w_i/\mathrm{d}t$ for $\mathrm{d}^3y_i/\mathrm{d}t^3$).

After transforming the system into a first-order system, however, we must still put it in the 'normal' form (23.1) or (23.2) as the case may be. Suppose that we have the first-order systems

$$\varphi_i\left(y_1, y_2, ..., y_n; y_1', y_2', ..., y_n'\right) = 0 ,$$

or

$$\varphi_i\left(y_{1_t}, y_{2_t}, ..., y_{n_t}; y_{1_{t+1}}, y_{2_{t+1}}, ..., y_{n_{t+1}}\right) = 0 ,$$

where the functions φ_i have continuous first partial derivatives. A sufficient condition that such systems may be put, at least in principle, in the 'normal' forms (23.1) or (23.2) is that the Jacobian of the φ_i with respect to the y_i' or to the $y_{i_{t+1}}$ be different from zero[4]. Otherwise, the normal form cannot be obtained with certainty.

Liapunov's second method can be applied profitably to the study of some general problems of stability, such as that of *structural stability* (see Chap. 21, Sect. 21.2.3). The interested reader is referred to Krasowskiĭ (1963, Chap. 4).

23.3 Some economic applications

The first extensive application of Liapunov's second method to economics was made by Arrow and Hurwicz (1958) and Arrow, Block and Hurwicz (1959) in their classic papers on the global stability of general competitive equilibrium. However, this method had been previously mentioned in the economic literature: see, for example, Bushaw and Clower (1954). Weintraub (1987, 1991) examines earlier applications of Liapunov's second method to stability of economic equilibria by the Japanese economist T. Yasui and the French economist M. Allais.

23.3.1 Global stability of Walrasian general equilibrium

Let us recall from Part II, Chap. 19, Sect. 19.1, that Walrasian '*tâtonnement*' can be formalized in the following system of differential equations:

$$\frac{\mathrm{d}p_i}{\mathrm{d}t} = k_i E_i\left(p_1, p_2, ..., p_n\right) , \qquad i = 1, 2, ..., n ,$$

where the p's are the prices, the E's the aggregate excess demand functions and the k's positive constants.

Now, choose as the Liapunov function the square of the modified Euclidean distance, i.e.

$$V\left(p_1 - p_1^e, p_2 - p_2^e, ..., p_n - p_n^e\right) = \sum_{i=1}^{n} \frac{1}{k_i}\left(p_i - p_i^e\right)^2 .$$

[4]This is the implicit function theorem in its 'local' form. If we want a 'global' univalence, then the condition is that all the principal minors of the said Jacobian be everywhere positive. See the article by Gale and Nikaidô (1965) and Nikaidô (1968).

We have (remember that the p_i^e's are constants)

$$\frac{dV}{dt} = 2\sum_{i=1}^{n} \frac{1}{k_i}(p_i - p_i^e)\frac{dp_i}{dt} ,$$

so that, since

$$\frac{dp_i}{dt} = k_i E_i(p_1, p_2, ..., p_n) ,$$

we have

$$\begin{aligned}\frac{dV}{dt} &= 2\sum_{i=1}^{n}(p_i - p_i^e) E_i(p_1, p_2, ..., p_n)\\ &= 2\sum_{i=1}^{n} p_i E_i(p_1, p_2, ..., p_n) - 2\sum_{i=1}^{n} p_i^e E_i(p_1, p_2, ..., p_n) .\end{aligned}$$

From static general equilibrium theory we know that

$$\sum_{i=1}^{n} p_i E_i = 0 \qquad \text{(Walras' law)} ,$$

and so

$$\frac{dV}{dt} = -2\sum_{i=1}^{n} p_i^e E_i(p_1, p_2, ..., p_n) .$$

Now, let us state without proof the following[5]:

Lemma. If the equilibrium prices are all positive, gross substitutability prevails, and Walras' law, together with positive homogeneity, holds, then

$$\sum_{i=1}^{n} p_i^e E_i(p_1, p_2, ..., p_n)$$

is always positive for any non-equilibrium positive price vector $(p_1, p_2, ..., p_n)$.

Positive homogeneity (of degree zero) means that, if all prices are multiplied by the same positive constant, the excess demands do not vary, and this is a well-known consequence of the utility maximization postulate. Gross substitutability means, as we saw Chap. 19, Sect. 19.1.2, that $\partial E_j/\partial p_i > 0$ for all i, j; $i \neq j$.

Using the lemma, it follows immediately that dV/dt is always negative out of equilibrium, becoming zero only at the equilibrium point (by definition, $E_i = 0$ for $p_i = p_i^e$, all i).

Gross substitutability, then, implies global stability of general equilibrium.

[5]See Arrow, Block and Hurwicz (1959, p. 90). Note that the positivity of the expression $\sum_{i=1}^{n} p_i^e E_i(p_1, p_2, ..., p_n)$ means that the aggregate excess demand functions satisfy the weak axiom of revealed preference.

Another interesting case of global stability is when

$$\frac{\partial E_j}{\partial p_j} < 0 , \qquad \left|\frac{\partial E_j}{\partial p_j}\right| > \sum_{\substack{s=1\\ s\neq j}}^{n}\left|\frac{\partial E_j}{\partial p_s}\right| \qquad \text{for all } j ,$$

i.e. the case of the *dominant negative diagonal.* For the Liapunov function choose

$$V = \max_j |k_j E_j| , \qquad k_j \text{ positive constants} .$$

Let $|k_w E_w| \geq |k_j E_j|$ for all j, where w is a subscript belonging to some excess demand function. Then

$$V = |k_w E_w| ,$$

and so, wherever dV/dt exists[6], we have[7]

$$\frac{dV}{dt} = k_w(sgn\ E_w)\sum_j \frac{\partial E_w}{\partial p_j}\frac{dp_j}{dt} = k_w(sgn\ E_w)\sum_j \frac{\partial E_w}{\partial p_j} k_j E_j . \tag{23.10}$$

It can immediately be checked that $dV/dt = 0$ in equilibrium, where $E_j = 0$ for all j. If we are out of equilibrium, then $E_j \neq 0$ for at least one j, and so $|E_w| > 0$. Given the assumption made at the beginning, we have, for $j = w$,

$$\left|\frac{\partial E_w}{\partial p_w}\right| > \sum_{s\neq w}\left|\frac{\partial E_w}{\partial p_s}\right| ,$$

and, multiplying both members by $|E_w|$, we have

$$\left|\frac{\partial E_w}{\partial p_w}\right| |E_w| > \sum_{s\neq w}\left|\frac{\partial E_w}{\partial p_s}\right| |E_w| . \tag{23.11}$$

Now, since $|k_w E_w| \geq |k_j E_j|$, we have $|E_w| \geq (k_j/k_w)|E_j|$ and so, using s instead of j,

$$\sum_{s\neq w}\left|\frac{\partial E_w}{\partial p_s}\right| |E_w| \geq \sum_{s\neq w}\left|\frac{\partial E_w}{\partial p_s}\right| \frac{k_s}{k_w}|E_s| = \frac{1}{k_w}\sum_{s\neq w}\left|\frac{\partial E_w}{\partial p_s}\right| k_s |E_s| . \tag{23.12}$$

From (23.11) and (23.12) we have

$$\left|\frac{\partial E_w}{\partial p_w}\right| |E_w| > \frac{1}{k_w}\sum_{s\neq w}\left|\frac{\partial E_w}{\partial p_s}\right| |E_s| k_s ;$$

[6]In the following proof we assume that dV/dt exists everywhere. For the case where it does not exist, see Arrow, Block and Hurwicz (1959, p. 106).

[7]The notation $(sgn\ E_w)$ means 'the sign of E_w', i.e. $(sgn\ E_w) = +1$ if $E_w > 0$, $(sgn\ E_w) = -1$ if $E_w < 0$, $(sgn\ E_w) = 0$ if $E_w = 0$. Note that $|k_w E_w| = k_w |E_w|$ since $k_w > 0$. Now, $k_w |E_w| = k_w (sgn E_w) E_w$, from which, differentiating with respect to t, we have relation (23.10).

therefore

$$\left|\frac{\partial E_w}{\partial p_w}\right| |E_w|\, k_w > \sum_{s \neq w} \left|\frac{\partial E_w}{\partial p_s}\right| |E_s|\, k_s \, .$$

Since $\partial E_w/\partial p_w < 0$ by assumption, the last inequality may be written as

$$-\frac{\partial E_w}{\partial p_w}\, (sgn\ E_w)\, E_w k_w > \sum_{s \neq w} \left|\frac{\partial E_w}{\partial p_s}\right| |E_s|\, k_s \, . \tag{23.13}$$

Now, the quantity

$$\sum_{s \neq w} \left|\frac{\partial E_w}{\partial p_s}\right| |E_s|\, k_s$$

is certainly not smaller than the quantity

$$\sum_{s \neq w} (sgn\ E_w)\, \frac{\partial E_w}{\partial p_s} E_s k_s \, .$$

Indeed, three cases are possible[8]:

(1) For one or more subscripts s, it happens that $\partial E_w/\partial p_s$ and E_s have the same sign, and so

$$\frac{\partial E_w}{\partial p_s} E_s = \left|\frac{\partial E_w}{\partial p_s}\right| |E_s| \, .$$

Now, if $(sgn\ E_w) = +1$, the corresponding elements in the two sums are equal, whereas, if $(sgn\ E_w) = -1$, the elements in the first sum are greater than the corresponding elements in the second sum.

(2) For one or more subscripts s, it happens that $\partial E_w/\partial p_s$ and/or E_s are equal to zero: in this case the corresponding elements in the two sums are equal.

(3) For one or more subscripts s, it happens that $\partial E_w/\partial p_s$ and E_s are of opposite sign. Now, if $(sgn\ E_w) = -1$, then

$$(sgn\ E_w)\, \frac{\partial E_w}{\partial p_s} E_s = \left|\frac{\partial E_w}{\partial p_s}\right| |E_s| \, ,$$

and the corresponding elements in the two sums are equal; if $(sgn\ E_w) = +1$, then $(sgn\ E_w)\, (\partial E_w/\partial p_s)\, E_s$ is a negative quantity and the elements in the first sum are greater than the corresponding elements in the second sum.

This proves that the elements in the first sum are not smaller than the corresponding elements in the second sum, and so[9]

$$\sum_{s \neq w} \left|\frac{\partial E_w}{\partial p_s}\right| |E_s|\, k_s \geq (sgn\ E_w) \sum_{s \neq w} \frac{\partial E_w}{\partial p_s} E_s k_s \, . \tag{23.14}$$

[8] Apart from the trivial case $(sgnE_w) = 0$, which anyway is excluded since we are out of equilibrium. Of course, the possible cases are not mutually exclusive.

[9] Since $(sgn\ E_w)$ does not depend on s, we may write

$$\sum_{s \neq w} (sgn\ E_w)\, \frac{\partial E_w}{\partial p_s} E_s k_s$$

in the form

$$(sgn\ E_w) \sum_{s \neq w} \frac{\partial E_w}{\partial p_s} E_s k_s .$$

Now, from (23.13) and (23.14) it follows that

$$-\frac{\partial E_w}{\partial p_w}\, (sgn\ E_w)\, E_w k_w > (sgn\ E_w) \sum_{s \neq w} \frac{\partial E_w}{\partial p_s} E_s k_s \, ; \tag{23.15}$$

therefore

$$0 > (sgn\ E_w)\, \frac{\partial E_w}{\partial p_w} E_w k_w + (sgn\ E_w) \sum_{s \neq w} \frac{\partial E_w}{\partial p_s} E_s k_s \, ,$$

so that, putting under the $\sum$ sign the term $(sgn\ E_w)\, (\partial E_w/\partial p_w)\, E_w k_w$, we have

$$0 > (sgn\ E_w) \sum_{s} \frac{\partial E_w}{\partial p_s} E_s k_s \, . \tag{23.16}$$

Since the subscript s runs over the same set of indices as the subscript j, we have

$$0 > (sgn\ E_w) \sum_{j} \frac{\partial E_w}{\partial p_j} E_j k_j \, . \tag{23.17}$$

Thus we have proved that in any non-equilibrium point inequality (23.17) holds. From (23.17) and (23.10) it follows that $dV/dt < 0$ out of equilibrium, and this proves the global stability of equilibrium in the case under consideration. The proof can be easily extended to the case of a *quasi-dominant* negative diagonal, in which

$$\frac{\partial E_j}{\partial p_j} < 0 \, , \qquad c_j \left|\frac{\partial E_j}{\partial p_j}\right| > \sum_{s \neq j} c_s \left|\frac{\partial E_j}{\partial p_s}\right| \, , \qquad j = 1, 2, ..., n \, ,$$

where the c's are positive constants. In this case, take

$$V = \frac{k_w}{c_w} |E_w| \, ,$$

and then proceed as in the previous case.

The two cases of global stability examined—gross substitutability and negative diagonal dominance—generalize the results already obtained 'locally' (see Chap. 19, Sect. 19.1.2).

In their first paper on the stability of the competitive equilibrium, Arrow and Hurwicz (1958) said that 'in none of the cases studied have we found the system to be unstable under the (perfectly competitive) adjustment process' (p. 529), and suggested tentatively the proposition that under perfect competition the system is always stable, admitting, however, that 'it is conceivable...that...an example of unstable unique competitive equilibrium may be found' (p. 530). Some such examples were found by Scarf. Here we shall examine the first of Scarf's counter-examples to the stability of general equilibrium.

Consider an economy involving only three consumers and three goods. The utility functions are such that each consumer desires only two commodities in a fixed ratio (i.e. the two commodities are perfectly complementary), which is taken to be one to one, and has no desire for the remaining commodity. It is assumed that the first consumer desires only goods 1 and 2, the second consumer desires only goods 2 and 3 and the third consumer desires only goods 3 and 1. Formally, the utility functions can be written as

$$\begin{aligned} U_1(x_{11}, x_{12}, x_{13}) &= \min(x_{11}, x_{12}), \\ U_2(x_{21}, x_{22}, x_{23}) &= \min(x_{22}, x_{23}), \\ U_3(x_{31}, x_{32}, x_{33}) &= \min(x_{33}, x_{31}), \end{aligned}$$

where x_{ij} is the quantity of good j consumed by the $i-th$ individual. For example, the typical indifference curve of the first consumer is shown in Fig. (23.1). Finally, it is assumed that the initial endowments are

$$\bar{x}_{ij} = \begin{cases} 1 & \text{for} \quad i = j, \\ 0 & \text{for} \quad i \neq j, \end{cases}$$

that is, the first consumer possesses initially one unit of the first good and zero units of the goods 2 and 3, and so on.

Let us consider the first consumer. For any income M_1 he will demand the same quantity of goods 1 and 2 and, therefore, putting $x_{11} = x_{12}$ and $x_{13} = 0$ in his budget constraint $p_1x_{11} + p_2x_{12} + p_3x_{13} = M_1$, we obtain his demand functions

$$\begin{aligned} x_{11}(p_1, p_2, p_3, M_1) &= M_1/(p_1 + p_2), \\ x_{12}(p_1, p_2, p_3, M_1) &= M_1/(p_1 + p_2), \\ x_{13}(p_1, p_2, p_3, M_1) &= 0. \end{aligned}$$

Now the 'income' of the first consumer is derived from his initial holding of the first good, so that $M_1 = p_1$. Thus the excess demand functions of the first consumer are

$$\begin{aligned} E_{11}(p_1, p_2, p_3) &= x_{11}(p_1, p_2, p_3, M_1) - 1 &&= -p_2/(p_1 + p_2), \\ E_{12}(p_1, p_2, p_3) &= x_{12}(p_1, p_2, p_3, M_1) - 0 &&= p_1/(p_1 + p_2), \\ E_{13}(p_1, p_2, p_3) &= 0\ . \end{aligned}$$

In a similar way we can derive the excess demand functions of the second and third consumers, and adding the three excess demand functions for each good we obtain the following aggregate excess demand functions:

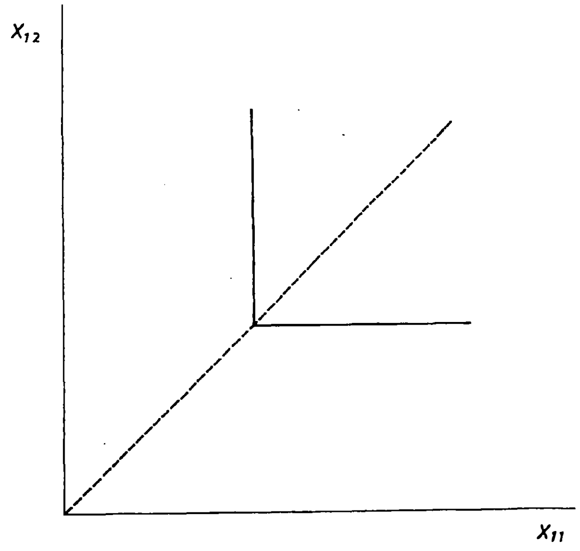


Figure 23.1: Perfect complementarity

$$\begin{aligned} E_1(p_1, p_2, p_3) &= \frac{-p_2}{p_1 + p_2} + \frac{p_3}{p_3 + p_1}, \\ E_2(p_1, p_2, p_3) &= \frac{-p_3}{p_2 + p_3} + \frac{p_1}{p_1 + p_2}, \\ E_3(p_1, p_2, p_3) &= \frac{-p_1}{p_3 + p_1} + \frac{p_2}{p_2 + p_3}. \end{aligned} \tag{23.18}$$

It can easily be verified that the only equilibrium situation is $p_1 = p_2 = p_3$. To determine 'absolute' prices we need a normalization condition, e.g. $p_1^2 + p_2^2 + p_3^2 = 3$ (alternatively we could choose one good as 'numéraire' and put its price equal to 1); the equilibrium point is then $(1, 1, 1)$.

Consider now the dynamic adjustment process

$$\frac{dp_i}{dt} = E_i(p_1, p_2, p_3)\ . \tag{23.19}$$

We shall show that

$$p_1p_2p_3 = \text{ constant} \tag{23.20}$$

for any solution of (23.19). Differentiation of (23.20) with respect to time yields

$$\frac{dp_1}{dt}p_2p_3 + \frac{dp_2}{dt}p_1p_3 + \frac{dp_3}{dt}p_1p_2 = 0, \tag{23.21}$$

and using (23.19) and (23.18) we have

$$\frac{p_3\left(p_1^2-p_2^2\right)}{p_1+p_2}+\frac{p_2\left(p_3^2-p_1^2\right)}{p_3+p_1}+\frac{p_1\left(p_2^2-p_3^2\right)}{p_2+p_3} \tag{23.22}$$
$$=p_3\left(p_1-p_2\right)+p_2\left(p_3-p_1\right)+p_1\left(p_2-p_3\right)=0\,,$$

which proves (23.20). It follows that equilibrium is not (asymptotically) stable. In fact the value of $p_1p_2p_3$ at equilibrium is 1, and, if $p_1(0)\,p_2(0)\,p_3(0) \neq 1$, equilibrium will never be reached, since by (23.20) we have $p_1(t)\,p_2(t)\,p_3(t) = p_1(0)\,p_2(0)\,p_3(0)$.

This completes our introduction to the study of the global stability of general equilibrium. The student who wishes to pursue the matter further may consult, as a first step, Negishi's (1962) survey article; see also Negishi (1972), Arrow and Hahn (1971), Dohtani (1993).

23.3.2 Rules of thumb in business management

The complete information required by a firm to exactly compute the optimum price as indicated by economic theory, is not freely available in real life; it is conceivable that it might be obtained, but only at a cost which would exceed the benefits. Business management, then, uses rules of thumb in its decision making.

Baumol and Quandt (1964) investigated how 'good' rules of thumb can be constructed and tested. A 'learning' rule of thumb is the following: a price change is made (this of course presupposes that the firm has some sort of market power), and the resulting change in profits is observed. If the profit change is positive, price is changed again in the same direction; if, instead, it is negative, price is changed in the opposite direction to that of the previous change; if, finally, the level of profits is stationary, price is not changed again. This rule of thumb can be proved to converge globally to the optimum point, provided that the latter exists and that no shifts occur in the profit function.

The rule under examination can be formalized as follows (a dot denotes $\mathrm{d}/\mathrm{d}t$):

$$\begin{aligned}\dot{p} &= g\left(\frac{\dot{\pi}}{\dot{p}}\right) && \text{if } \dot{p}\neq 0\,,\\ \dot{p} &= 0 && \text{otherwise}\,,\end{aligned}$$

where p is price, π is profit and g is a sign-preserving function, i.e. $g \gtreqless 0$ if $\dot{\pi}/\dot{p} \gtreqless 0$[10]. Let us now assume that the (unknown) profit function is a concave, differentiable function of p and that an optimum exists, i.e.

[10] This is a generalization of Baumol's and Quandt's analysis (they use the relation $\dot{p}=k\dot{\pi}/\dot{p}$, k a positive constant).

$$\begin{aligned}\pi &= f(p)\,, && f''<0 \text{ everywhere}\,,\\ f'(p) &= 0 && \text{for } p=p_e\,.\end{aligned}$$

Now, $\dot{\pi}=f'(p)\,\dot{p}$, and so the basic dynamical equation can be written as

$$\dot{p}=g\left(f'(p)\right)\,.$$

As a Liapunov function, Baumol and Quandt choose $V=\frac{1}{2}(\dot{p}-0)^2$, but this is not correct, since such a V does not necessarily satisfy requirement (ii) of Theorem 23.1. This illustrates the importance of checking all the requirements of the theorem when applying it. A correct Liapunov function is, for example, the square of the Euclidean distance:

$$V=\left(p-p_e\right)^2.$$

We have

$$\frac{\mathrm{d}V}{\mathrm{d}t}=2\left(p-p_e\right)\dot{p}=2\left(p-p_e\right)g\left(f'(p)\right)\,.$$

From the assumptions on the profit function it follows that

$$f'(p)\gtreqless 0 \quad \text{as} \quad p \lesseqgtr p_e,$$

and since g is a sign-preserving function, this implies that $\mathrm{d}V/\mathrm{d}t$ is always negative out of equilibrium, which proves global stability.

23.3.3 Price adjustment and oligopoly under product differentiation

We have already treated Cournot's oligopoly model in Chap. 10, Sect. 10.1. Here we examine a contribution by Okuguchi (1969) as an illustration of the application of Liapunov's second method to difference systems[11].

In this model price is the variable being adjusted by each oligopolist, on the assumption that its rivals' prices remain unchanged at the level of the previous period. There are n oligopolistic firms with differentiated products. The expected demand for the $i-$th firm at period t is

$$q_i(t)=a_i-\sum_{\substack{j=1\\ j\neq i}}^{n} b_{ij}p_{j_{t-1}}-b_{ii}\tilde{p}_{i_t}\,, \qquad i=1,2,\ldots,n\,, \tag{23.23}$$

where p_j is the $j-$th firm's actual price and $\tilde{p}_i$ is the target price which maximizes the $i-$th firm's profit; the assumptions on the various coefficients

[11] Although Okuguchi does not mention Liapunov's second method, it is the method he actually followed. We also warn that our proof will be slightly different from his.

are $a_i > 0$, $b_{ij} < 0$ for $i \neq j$ and $b_{ii} > 0$. These assumptions mean that the demand curve facing each firm is 'normal' (in the sense that the 'own' price effect is negative, a change in $\bar{p}_i$ causing a change in q_i in the opposite direction) and that all the n different products are substitutes (a change in $p_j, j \neq i$, causes a change in q_i in the same direction). The fact that the p_j's are referred to $t-1$ reflects the assumption made by each firm that its rivals' prices will remain unchanged at the level of the previous period while it changes its price to maximize its profit. The profit that the i-th firm expects to make at period t given the above demand function is

$$\pi_{i_t} = \bar{p}_{i_t} q_{i_t} - c_i(q_{i_t}) = \bar{p}_{i_t}\left(a_i - \sum_{j\neq i}^{n} b_{ij} p_{j_{t-1}} - b_{ii}\bar{p}_{i_t}\right) - c_i(q_{i_t}) \ , \tag{23.24}$$

where $c_i(q_i)$ is the total cost function of the i-th firm, with $\mathrm{d}c_i/\mathrm{d}q_i > 0$.

Each firm maximizes its expected profit with respect to $\bar{p}_i$. The first-order condition for a maximum of (23.24) is $\mathrm{d}\pi_i/\mathrm{d}\bar{p}_i = 0$, whence

$$a_i - \sum_{j\neq i}^{n} b_{ij} p_{j_{t-1}} - b_{ii}\bar{p}_{i_t} - b_{ii}\bar{p}_{i_t} - \frac{\mathrm{d}c_i}{\mathrm{d}q_i}\frac{\mathrm{d}q_i}{\mathrm{d}\bar{p}_i} = 0 \ , \qquad i = 1, 2, ..., n \ , \tag{23.25}$$

that is, being $\mathrm{d}q_i/\mathrm{d}\bar{p}_i = -b_{ii}$,

$$a_i - \sum_{j\neq i}^{n} b_{ij} p_{j_{t-1}} - 2b_{ii}\bar{p}_{i_t} + \frac{\mathrm{d}c_i}{\mathrm{d}q_i} b_{ii} = 0 \ , \qquad i = 1, 2, ..., n \ . \tag{23.26}$$

The second-order condition gives

$$-2b_{ii} - b_{ii}^2 \frac{\mathrm{d}^2 c_i}{\mathrm{d}q_i^2} < 0 \ ,$$

that is

$$2b_{ii} + b_{ii}^2 \frac{\mathrm{d}^2 c_i}{\mathrm{d}q_i^2} > 0 \ . \tag{23.27}$$

Let us note that (23.27) is certainly satisfied if $\mathrm{d}^2 c_i/\mathrm{d}q_i^2 \geq 0$ (increasing or constant marginal cost); it may also be satisfied if $\mathrm{d}^2 c_i/\mathrm{d}q_i^2 < 0$, provided that $\mathrm{d}^2 c_i/\mathrm{d}q_i^2 > -2/b_{ii}$.

Equation (23.26) can be used to express $\bar{p}_{i_t}$ as a single-valued and differentiable function of the prices of all the other firms lagged one period[12]

$$\bar{p}_{i_t} = f^i\left(p_{1_{t-1}}, p_{2_{t-1}}, ..., p_{i-1_{t-1}}, p_{t+1_{t-1}}, ..., p_{n_{t-1}}\right) \ , \qquad i = 1, 2, ..., n \ . \tag{23.28}$$

To compute the partial derivatives of (23.28)—the knowledge of their sign is needed in the course of the stability analysis—we differentiate the first-order condition (23.26) totally with respect to $p_{j_{t-1}}$, keeping in mind that $\bar{p}_{i_t}$ is a function of $p_{j_{t-1}}$ as stated in (23.28). We obtain

[12]This in an application of the implicit function theorem. See Chap. 20, Sect. 20.2.

$$-b_{ij} - 2b_{ii}\frac{\partial \bar{p}_{i_t}}{\partial p_{j_{t-1}}} + b_{ii}\frac{\mathrm{d}^2 c_i}{\mathrm{d}q_i^2}\frac{\mathrm{d}q_i}{\mathrm{d}p_{j_{t-1}}} = 0 \ , \qquad i, j = 1, 2, ..., n \ ; \ i \neq j \ ,$$

whence, being $\mathrm{d}q_i/\mathrm{d}p_{j_{t-1}} = -b_{ij} - b_{ii}\partial\bar{p}_{i_t}/\partial p_{j_{t-1}}$,

$$\frac{\partial \bar{p}_{i_t}}{\partial p_{j_{t-1}}} = \frac{\partial f^i}{\partial p_{j_{t-1}}} = -\frac{b_{ij} + b_{ii}b_{ij}\mathrm{d}^2 c_i/\mathrm{d}q_i^2}{2b_{ii} + b_{ii}^2\mathrm{d}^2 c_i/\mathrm{d}q_i^2} > 0 \ , \qquad i, j = 1, 2, ..., n \ ; \ i \neq j \ . \tag{23.29}$$

The positive sign is a consequence of the second-order condition (23.27) and of the assumptions on the signs of the b's.

Let us now introduce the following price adjustment mechanism:

$$p_{i_t} - p_{i_{t-1}} = k_i\left(\bar{p}_{i_t} - p_{i_{t-1}}\right) \ , \qquad 0 < k_i \leq 1 \ , \qquad i = 1, 2, ..., n \ , \tag{23.30}$$

which means that when the actual price is different from the target price the firm adjusts the former towards the latter. It is plausible to think that the gap is not made up completely in one period, so that probably the adjustment coefficient k_i is smaller than one; however, only the case $k_i > 1$ (overadjustment) has been excluded. Substituting (23.28) into (23.30) and shifting the time subscripts forwards by one period, we obtain

$$p_{i_{t+1}} = k_i f^i(p_{1_t}, p_{2_t}, ..., p_{i-1_t}, p_{i+1_t}, ..., p_{n_t}) + (1 - k_i)p_{i_t} \ . \tag{23.31}$$

Applying the mean value theorem to Eq. (23.31) we obtain

$$\begin{aligned} &[k_i f^i(p_{1_t}, p_{2_t}, ..., p_{i-1_t}, p_{i+1_t}, ..., p_{n_t}) + (1 - k_i)p_{i_t}] \\ &- [k_i f^i(p_{1_e}, p_{2_e}, ..., p_{i-1_e}, p_{i+1_e}, ..., p_{n_e}) + (1 - k_i)p_{i_e}] \\ &= \textstyle\sum_{j\neq i} k_i \frac{\partial f^i}{\partial p_j}(p_{j_t} - p_{j_e}) + (1 - k_i)(p_{i_t} - p_{i_e}) \ , \end{aligned} \tag{23.32}$$

where the partial derivatives $\partial f^i/\partial p_j$ are evaluated at an intermediate point between p_{j_t} and p_{j_e} $(j = 1, 2, ..., n; j \neq i)$, say at $p_j^* = p_{j_e} + \theta(p_{j_t} - p_{j_e})$, $0 < \theta < 1, j = 1, 2, ..., n; j \neq i$. The symbol p_{i_e} denotes the i-th firm's equilibrium price.

As $f^i(p_{1_e}, p_{2_e}, ..., p_{i-1_e}, p_{i+1_e}, ..., p_{n_e}) = p_{i_e}$, we have, substituting in (23.32) and then in (23.30),

$$p_{i_{t+1}} - p_{i_e} = \sum_{j\neq i} k_i \frac{\partial f^i}{\partial p_j}(p_{j_t} - p_{j_e}) + (1 - k_i)(p_{i_t} - p_{i_e}) \ , \qquad i = 1, 2, ..., n \ . \tag{23.33}$$

Let us define

$$\alpha_{ij} = \begin{cases} k_i \dfrac{\partial f^i}{\partial p_j} , & i \neq j , \\ 1 - k_j , & i = j . \end{cases} \tag{23.34}$$

It is now easy to show that a *sufficient* stability condition is

$$\sum_{i=1}^{n} \alpha_{ij} < 1 , \qquad j = 1, 2, ..., n . \tag{23.35}$$

In fact, when condition (23.35) holds, system (23.34) defines a *contraction* and so the equilibrium point is globally stable. As a Liapunov function we choose the absolute value distance, i.e.

$$V = \sum_{i=1}^{n} \left| p_{i_t} - p_{i_e} \right| , \tag{23.36}$$

whence

$$\Delta V = \sum_{i=1}^{n} \left| p_{i_{t+1}} - p_{i_e} \right| - \sum_{i=1}^{n} \left| p_{i_t} - p_{i_e} \right| ,$$

which becomes zero at the equilibrium point. To show that $\Delta V < 0$ out of equilibrium, consider

$$\begin{aligned} \left| p_{i_{t+1}} - p_{i_e} \right| &= \left| \sum_{j \neq i} k_i \frac{\partial f^i}{\partial p_j} (p_{j_t} - p_{j_e}) + (1 - k_i)(p_{i_t} - p_{i_e}) \right| & (23.37) \\ &\leq \sum_{j \neq i} \left| k_i \frac{\partial f^i}{\partial p_j} \right| \left| p_{j_t} - p_{j_e} \right| + \left| 1 - k_i \right| \left| p_{i_t} - p_{i_e} \right| \\ &= \sum_{i=1}^{n} \alpha_{ij} \left| p_{i_t} - p_{i_e} \right| , & (23.38) \end{aligned}$$

where the last equality derives from the definition of α_{ij} in (23.34) and from the fact, due to (23.29) and to the definition of k_i in (23.30), that

$$\left| k_i \frac{\partial f^i}{\partial p_j} \right| = k_i \frac{\partial f^i}{\partial p_j} , \qquad \left| 1 - k_i \right| = 1 - k_i . \tag{23.39}$$

Therefore, defining

$$\beta = \max_j \sum_{i=1}^{n} \alpha_{ij} , \tag{23.40}$$

we have ,

$$\begin{aligned} &\textstyle\sum_i \left| p_{i_{t+1}} - p_{i_e} \right| \leq \sum_i \sum_j \alpha_{ij} \left| p_{j_t} - p_{j_e} \right| \\ &= \textstyle\sum_j \sum_i \left| p_{j_t} - p_{j_e} \right| \alpha_{ij} \qquad\qquad (23.41) \\ &= \textstyle\sum_j \left| p_{j_t} - p_{j_e} \right| \sum_i \alpha_{ij} \leq \beta \sum_j \left| p_{j_t} - p_{j_e} \right| , \end{aligned}$$

whence it follows, being $\beta < 1$ because of (23.40) and (23.35),

$$\sum_j \left| p_{i_{t+1}} - p_{i_e} \right| < \sum_j \left| p_{j_t} - p_{j_e} \right| , \qquad i, j = 1, 2, ..., n , \tag{23.42}$$

and so $\Delta V < 0$ out of equilibrium, Q.E.D.

Let us now examine the economic meaning of condition (23.35). Given (23.34), we can rewrite (23.35) as

$$\sum_{i \neq j} k_i \frac{\partial f^i}{\partial p_j} + (1 - k_j) < 1 , \qquad j = 1, 2, ..., n \tag{23.43}$$

or

$$\sum_{i \neq j} k_i \frac{\partial f^i}{\partial p_j} < k_j , \qquad j = 1, 2, ..., n , \tag{23.44}$$

that is, the adjustment coefficients of all other firms multiplied by the derivatives measuring the responsiveness of each of these other firms to a change in the price of the firm under consideration (the j-th firm), must sum up to a value smaller than the adjustment coefficient of the j-th firm, and this must hold for all j. In the particular case in which all adjustment coefficients are assumed equal to one, the condition becomes

$$\sum_{i \neq j} \frac{\partial f^i}{\partial p_j} < 1 , \qquad j = 1, 2, ..., n , \tag{23.45}$$

which will be satisfied if (but not only if)

$$\frac{\partial f^i}{\partial p_j} < \frac{1}{n-1} , \qquad i, j = 1, 2, ..., n ; \qquad i \neq j , \tag{23.46}$$

that is, a sufficient stability condition is that the responsiveness of any firm to a change in any other firm's price be less than a critical value given by the reciprocal of the number of oligopolists less one.

23.4 Exercises

1. Consider the growth model with human capital (Chap. 19, Sect. 19.2), and show that the equilibrium point (1, 1) is globally stable (Hint: as a Liapunov function use $V = \frac{1}{2c}(z_1 - 1)^2 + \frac{1}{2c}(z_2 - 1)^2$.)

2. In the von Stackelberg duopoly model (von Stackelberg, 1933; Varian, 1992, Chap. 16, Sect. 16.6) one firm (the follower) behaves as a Cournot duopolist (see Chap. 10) while the other firm (the leader) knows this fact and chooses its optimal quantity taking into account that the follower will react along its Cournot reaction curve. Let $p = f(Y)$ be the market demand curve,

where $Y = y_1 + y_2$, and $\pi_i = py_i - C_i(y_i)$ the profit of firm i, $i = 1, 2$,where y_i is the quantity produced. Firm 2 is the follower, and will maximise profits considering firm 1's output as given, so that

$$\max_{y_{2t}} \pi_{2t} = f(y_{1t-1} + y_{2t})y_{2t} - C_2(y_{2t}),$$

from which

$$\frac{\partial \pi_{2t}}{\partial y_{2t}} = \frac{\mathrm{d}f}{\mathrm{d}Y} y_{2t} + f(y_{1t-1} + y_{2t}) - \frac{\mathrm{d}C_2}{\mathrm{d}y_{2t}} = 0.$$

Assuming that the second-order condition for an interior maximum are satisfied, from the first-order condition we can express y_{2t} as a function of y_{1t-1}, say $y_{2t} = \phi_2(y_{1t-1})$, which is the reaction curve of the follower. The leader takes this function into account, and maximises

$$\max_{y_{1t}} \pi_{1t} = f(y_{1t} + \phi_2(y_{1t-1}))y_{1t} - C_1(y_{1t}),$$

from which

$$\frac{\partial \pi_{1t}}{\partial y_{1t}} = \frac{\mathrm{d}f}{\mathrm{d}Y} y_{1t} + f(y_{1t} + \phi_2(y_{1t-1})) - \frac{\mathrm{d}C_1}{\mathrm{d}y_{1t}} = 0.$$

This last equation, on the assumption that the second-order condition for an interior maximum is satisfied, gives y_{1t} as a function of y_{1t-1}, say $y_{1t} = h(y_{1t-1})$. Thus we have the difference equation system

$$\begin{aligned} y_{1t} &= h(y_{1t-1}), \\ y_{2t} &= \phi_2(y_{1t-1}). \end{aligned}$$

Examine the global stability of this system (Hint: use the contraction mapping. See also Okuguchi, 1976, Chap. 5).

23.5 References

Andronov, A.A., A.A. Vitt and S.E. Khaikin, 1966, *Theory of Oscillators*, pp. 374 ff.

Arrow, K.J., H.D. Block and L. Hurwicz, 1959, On the Stability of the Competitive Equilibrium, II.

Arrow, K.J. and F.H. Hahn, 1971, *General Competitive Analysis*, Chaps. 11, 12, 13.

Arrow, K.J. and L. Hurwicz, 1958, On the Stability of the Competitive Equilibrium, I.

Baumol, W.J. and R.E. Quandt, 1964, Rules of Thumb and Optimally Imperfect Decisions, Appendix A.

Brock, W.A. and J.A. Scheinkman, 1975, Some Results on Global Asymptotic Stability of Difference equations.

Clower, R.W. and D.W. Bushaw, 1954, Price Determination in a Stock-Flow Economy.

Dohtani, A., 1993, Global Stability of the Competitive Economy Involving Complementary Relations among Commodities.

Gale, D. and H. Nikaidô, 1965, The Jacobian Matrix and Global Univalence of Mappings.

Hahn, W., 1967, *Stability of Motion.*

Kalman, R.E. and J.E. Bertram, 1960, Control System Analysis and Design via the 'Second Method' of Lyapunov, I and II.

Karlin, S., 1959, *Mathematical Methods and Theory in Games, Programming and Economics*, Vol. I, Chap. 9, §§ 9.4, 9.5.

Krasovskiĭ, N.N., 1963, *Stability of Motion: Application of Lyapunov's Second Method to Differential Systems and Equations with Delay.*

LaSalle, J.P. and S. Lefschetz, 1961, *Stability by Liapunov's Direct Method with Applications.*

Liapounoff, A., 1907, *Problème général de la stabilité du mouvement*, pp. 255-67.

Morishima, M., 1964, *Equilibrium Stability, and Growth,* Chap. II.

Negishi, T., 1962, The Stability of a Competitive Economy: A Survey Article.

Negishi, T., 1972, *General Equilibrium Theory and International Trade,* Chap. 13, Sect. 6; Chap. 14.

Newman, P.K., 1961, Approaches to Stability Analysis.

Nikaidô, H., 1968, *Convex Structures and Economy Theory*, Chap. VII.

Okuguchi, K., 1969, On the Stability of Price Adjusting Oligopoly Equilibrium under Product Differentiation.

Okuguchi, K., 1976, *Expectations and Stability in Oligopoly Models.*

Sansone, G. and R. Conti, 1964, *Non-Linear Differential Equations*, Chap. I, Sect. 3.3; Chap. IX, Sects. 1, 3.1.

Scarf, H., 1960, Some Examples of Global Instability of the Competitive Equilibrium.

Szidarovszky, F., 1990, Global Asymptotic Stability of Non-Linear Difference Equations with Non-Differentiable Functions.

Szidarovszky, F. and K. Okuguchi, 1988, A Note on Global Asymptotic Stability of Non-Linear Difference Equations.

Takayama, A., 1985, *Mathematical Economics*, Chap. 3, Sects. D-H.

Varian, H.L., 1992, *Microeconomic Analysis,* Chap. 16.

von Stackelberg, H., 1933, Sulla teoria del duopolio e del polipolio.

Weintraub, E.R., 1987, Stability Theory via Liapunov's Method: A Note on the Contribution of Takuma Yasui.

Weintraub, E.R., 1991, Allais, Stability, and Liapunov Theory.
Wu, J.-W. and D.P. Brown, 1989, Global Asymptotic Stability in Discrete Systems.

Chapter 24

Introduction to Nonlinear Dynamics

24.1 Preliminary remarks

Economic phenomena are not necessarily linear, so that linear and constant-coefficient differential and difference equations (l.c. equations for brevity) cannot be considered as a *generally* adequate tool to analyse dynamic problems. One reason for the widespread use of l.c. equations is of course that such equations are always solvable, whereas as soon as we venture into the field of non-l.c. equations we are not sure of finding the explicit solution. But there is more to it than that. While linearity is one, the possible non-linearities are infinite. When one abandons linearity (and related functional forms that can be reduced to linearity by a simple transformation of variables, such as log-linear equations), in general it is not clear which non-linear form one should adopt. Further to clarify the matter, let us distinguish between *purely qualitative* non-linearity and *specific* non-linearity.

By *purely qualitative* non-linearity we mean the situation in which we only know that a generic non-linear functional relation exists with certain *qualitative* properties, such as continuous first-order partial derivatives with a given sign and perhaps certain bounds (e.g., aggregate consumption depends on national income, with a positive but smaller-than-one partial derivative). This is the typical situation in pure economic theory.

By *specific* non-linearity we mean the situation in which we assume a specific non-linear functional relationship. Since in general it is not clear from the theoretical point of view which non-linear form one should adopt, the choice of a form is often made for convenience (e.g., to satisfy the requirements of a certain theorem which serves to obtain a certain type of motion of the model under investigation).

Since, from the theoretical point of view, specific non-linearity looks as arbitrary as linearity, we believe that in economic theory the true generalization of linear dynamics is purely qualitative non-linearity. Specific non-linearity simply shows that a certain motion is possible, not that the model is more general or more realistic than the corresponding linear model. It goes without saying that specific non-linearity is well founded when there are compelling theoretical or empirical reasons showing that the functional form is of a certain type.

Finally, adherence to the one or the other economic school of thought may also have played a role in the dominance of linear dynamical systems, as suggested by Lorenz (1993, Chap. 1, Sects. 1.1 and 1.3).

Turning to the mathematics, when an equation is non-l.c., more often than not it turns out that, although we know the solution exists (by the existence and uniqueness theorem[1]), we are not able to 'find' it, since it cannot be expressed in terms of (a finite number of) known functions. There are, however, several types of non-l.c. equations which are explicitly integrable. The types of non l.c. equations whose closed-form solution is known have been codified long ago: for differential equations see the first four volumes of the monumental six-volume treatise by Forsyth (1890-1906), and the book by Ince (1926); for difference equations see the books by Boole (1960) and Milne-Thomson (1933) .

It is now as well to recall that the analytical study of differential equations follows two main approaches: the *qualitative* (or *topological)* and the *quantitative*. The *qualitative approach* consists in the analysis of the properties of the solutions of a differential equation (system) without actually knowing the solution itself or trying to approximate it; this is based on phase diagrams, Liapunov's second method, etcetera. On the contrary, the *quantitative approach* consists in trying to find the explicit analytical solution of the differential equation or to approximate it by using power series and other methods. It is then easy to check the nature of the motion of the system. This distinction applies equally well if one considers difference equations instead of differential equations.

It is important to observe that the quantitative approach remains an analytical method, and should not be confused with the *numerical* method(s) of integration, a topic that lies outside the scope of the present book. Numerical integration is also used in the qualitative approach, to obtain an idea (often in the form of phase paths) of the behaviour of the solution.

[1]Consider the differential equation $\mathrm{d}y/\mathrm{d}t = f(y,t)$ and let $f(y,t)$ be a single-valued and continuous function of y and of t in a rectangular domain D surrounding a point (y_0, t_0) and defined by the inequalities $|y - y_0| \leq b$, $|t - t_0| \leq a$. Let M be the upper bound of $|f(y,t)|$ in D, and let h be the smaller of a and of b/M. If $h < a$, the more stringent restriction $|t - t_0| \leq h$ is imposed upon t. Moreover, $f(y,t)$ is Lipschitzian, that is, if (y,t) and (Y,t) are two points in D with the same abscissa, then $|f(Y,t) - f(y,t)| < K|Y - y|$, where K is a constant. All these conditions being satisfied, there exists a unique continuous function of t, say $y(t)$, defined for all values of t such that $|t - t_0| < h$, which satisfies the differential equation and reduces to y_0 when $t = t_0$. This theorem can be extended to system of first-order equations. For proofs see, for example, Ince (1956), Chap. III. For existence and uniquess theorems concerning difference equations see, for example, Milne-Thomson (1933).

In theoretical mathematics since a long time there has been a shift of emphasis from quantitative to qualitative methods. This reflects the above mentioned fact that in most cases, although we know that the solution exists, we are not able to 'find' it. Thus, we shall just treat a few types of integrable non-l.c. equations that have had some, although limited, application in economic dynamics. We shall only examine differential equations, both because in the field of non-l.c. equations only differential equations are normally met in economic applications, and because a difference equation can, if necessary, be easily analysed numerically, obtaining the *precise* time path corresponding to the given numerical values of the parameters and of the initial conditions[2].

We shall then go on to the elements of the *qualitative* theory of non-linear equations, such as non-linear oscillations. We have already treated qualitative methods while studying stability: phase diagrams (Chap. 21, Sects. 21.3 and 21.5) and Liapunov's second method (Chap. 23). These are of course assumed known in the treatment of the present chapter.

Non-linear qualitative dynamics is the field where the most exciting new developments, such as chaotic dynamics, can be found. These will be treated in the two next chapters (where we shall also examine difference equations). Hence our introduction to non-linear dynamics consists of the present chapter and the two following.

24.2 Some integrable differential equations

24.2.1 First-order and first-degree exact equations

The general equation of the first order and of the first degree can be written as

$$P(t,y) + Q(t,y)\frac{\mathrm{d}y}{\mathrm{d}t} = 0, \tag{24.1}$$

and is called of the first degree because the power to which the derivative $\mathrm{d}y/\mathrm{d}t$ is raised is 1. Eq. (24.1) can also be written in the form of a total differential equation:

$$P(t,y)\,\mathrm{d}t + Q(t,y)\,\mathrm{d}y = 0. \tag{24.2}$$

[2]Of course, differential equations can also be analysed numerically, but the method is not easy and in any case yields only approximate results. On the other hand, the precise numerical solution of a difference equation can always and easily be obtained by applying the recurrent method. If, for example, the equation is of the type $y_{t+1} = f(y_t, t)$, then, given y_0, we compute $y_1 = f(y_0, 0)$; from y_1 we compute $y_2 = f(y_1, 1)$, and so on.

When the left-hand side of (24.2) is *immediately* (that is, without previous multiplication by any factor) recognizable as the total differential dz of a function $z(t,y)$, equation (24.2) is said to be *exact* and its solution is

$$z(t,y) = A, \tag{24.3}$$

where A is an arbitrary constant. For this to be true it must, of course, be the case that

$$\begin{aligned} P(t,y) &= \frac{\partial z(t,y)}{\partial t}, \\ Q(t,y) &= \frac{\partial z(t,y)}{\partial y}. \end{aligned} \tag{24.4}$$

Now, for (24.4) to hold, P and Q being differentiable functions, it must be true that

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial t}, \tag{24.5}$$

since by the *commutative theorem* on partial differentiation the following relation:

$$\frac{\partial^2 z}{\partial t \partial y} = \frac{\partial^2 z}{\partial y \partial t}$$

must hold, so that (24.5) is a consequence of (24.4).

Condition (24.5)—called the *integrability condition*—is not only necessary but also sufficient for Eq. (24.2) to be an exact equation. When the integrability condition is satisfied, to find the solution we proceed as follows. Starting from

$$\frac{\partial z(t,y)}{\partial t} = P(t,y),$$

integrating and remembering that—since the differentiation with respect to t was partial—the inverse process of integration will introduce, in addition to an arbitrary constant, also an arbitrary function of y, we have

$$z(t,y) = \int_{t_0}^{t} P(t,y)\,\mathrm{d}t + \phi(y),$$

where t_0 is an arbitary constant. If we differentiate this expression with respect to y we have

$$\frac{\partial z(t,y)}{\partial y} = \int_{t_0}^{t} \frac{\partial P(t,y)}{\partial y}\mathrm{d}t + \frac{\mathrm{d}\phi}{\mathrm{d}y}.$$

We now observe that $\partial z(t,y)/\partial y = Q(t,y)$ since the equation is exact, and that $\partial P(t,y)/\partial y = \partial Q(t,y)/\partial t$ by the integrability condition. Hence

$$\begin{aligned} Q(t,y) &= \int_{t_0}^{t} \frac{\partial Q(t,y)}{\partial t}\mathrm{d}t + \frac{\mathrm{d}\phi}{\mathrm{d}y} \\ &= [Q(t,y) - Q(t_0,y)] + \frac{\mathrm{d}\phi}{\mathrm{d}y}, \end{aligned}$$

which gives

$$\frac{\mathrm{d}\phi}{\mathrm{d}y} = Q(t_0,y).$$

Integrating with respect to y we obtain the hitherto undetermined function $\phi(y)$, namely

$$\phi(y) = \int_{y_0}^{y} Q(t_0,y)\mathrm{d}y,$$

where y_0 is an arbitrary constant.

Thus we can write the solution of (24.3) as

$$\int_{t_0}^{t} P(t,y)\,\mathrm{d}t + \int_{y_0}^{y} Q(t_0,y)\,\mathrm{d}y = A, \tag{24.6}$$

where A is an arbitrary constant and t_0, y_0 may be chosen as convenient. Let us note that this does not mean that there are three arbitrary constants. The arbitrary constant is actually only one, since a change in t_0 or in y_0 is the same as adding an arbitrary constant to the left-hand side of (24.6), which can be shifted to the right-hand side and absorbed in A. For this reason the limits of integration can also be left indeterminate and omitted, the solution being written as

$$\int P(t,y)\,\mathrm{d}t + \int Q(t,y)\,\mathrm{d}y = A. \tag{24.7}$$

It is however to be noted that the possibility of conveniently choosing t_0, y_0 in (24.6) may simplify the calculations (Ince, 1926, p. 17).

When condition (24.5) does not hold, the solution cannot be given by (24.6), and it is necessary to find an *integrating factor*, that is, a function $\mu(t,y)$ such that the expression

$$\mu(t,y)\,(P\mathrm{d}t + Q\mathrm{d}y)$$

is a total differential. In other words, the integrating factor must be such that the integrability condition is satisfied with respect to $\mu(t,y)\,P(t,y)$ and to $\mu(t,y)\,Q(t,y)$, i.e.

$$\frac{\partial\,[\mu(t,y)\,P(t,y)]}{\partial y} = \frac{\partial\,[\mu(t,y)\,Q(t,y)]}{\partial t}.$$

When an integrating factor has been found, the solution can be given as

$$\int [\mu(t,y)\,P(t,y)]\mathrm{d}t + \int [\mu(t,y)\,Q(t,y)]\mathrm{d}y = A.$$

Although it can be proved[3] that there exists an infinite number of integrating factors (so that Eq. (24.1) is always integrable, at least in principle), it turns out that the direct evaluation of μ requires the solution of a partial differential equation. That is, an equation of a more advanced type than the equation whose solution we are looking for. Fortunately, in many cases the partial differential equation has an obvious solution which gives the required integrating factor. We shall not pursue this matter further, turning instead to study two special cases of exact equations.

The first occurs when P depends on t alone and Q on y alone, that is

$$P(t)\,\mathrm{d}t + Q(y)\,\mathrm{d}y = 0. \tag{24.8}$$

The equation is then said to have *separated variables,* and the integrability condition is necessarily satisfied, since

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial t} = 0. \tag{24.9}$$

Therefore, the solution is

$$\int P(t)\,\mathrm{d}t + \int Q(y)\,\mathrm{d}y = A. \tag{24.10}$$

The second case occurs when P and Q are both functions of t and y, but such that they can be factorized into the product of a function of t alone and of a function of y alone, that is, say

$$\begin{aligned} P(t,y) &= T(t) \times Y_1(y), \\ Q(t,y) &= T_1(t) \times Y(y). \end{aligned} \tag{24.11}$$

The equation is than said to have *separable variables*, since it may be written—if we divide through by $T_1(t) \times Y_1(y)$—in the form

$$\frac{T(t)}{T_1(t)}\mathrm{d}t + \frac{Y(y)}{Y_1(y)}\mathrm{d}y = 0, \tag{24.12}$$

which has separated variables.

24.2.2 Linear equations of the first order with variable coefficients

The general first-order linear equation is

$$a_1(t)\frac{\mathrm{d}y}{\mathrm{d}t} + a_0(t)\,y = g(t), \tag{24.13}$$

[3]See, e.g., Ince (1956, pp. 27-9). The only condition is that Eq. (24.2) has a unique solution, that is, the conditions of the existence and uniqueness theorem are assumed to hold.

where $a_1(t), a_0(t), g(t)$ are known functions. As in constant-coefficient linear equations, Eq. (24.13) is said to be non-homogeneous; the corresponding homogeneous equation is

$$a_1(t)\frac{\mathrm{d}y}{\mathrm{d}t} + a_0(t)\,y = 0. \tag{24.14}$$

Setting

$$\begin{aligned} h(t) &\equiv \frac{a_0(t)}{a_1(t)}, \\ \varphi(t) &\equiv \frac{g(t)}{a_1(t)}, \end{aligned}$$

we can rewrite (24.13) and (24.14) as

$$\frac{\mathrm{d}y}{\mathrm{d}t} + h(t)\,y = \varphi(t), \tag{24.15}$$

$$\frac{\mathrm{d}y}{\mathrm{d}t} + h(t)\,y = 0. \tag{24.16}$$

Let us consider first Eq. (24.16). It can be written as

$$\mathrm{d}y + h(t)\,y\mathrm{d}t = 0, \tag{24.17}$$

which has separable variables: in fact, dividing through by y we have

$$\frac{1}{y}\mathrm{d}y + h(t)\,\mathrm{d}t = 0, \tag{24.18}$$

whose solution—see Eq. (24.10) in the previous section—is

$$\int \frac{1}{y}\mathrm{d}y + \int h(t)\,\mathrm{d}t = A, \tag{24.19}$$

that is,

$$\log_e y = A - \int h(t)\,\mathrm{d}t,$$

so that

$$y(t) = \exp\left[A - \int h(t)\,\mathrm{d}t\right]. \tag{24.20}$$

Setting $e^A \equiv C$, we have

$$y(t) = C\exp\left[-\int h(t)\,\mathrm{d}t\right]. \tag{24.21}$$

Let us now tackle Eq. (24.15). To solve it we shall apply Lagrange's method of *variation of parameters*[4]. This is a general method of solving

[4]This method was devised by John Bernoulli in relation to the first-order linear equation, but its generalization is due to Lagrange, and this explains why the method is usually referred to as Lagrange's.

an equation by considering as variables (as a function of t) the arbitrary constants which appear in the solution of a simpler equation, then trying to determine them in such a way that the equation to solve is identically satisfied. In our case we set

$$y(t) = C(t)\exp\left[-\int h(t)\,\mathrm{d}t\right], \tag{24.22}$$

where $C(t)$ is an undetermined function. Differentiating we have

$$\frac{\mathrm{d}y}{\mathrm{d}t} = \frac{\mathrm{d}C}{\mathrm{d}t}\exp\left[-\int h(t)\,\mathrm{d}t\right] - h(t)\,C(t)\exp\left[-\int h(t)\,\mathrm{d}t\right]. \tag{24.23}$$

Substituting (24.22) and (24.23) in (24.15) we obtain

$$\frac{\mathrm{d}C}{\mathrm{d}t}\exp\left[-\int h(t)\,\mathrm{d}t\right] = \varphi(t), \tag{24.24}$$

that is,

$$\frac{\mathrm{d}C}{\mathrm{d}t} = \varphi(t)\exp\left[\int h(t)\,\mathrm{d}t\right]; \tag{24.25}$$

which gives

$$C(t) = \int\left\{\varphi(t)\exp\left[\int h(t)\,\mathrm{d}t\right]\right\}\mathrm{d}t + B, \tag{24.26}$$

where B is an arbitrary constant. From (24.22) and (24.26) we have

$$y(t) = B\exp\left[-\int h(t)\,\mathrm{d}t\right] + \exp\left[-\int h(t)\,\mathrm{d}t\right]\int\left\{\varphi(t)\exp\left[\int h(t)\,dt\right]\right\}\mathrm{d}t, \tag{24.27}$$

which is the required solution of (24.15).

Let us note, finally, that the method expounded in this section can be applied also when the coefficients are constant. In fact, we have applied the method of variation of parameters to find a particular solution of the first-order linear constant-coefficient non-homogeneous differential equation (see Part II, Chap. 12, Sect. 12.2.6).

24.2.3 The Bernoulli equation

The non-linear first-order equation

$$\frac{\mathrm{d}y}{\mathrm{d}t} + h(t)\,y = \varphi(t)\,y^n, \tag{24.28}$$

where n is any real number (different from 0 and from +1), is known in the mathematical literature as the Bernoulli equation[5]. It can be put in linear form by change of dependent variable. If we put

$$z = y^{1-n}, \tag{24.29}$$

[5]The name derives from the fact that it was proposed for solution by James Bernoulli (in 1695); the method of solution expounded here was discovered by Leibniz in the following year.

then

$$\frac{\mathrm{d}z}{\mathrm{d}t} = (1-n)\,y^{-n}\frac{\mathrm{d}y}{\mathrm{d}t}. \tag{24.30}$$

Multiplying both members of (24.28) by $(1-n)\,y^{-n}$, we have

$$(1-n)\,y^{-n}\frac{\mathrm{d}y}{\mathrm{d}t} + (1-n)\,h(t)\,y^{1-n} = (1-n)\,\varphi(t), \tag{24.31}$$

and so, taking account of (24.29) and (24.30), we have

$$\frac{\mathrm{d}z}{\mathrm{d}t} + (1-n)\,h(t)\,z = (1-n)\,\varphi(t), \tag{24.32}$$

which is linear in z and can be solved by the methods expound in the previous section. Having found the function $z(t)$, from (24.29) we obtain

$$y(t) = [z(t)]^{\frac{1}{(1-n)}}, \tag{24.33}$$

which is the required solution of Eq. (24.28).

Bernoulli equations are often used in the theory of economic growth: see, for example, Chap. 13, Sect. 13.2.

24.3 Limit cycles and relaxation oscillations

Limit cycles have already been introduced in Chap 21, Sect. 21.3.2.3 and Fig. 21.13. We shall first give the general theory and then examine a specific case, that of relaxation equations, which are known to give rise to limit cycles under certain conditions.

24.3.1 Limit cycles: the general theory

Let us recall that a limit cycle is an isolated closed integral curve (also called *orbit*) to which all nearby paths approach from both sides in a spiral fashion. If the direction of the movement along the nearby path is towards (away from) the limit cycle, the latter is called *orbitally stable (unstable)*—see Chap. 21, Fig. 21.13.

Intermediate cases may also occur in which all paths on one side of the limit cycle approach it, while on the other side they move away from it; in such cases the limit cycle can be called *semi-stable*. The (local) stability of a limit cycle can be investigated by means of the following theorem:

Orbital stability theorem. Let $y_1 = y_1(t), y_2 = y_2(t)$ be a periodic motion (limit cycle) of the non-linear system

$$\begin{aligned}\frac{\mathrm{d}y_1}{\mathrm{d}t} &= \varphi_1(y_1, y_2),\\ \frac{\mathrm{d}y_2}{\mathrm{d}t} &= \varphi_2(y_1, y_2).\end{aligned} \tag{24.34}$$

This limit cycle is locally stable (unstable) if its *characteristic exponent*

$$h=\frac{1}{T}\int_0^T\left\{\frac{\partial\varphi_1\left[y_1(t),y_2(t)\right]}{\partial y_1}+\frac{\partial\varphi_2\left[y_1(t),y_2(t)\right]}{\partial y_2}\right\}\mathrm{d}t,$$

where T is the period of the oscillation, is respectively negative (positive). For a proof the reader is referred, for example, to Andronov et al., 1966, pp. 289-90 and 296-300.

In the above theorem we started with an existing limit cycle. But to determine whether a given system has limit cycles is not an easy matter. A sufficient condition for the *absence* of closed integral curves (that is, the non-satisfaction of this condition is a necessary condition for the presence of limit cycles) is given by the following theorem:

Negative criterion for existence of limit cycles. If a system has no singular points, then it cannot have limit cycles (Andronov et al., 1966, p. 306).

Another negative criterion is the more famous

Bendixson's negative criterion. Given system (24.34), if the expression $\partial\varphi_1/\partial y_1+\partial\varphi_2/\partial y_2$ does not change its sign (or vanish identically) within a region D of the phase plane, no closed path can exist in D (Andronov et al., 1966, p. 305; Minorsky, 1962, pp. 82-4).

Let us note that the expression $\partial\varphi_1/\partial y_1+\partial\varphi_2/\partial y_2$ is simply the *trace* of the Jacobian of system (24.34), which plays such an important role in the criterion. To show why, consider the linear approximation to the system. To get periodic motions we assume complex roots. We know (see Chap. 21, Table 21.1) that a zero trace means that the singular point is a centre. A centre is a closed path.

If, however, we want a limit cycle, identical vanishing of the trace will not do: all paths will be closed orbits (see the next section on the Lotka-Volterra equations). What we need are spirals converging to a closed orbit (see Chap. 21, Fig. 21.13). The inward spiral must be divergent from the singular point, which means a positive trace. The outward spiral must be convergent to the closed orbit, which means a negative trace. Hence the trace must change its sign in the region where the limit cycle lies.

We shall now state a set of *sufficient* conditions for the existence of limit cycles (Minorsky, 1962, p. 84; Sansone and Conti, 1964, Chap. IV, § IV.2.9):

Poincaré-Bendixson positive criterion. Let D indicate the finite domain in the phase plane, contained between two closed curves C_1 and C_2. Then, if (1) in D and on C_1 and C_2 no singular points exist, and (2) the integral curves passing through the points of C_1 and C_2 penetrate in D all for t increasing or for t decreasing, then D contains at least one limit cycle.

It can also be shown that the limit cycles contained in D must be alternately stable and unstable (the outermost and innermost being stable), so that if there is only one limit cycle it is necessarily stable.

It is important to note two points, that will be useful in practical applications of the Poincaré-Bendixson theorem.

(i) The expression 'closed curves' is not to be taken literally. A closed polygon will do as well.

(ii) Without loss of generality we can take C_2 as the outermost and C_1 as the innermost closed curve. Now suppose that the isolated singular point around which we are interested in finding a limit cycle is *locally unstable*. Then it only remains to find C_2, since C_1 automatically exists. In fact, any ϵ-neighbourhood of the singular point can be C_1: instability means that the integral curves passing through the points of C_1 will penetrate in D all for t increasing.

Necessary conditions for the existence of a closed trajectory (limit cycle) can be obtained using the concept of *index* introduced by Poincaré, which can also be useful to obtain some information about the phase portrait of a system: see Andronov et al. (1966, pp. 300-305) and Minorsky (1962, pp. 77-80).

Another useful theorem, which allows us to establish simultaneously the existence *and* the stability (or instability) of a limit cycle, is the following (Minorsky, 1962, p. 90):

Existence and stability of limit cycles (Nemitzky). If it is possible to determine two positive constants ρ_1 and ρ_2 ($\rho_2>\rho_1$) such that for ρ_1 the expression $y_1\varphi_1+y_2\varphi_2\geq 0$, and, for ρ_2, $y_1\varphi_1+y_2\varphi_2\leq 0$, and, if, moreover, the circular ring D between the circles of radii ρ_1 and ρ_2 has no singular points, then there exists a stable limit cycle in D. If the signs of $y_1\varphi_1+y_2\varphi_2$ are reversed, other conditions being the same, the limit cycle is unstable.

24.3.2 Limit cycles: relaxation oscillations

A topic closely related to limit cycles is that of periodic motions due to the so-called relaxation oscillations. Mathematically, these derive from a specific type of second-order non-linear autonomous differential equation, and we know that a second-order equation is equivalent to a 2×2 first-order system.

The name derives from the fact that the basic equation can be used in physics to express many relaxation phenomena, namely phenomena in which periodically energy is first accumulated and then released. An example is that of a hammer striking an object periodically. Actually the movement of the hammer takes place under *two* different regimes, one before and one after the hammer hits the object, i.e. before and after the energy is released through the shock. This discontinuity will introduce a discontinuity in the *speed* of the movement in the two phases, which means that the movement in each phase will be described by a *different* function (Georgescu-Roegen, 1951, pp. 116-117).

The aim of van der Pol's contribution was to approximate these two different functions by a single analytic function, which he achieved by considering the differential equation

$$\frac{d^2y}{dt^2} + \mu\left(y^2-1\right)\frac{dy}{dt} + y = 0,$$

for μ large. However, as Georgescu-Roegen observes (1951, p. 117), this veiled the real meaning of relaxation oscillations, which is the regime switch and the different speeds of movement in the two regimes. We shall come back to this problem in a later chapter (see Sect. 26.5.2).

From the mathematical point of view, relaxation oscillations are one of the few cases in which we are able to establish sufficient conditions for the existence of a *unique* periodic motion. We first consider the periodic motions of the second-order equation

$$\frac{d^2y}{dt^2} + f(y)\frac{dy}{dt} + g(y) = 0, \tag{24.35}$$

which is an equation of the Liénard type. This is a generalization of the van der Pol equation, which is clearly a particular case of the Liénard equation when $f(y) = \mu\left(y^2-1\right)$ and $g(y) = y$.

We then consider the periodic motions of the second-order equation

$$\frac{d^2y}{dt^2} + f\left(y, \frac{dy}{dt}\right)\frac{dy}{dt} + g(y) = 0, \tag{24.36}$$

which, in turn, is a generalization of Liénard equation. The following theorems are due to Levinson and Smith (1942) and have been successively sharpened by Dragilev, de Castro, Sansone and Conti and others (Minorsky, 1962, Chap. 4, Sect.2; Sansone and Conti, 1964, Chap. VI, Sects. VI.3 and VI.4).

Periodic motions in Liénard equations. The differential equation (24.35) has a unique periodic solution if the following conditions are satisfied:

(1) $f(y)$ and $g(y)$ are differentiable;

(2) there exist two positive numbers y_1, y_2 such that $f(y) < 0$ for $-y_1 < y < y_2$ and $f(y) \geq 0$ otherwise;

(3) $yg(y) > 0$ for $y \neq 0$;

(4) $\lim\limits_{y\to\pm\infty} F(y) = \lim\limits_{y\to\pm\infty} G(y) = +\infty$,
where

$$F(y) = \int_0^y f(y)\,dy, \quad G(y) = \int_0^y g(y)\,dy;$$

(5) $G(-y_1) = G(y_2)$.

Periodic motions in generalized Liénard equations. The differential equation (24.36) has at least one limit cycle if the following conditions are satisfied:

(1) $yg(y) > 0$ for $y \neq 0$;

(2) $\lim\limits_{y\to\pm\infty} G(y) = +\infty$, where

$$G(y) = \int_0^y g(y)\,dy;$$

(3) $f(0,0) < 0$;

(4) there exists a positive number y_0 such that $f(y, dy/dt) \geq 0$ for $|y| \geq y_0$;

(5) there exists a constant M such that, for $|y| \leq y_0$, $f(y, v) \geq -M$ for any v;

(6) there exists a $y_1 > y_0$ such that

$$\int_{y_0}^{y_1} f(y, v(y))\,dy \geq 10My_0,$$

where $v(y)$ is an arbitrary positive decreasing function of y.

The limit cycle is unique if, in addition:

(7) there exist two positive numbers y_1, y_2 such that $f(y, dy/dt) < 0$ for $-y_1 < y < y_2$ and $f(y, dy/dt) \geq 0$ otherwise;

(8) $\dfrac{dy}{dt}\dfrac{\partial f}{\partial(dy/dt)} \geq 0$;

(9) $G(-y_1) = G(y_2)$.

24.3.3 Kaldor's non-linear cyclical model

24.3.3.1 The model

A model that well lends itself to be studied by the various methods explained in this section is Kaldor's non-linear model of the business cycle (Kaldor, 1940). In fact, it is invariably present (under various forms) in modern treatments of business cycle theory (e.g. Gabisch and Lorenz, 1989; Sordi, 1990; Thygesen et al., 1991; Zarnowitz, 1992; Dore, 1993) notwithstanding its venerable age.

The original presentation was non-mathematical, and its early study was based on graphical techniques (see, e.g., Marrama, 1946). The first rigorous mathematical study is due to Ichimura (1955). Subsequent work includes Chang and Smyth (1971) and the modern treatments cited above. It is an impressive tribute to its author, and a demonstration of the importance of non-linearity, that Kaldor's business cycle model still yields stimuli to research: recently, Grasman and Wentzel (1994) have shown that in this model coexistence of a stable equilibrium and a stable limit cycle is possible.

The basic idea of this model, which is entirely Keynesian in spirit, is that the investment and the saving functions are both non-linear functions of output (income) and the capital stock.

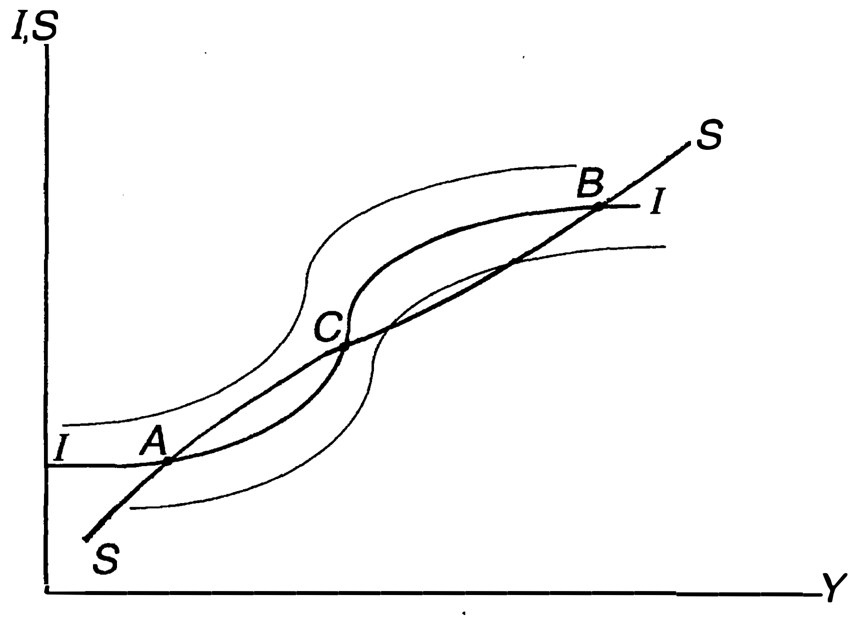


Figure 24.1: Kaldor's business cycle model: the investment and saving functions

Let us begin with the investment function, $I = I(Y, K)$. The marginal propensity to invest, $I_Y \equiv \partial I/\partial Y$, is positive but varying. More precisely, it has a 'normal' value for 'normal' mid-range values of Y. For decreasing values of Y below this normal interval, the marginal propensity to invest decreases, due to missing profit opportunities in periods of low activity levels relative to the 'normal' level. It also decreases for increasing values of Y above the normal interval, because of decreasing economies of scale and increasing financial costs.

It follows that the investment function has an S-shaped form as shown in Fig. 24.1, where the dependence of investment on the capital stock is also shown in the form of various curves. These are the projection on the (Y, I) plane of the function $I = I(Y, K)$ for different values of the capital stock. Kaldor assumes a negative relation between investment and the capital stock,

$I_K \equiv \partial I/\partial K < 0$. A higher capital stock means a lower marginal efficiency of capital, hence a lower level of investment for each income level.

Kaldor also assumes a non-linear saving function, $S = S(Y, K)$. The marginal propensity to save is positive and smaller than one but varying. This can be justified on the following grounds: suppose that there is a 'normal' level of the propensity to save corresponding to a normal range of income. Below that range of income, saving will be cut drastically in favour of consumption, while above that range it will be increased considerably. Thus we have an inverted S-shape that is the mirror image of the investment function.

Kaldor assumes that, in the normal range, $I_Y > S_Y$, so that the equilibrium point falling in the normal range (point C in Fig. 24.1) is unstable on the usual Keynesian behaviour assumption $Y' = \alpha[I(Y, K) - S(Y, K)]$, while extreme points like A and B are stable.

In addition, Kaldor assumes that $S_K \equiv \partial S/\partial K > 0$, which means that the saving function shifts upwards when the capital stock increases. Most commentators seem to agree with Ichimura (1955) that this effect is not explained satisfactorily by Kaldor. We do not agree with this view, because formalisations often compel us to cast a verbal model in a straitjacket. Now, Kaldor did not use Y as the argument of his investment and saving functions, but the 'level of activity (measured in terms of employment)' (1940; p. 178 of the 1960 reprint). Hence it was perfectly legitimate to assume that 'When activity is high, the level of investment is high, the total amount of equipment gradually increases, and so, in consequence, the amount of consumers' goods produced at a given level of activity. As a result the S curve gradually shifts upwards (because there will be more consumption, and hence more saving, for any given activity)' (1940; p. 182 of the 1960 reprint). In other words, at any given level of 'activity' (employment) there will be more output when the capital stock is higher, which means both more consumption and more saving.

By choosing to use Y (a necessary choice if we wish to avoid the explicit introduction of employment into the formal model) we lose this explanation, which nonetheless remains perfectly plausible in the context of the model. To identify S_K with a standard wealth effect, as Chang and Smyth (1971) do (consequently they take $S_K < 0$ and are compelled to make additional assumptions to be able to apply the Poincaré-Bendixson theorem) is not correct. Hence we shall keep with Kaldor in assuming $S_K > 0$.

The graphical analysis of the model is fairly simple. In Fig. 24.2, adapted from Kaldor (1940, Fig. 6) we start from stage I in which the model is in equilibrium; since C is an unstable equilibrium point, the model will converge rapidly to either A or B. Let us assume that it is at B. At this high level of activity, capital accumulation will shift the saving curve upwards and the investment curve downwards, causing a decrease in activity (stage II). These gradual shifts will bring the I and S curves tangent to each other; since point B+C is unstable in the downward direction, the system will rapidly

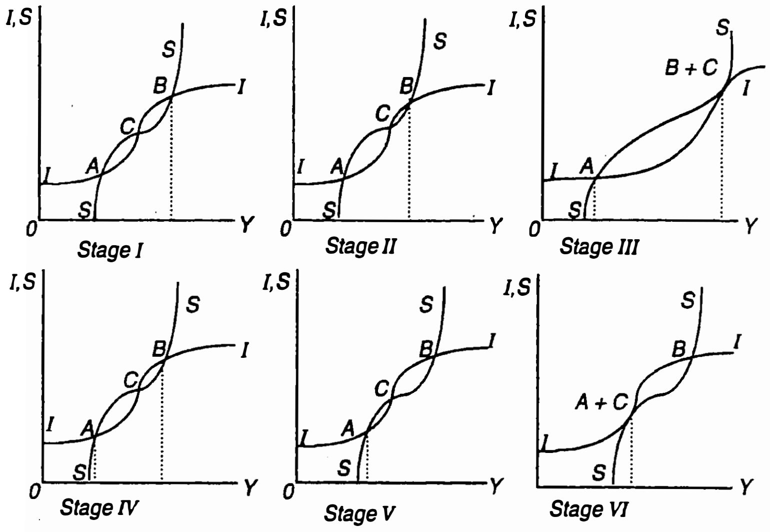


Figure 24.2: The dynamics of Kaldor's cycle

drop to the stable equilibrium A (stage III). At this low level of activity, investment is not sufficient to cover replacement, hence capital decumulation will lower the S curve and shift the I curve upwards, which means that point A will move gradually to the right and points B and C will again become separated (stages IV and V). These shifts will bring the investment and saving curves again tangent to each other; since point $A+C$ is unstable in the upward direction, an upward cumulative movement will follow which can only come to rest when position B is reached (stage VI). Thereafter the curves gradually return to the position shown in stage I, and the cyclical movement is repeated.

Let us note an important point in this verbal description: when an equilibrium becomes unstable, the movement to a stable equilibrium takes place *before* the curves shift again, namely it takes less time to adjust output than it takes the change in the stock of capital to have an appreciable effect on the level of investment and saving at a given level of output. This condition on the faster adjustment speed of output with respect to other variables is implicit in all short-run Keynesian models, and will be found very important in the subsequent mathematical analysis.

Let us now come to the mathematics. The model can be summarized in

the following equations:

$$\begin{aligned} Y' &= \varphi_1(Y,K) \equiv \alpha[I(Y,K) - S(Y,K)], && I_Y > 0, S_Y > 0, I_K < 0, S_K > 0, \\ & I_Y > S_Y \text{ for } Y_n \le Y \le Y_N, && I_Y < S_Y \text{ elsewhere}, \\ K' &= \varphi_2(Y,K) \equiv \overline{I}(Y,K). \end{aligned} \tag{24.37}$$

The range $Y_n \le Y \le Y_N$ is the normal range which was explained above. $\overline{I}(Y,K)$ is *realized investment*, to be distinguished from $I(Y,K)$, which is *planned* (or *ex ante*) investment. Note that for simplicity's sake we have neglected depreciation. The actual specification of realized investment is different according to the various formal treatments of the model. To these we now turn.

Kaldor's model can be examined, and has in fact been examined, by using all the main tools for detecting oscillations: the Liénard equation on relaxation oscillations (Ichimura, 1955), the Poincaré-Bendixson theorem (Chang and Smyth, 1971), and the Hopf bifurcation (see the next chapter, Sect. 25.2.4).

24.3.3.2 Kaldor via relaxation oscillations

Relaxation oscillations as a possible explanation of business cycles were called to the attention of economists as long as the early 1930s by the economist Hamburger (1931, 1935: see Sordi, 1988) and the physicist Le Corbeiller (1933), but without much follow-up. If we except Goodwin's (1951) nonlinear business cycle model and Georgescu-Roegen's 1951 paper, Ichimura (1955) was the first systematically to exploit relaxation oscillations in the study of various business cycle models, amongst which the Kaldor model.

In model (24.37) we differentiate the goods market adjustment equation with respect to time and obtain

$$Y'' = \alpha(I_Y Y' + I_K K' - S_Y Y' - S_K K'). \tag{24.38}$$

Rearrangement of terms and substitution from the second equation yields

$$Y'' - \alpha(I_Y - S_Y)Y' - \alpha(I_K - S_K)\overline{I}(Y,K) = 0. \tag{24.39}$$

This is not yet in the required form, since the capital stock still appears as a second variable. Besides the presence of realized investment, we must also take into account that, except in the linear case, the partial derivatives of a function depend on the same variables which are the arguments of the function being differentiated. Hence I_Y, I_K, S_Y, S_K are functions of both Y *and* K.

There are various possibilities for dealing with this problem.

One is to assume that realized investment coincides with *ex ante* investment, which is taken to be independent of the capital stock, and that the saving function (whose shifts will be the only cause of the cycle) is linear in the capital stock. A second one is to assume that the actual change in the capital stock is determined by savings decisions, namely $K' = \overline{I} = S$, where S is taken to be independent of the capital stock, and that the investment function (whose shifts will now be the only cause of the cycle) is linear in K.

Other possibilities also exist (Kosobud and O'Neil, 1972), but we follow Ichimura (1955) and Gabisch and Lorenz (1989, Sect. 4.3.4.2) in considering the second alternative. Hence we have

$$Y'' - \alpha(I_Y - S_Y)Y' - \alpha I_K S(Y) = 0. \tag{24.40}$$

Given the assumptions, I_Y, S_Y, I_K are independent of K.

By a change of coordinates we can take the singular point as the origin, namely

$$y'' - \alpha(i_y - s_y)y' - \alpha i_k s(y) = 0, \tag{24.41}$$

where the lower-case letters denote the deviations of the corresponding upper-case letters from their respective equilibrium values. Equation (24.41) is an equation of the Liénard type, with $f(y) = -\alpha(i_y - s_y)$ and $g(y) = -\alpha i_k s(y)$. To show that a periodic solution exists, we have to check the conditions required by the theorem stated in Sect. 24.3, that we rewrite here for the reader's convenience, together with the proof of the satisfaction of each condition.

(1) $f(y)$ and $g(y)$ are differentiable.

Satisfied.

(2) there exist two positive numbers y_1, y_2 such that $f(y) < 0$ for $-y_1 < y < y_2$ and $f(y) \geq 0$ otherwise.

Let us consider the 'normal' range $Y_n \leq Y \leq Y_N$, that includes the equilibrium point Y_e. We take $y_1 = Y_e - Y_n$ and $y_2 = Y_N - Y_e$; by definition y_1, y_2 are both positive. In the normal range by definition we have $i_y - s_y > 0$, hence $f(y) < 0$. Outside the normal range we have $i_y - s_y < 0$, hence $f(y) > 0$. Thus the condition is satisfied.

(3) $yg(y) > 0$ for $y \neq 0$.

Since $S(Y) > S(Y_e)$ for $Y > Y_e$ and $S(Y) < S(Y_e)$ for $Y < Y_e$, we have $s(y) \gtrless 0$ for $y \gtrless 0$. Hence condition (3) is satisfied.

(4) $\lim\limits_{y\to\infty} F(y) = \lim\limits_{y\to\infty} G(y) = \infty$,

where

$$F(y) = \int_0^y f(y)\,\mathrm{d}y, \quad G(y) = \int_0^y g(y)\,\mathrm{d}y.$$

We have $F(y) = \int_0^y [-\alpha(i_y - s_y)]\mathrm{d}y = \alpha[s(y) - i(y)]$. Given the shapes of the saving and investment functions (see Fig. 24.1), this expression obviously

tends to infinity as y tends to infinity. We also have $G(y) = \int_0^y [-\alpha i_k s(y)]\mathrm{d}y = -\alpha i_k \int_0^y s(y)\mathrm{d}y$. Since $-\alpha i_k > 0$, and $s(y)$ is increasing as y increases, $G(y)$ tends to infinity as $y \to \infty$. Condition (4) is satisfied.

(5) $G(-y_1) = G(y_2)$.

This condition requires $\int_0^{-y_1} s(y)\mathrm{d}y = \int_0^{y_2} s(y)\mathrm{d}y$.

Since the extremes of the normal range can be taken as fairly symmetric with respect to Y_e, namely with respect to zero in terms of the transformed coordinates, the integration intervals will have the same (absolute) length.

Thus the condition will be satisfied if the saving function is symmetrical in the normal range, namely $s(y) = -s(-y)$ or $-s(y) = s(-y)$ (an *odd* function in mathematical terms).

This means that when income is *higher* than equilibrium income by a certain amount, saving will be *higher* than equilibrium saving by an amount which is equal to the (absolute value of the) amount by which saving falls short of equilibrium saving when income is *lower* than equilibrium income by the same amount as it was higher before. An S-shaped function like that postulated by Kaldor might well satisfy this requirement, but it is not certain.

In conclusion, while conditions (1)-(4) are automatically satisfied, the symmetry condition (5) requires an additional assumption for the emergence of a cycle.

24.3.3.3 Kaldor via Poincaré's limit cycle

In the Chang and Smyth (1971) interpretation, realized and planned investment coincide, hence the second equation of the model is $K' = I(Y, K)$. Let us first build the phase diagram of the model around the (unstable) intermediate equilibrium point, using the technique explained in Chap. 21, Sect. 21.3.2.3. From $\varphi_1(Y, K) = \alpha[I(Y, K) - S(Y, K)] = 0$ we can compute

$$\left.\frac{\mathrm{d}K}{\mathrm{d}Y}\right|_{\varphi_1=0} = \frac{S_Y - I_Y}{I_K - S_K}. \tag{24.42}$$

Since $I_K - S_K < 0$, while $S_Y - I_Y < 0$ in the neighbourhood of the equilibrium point, i.e. in the intermediate range of 'normal' values of income, $Y_n \leq Y \leq Y_N$, and $S_Y - I_Y > 0$ elsewhere, this curve will be positively sloped in the 'normal' range and negatively sloped elsewhere.

From $\varphi_2 = I(Y, K) = 0$ we compute

$$\left.\frac{\mathrm{d}K}{\mathrm{d}Y}\right|_{\varphi_2=0} = -\frac{I_Y}{I_K} < 0, \tag{24.43}$$

which shows that this curve is positively sloped everywhere.

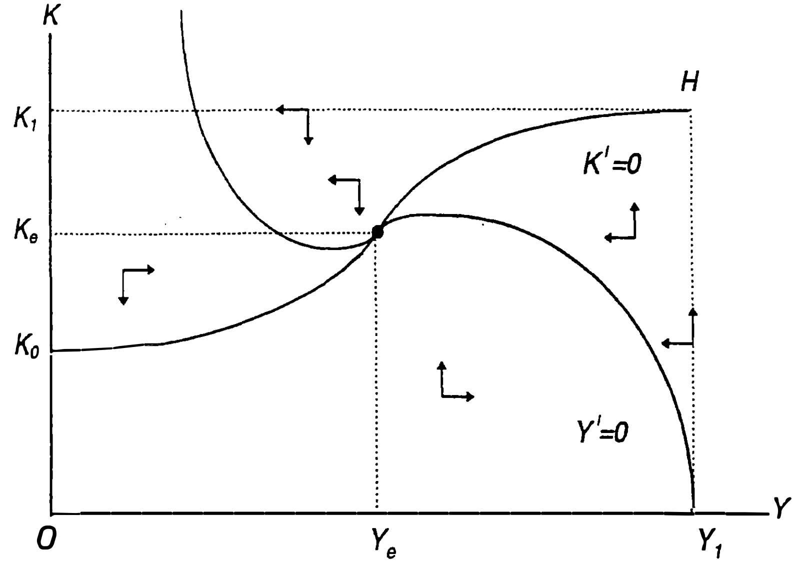


Figure 24.3: Kaldor's model via Poincaré-Bendixson

We now consider points not on the $\varphi_1 = 0$ curve. Since $\partial\varphi_1/\partial K = I_K - S_K < 0$, at all points *above (below)* the $\varphi_1 = 0$ we have $\varphi_1 < 0$ ($\varphi_1 > 0$), hence the horizontal arrows respectively pointing to the left (to the right).

By considering points not on the $\varphi_2 = 0$ we can similarly see that, since $\partial\varphi_2/\partial K = I_K < 0$, at all points *above (below)* the $\varphi_2 = 0$ we have $\varphi_2 < 0$ ($\varphi_2 > 0$), hence the vertical arrows respectively pointing downwards (upwards).

Let us now consider the Jacobian matrix of the model

$$\begin{bmatrix} \alpha(I_Y - S_Y) & \alpha(I_K - S_K) \\ I_Y & I_K \end{bmatrix}, \tag{24.44}$$

whose characteristic equation is

$$\begin{aligned} &\lambda^2 + a_1\lambda + a_2 = 0, \\ &\text{where} \\ &a_1 \equiv -\left[\alpha(I_Y - S_Y) + I_K\right], \\ &a_2 \equiv \alpha\left[(I_Y - S_Y)I_K - I_Y(I_K - S_K)\right]. \end{aligned} \tag{24.45}$$

With the assumed signs of the partial derivatives we have $a_2 > 0$. As regards a_1, the assumption $(I_Y - S_Y) > 0$ is not sufficient to determine its

sign. To have local instability we require $a_1 < 0$, namely

$$\alpha(I_Y - S_Y) + I_K > 0. \tag{24.46}$$

This assumption—which is not stated explicitly in the text of Kaldor's paper, but recalled in the appendix, where he compares his model with a similar model by Kalecki (Kaldor, 1940, p. 90; p. 190 of the 1960 reprint)—means that a movement along the I or S curve proceeds more speedily in time than the rate of shift of these curves due to $\mathrm{d}K/\mathrm{d}t$. In other words it takes less time to adjust output to a change in investment than it takes to change investment (at a given level of output) on account of the change in the stock of capital.

This, as Kaldor correctly noted in his reply to Chang and Smyth (Kaldor, 1971) is implicit in *all* short-period Keynesian models.

To apply the Poincaré-Bendixson theorem is now easy. Since the singular point is locally unstable, we only have to find a closed 'curve' C_2 such that the trajectories point inwards. This will be a polygon, and precisely the rectangle OK_1HY_1. The integral curves, which are (strictly) included between the two arrows, point to the interior. Hence the conditions of the theorem are satisfied, and a limit cycle exists.

The elegant model by Kaldor does have a cycle under Kaldor's assumptions. Hence we do not agree with the statement by Chang and Smyth (1971), usually taken for granted by subsequent commentators, that Kaldor's analysis is rather incomplete and the existence of a cycle cannot be established unless additional assumptions are made. The reason for this incorrect statement is that Chang and Smyth didn't see the presence of Kaldor's assumption (24.46), and besides changed Kaldor's assumption $S_K > 0$ into $S_K < 0$ (see above, Sect. 24.3.3.1).

24.4 The Lotka-Volterra equations

Here we are in a case in which the system possesses infinitely many closed orbits. Hence, whatever the initial position (except the origin), the motion is always periodic along the closed orbit on which the initial point happens to be. These equations can be completely analysed, in the sense that we can obtain the analytical form of the integral curves although the closed-form solution for $y_1(t), y_2(t)$ cannot be obtained. Hence a complete analytical characterization of the properties of the solution can be given.

In the mathematical literature[6] such equations are sometimes attributed solely to Volterra. However, this particular aspect of Volterra's work on mathematical biology (prey-predator interaction) was anticipated by Lotka,

[6]See, for example, Andronov et al., (1966, pp. 142-5), and Minorsky (1962, pp. 65-70). The following mathematical treatment is based on those works, as well as on Volterra's.

who developed and studied the same equations earlier than Volterra[7]. This is why we think that the denomination 'Lotka-Volterra equations' is more appropriate. Here we shall, of course, leave aside the biological problems[8] which gave rise to such equations, and examine them from the purely mathematical point of view.

The equations under consideration have the form

$$\begin{aligned} \frac{\mathrm{d}y_1}{\mathrm{d}t} &= (a_1 - b_1 y_2)\, y_1, \\ \frac{\mathrm{d}y_2}{\mathrm{d}t} &= -(a_2 - b_2 y_1)\, y_2, \end{aligned} \tag{24.47}$$

where a_1, b_1, a_2, b_2 are positive constants and only non negative values of y_1, y_2 are considered. The integral curves of (24.47) can be obtained as follows. Multiply the first equation by a_2/y_1, the second by a_1/y_2 and add, obtaining

$$\frac{a_2}{y_1}\frac{\mathrm{d}y_1}{\mathrm{d}t} + \frac{a_1}{y_2}\frac{\mathrm{d}y_2}{\mathrm{d}t} = -a_2 b_1 y_2 + a_1 b_2 y_1, \tag{24.48}$$

that is,

$$a_2 \frac{\mathrm{d}\log y_1}{\mathrm{d}t} + a_1 \frac{\mathrm{d}\log y_2}{\mathrm{d}t} = -a_2 b_1 y_2 + a_1 b_2 y_1, \tag{24.49}$$

where, of course, the logarithms are to the base e.

Now, multiply the first equation of (24.47) by b_2, the second by b_1 and add, obtaining

$$b_2 \frac{\mathrm{d}y_1}{\mathrm{d}t} + b_1 \frac{\mathrm{d}y_2}{\mathrm{d}t} = a_1 b_2 y_1 - a_2 b_1 y_2. \tag{24.50}$$

Since the right-hand members of (24.49) and of (24.50) are equal, the left-hand member must also be equal, so that

$$-a_2 \frac{\mathrm{d}\log y_1}{\mathrm{d}t} - a_1 \frac{\mathrm{d}\log y_2}{\mathrm{d}t} + b_2 \frac{\mathrm{d}y_1}{\mathrm{d}t} + b_1 \frac{\mathrm{d}y_2}{\mathrm{d}t} = 0. \tag{24.51}$$

This differential equation is directly integrable and yields the single valued integral

$$b_2 y_1 + b_1 y_2 - a_2 \log y_1 - a_1 \log y_2 = A, \tag{24.52}$$

where A is an arbitrary constant. Another way of arriving at the same result is to follow the standard procedure for obtaining the integral curves explained in Chap. 21, Sect. 21.3.2.1, that is, to eliminate $\mathrm{d}t$ from the original system

[7]See Lotka (1956, pp. 88-92); see also E.T. Whittaker's *Biography of Vito Volterra*, reprinted in Volterra's *Theory of Functionals*, etc., pp. 20-1 of the reprint, and Volterra (1927, p. 4, fn. 1; p. 2, fn. 2 of the reprint).

[8]The interested reader can consult Volterra (1927, 1931) and Lotka, *loc. cit.*

and to integrate the resulting differential equation. Eliminating $\mathrm{d}t$ from system (24.47) we have

$$\frac{\mathrm{d}y_2}{\mathrm{d}y_1} = -\frac{(a_2 - b_2 y_1)\, y_2}{(a_1 - b_1 y_2)\, y_1},$$

that is,

$$-(a_1 - b_1 y_2)\, y_1 \mathrm{d}y_2 - (a_2 - b_2 y_1)\, y_2 \mathrm{d}y_1 = 0.$$

The variables are separable (see above, Sect. 23.2.1) through the division by $y_1 y_2$, so that

$$-\left(a_1 y_2^{-1} - b_1\right) \mathrm{d}y_2 - \left(a_2 y_1^{-1} - b_2\right) \mathrm{d}y_1 = 0.$$

Integrating we have

$$-\int \left(a_1 y_2^{-1} - b_1\right) \mathrm{d}y_2 - \int \left(a_2 y_1^{-1} - b_2\right) \mathrm{d}y_1 = A,$$

that is,

$$-a_1 \log y_2 + b_1 y_2 - a_2 \log y_1 + b_2 y_1 = A,$$

which is the same as the Eq. (24.52).

Setting $B \equiv e^A$, Eq. (24.52) can be rewritten as

$$\exp(b_2 y_1) \exp(b_1 y_2)\, y_1^{-a_2} y_2^{-a_1} = B, \tag{24.53}$$

which gives

$$y_1^{-a_2} \exp(b_2 y_1) = B y_2^{a_1} \exp(-b_1 y_2). \tag{24.54}$$

Now, consider the functions

$$\begin{aligned} X_1 = X_1(y_1) &= y_1^{-a_2} \exp(b_2 y_1), \\ X_2 = X_2(y_2) &= y_2^{a_1} \exp(-b_1 y_2). \end{aligned} \tag{24.55}$$

The required integral curves are determined by the relation

$$X_1 = B X_2, \tag{24.56}$$

that is, are obtained equating the function X_1 to the function X_2 multiplied by an arbitrary constant. Of course, to each value of the arbitrary constant B there corresponds one integral curve.

In order to construct the integral curves, let us first investigate the form of the functions $X_1(y_1)$ and $X_2(y_2)$. We have

$$\frac{dX_1}{dy_1} = -a_2 y_1^{-a_2-1} \exp(b_2 y_1) + b_2 y_1^{-a_2} \exp(b_2 y_1) = X_1 \left(b_2 - \frac{a_2}{y_1}\right), \tag{24.57}$$

from which we see that $\mathrm{d}X_1/\mathrm{d}y_1 = 0$ for $y_1 = a_2/b_2$ and that $\mathrm{d}X_1/\mathrm{d}y_1$ is always negative for $0 \le y_1 < a_2/b_2$ and always positive for $y_1 > a_2/b_2$.

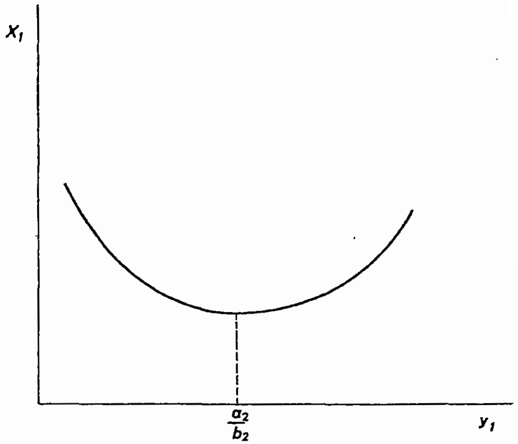


Figure 24.4: Lotka-Volterra equations, diagram 1

Thus the shape of X_1 is as shown in Fig. 24.4. The curve has been drawn everywhere convex to the y_1-axis since the second derivative $\mathrm{d}^2X_1/\mathrm{d}y_1^2$ is always positive for $y_1 \geq 0$.[9]

With regard to X_2, we have

$$\frac{\mathrm{d}X_2}{\mathrm{d}y_2} = a_1 y_2^{a_1-1} \exp\left(-b_1 y_2\right) - b_1 y_2^{a_1} \exp(-b_1 y_2) = X_2 \left(\frac{a_1}{y_2} - b_1\right), \quad (24.58)$$

from which we see that $\mathrm{d}X_2/\mathrm{d}y_2 = 0$. for $y_2 = a_1/b_1$ and that $\mathrm{d}X_2/\mathrm{d}y_2$ is always positive for $0 \leq y_2 < a_1/b_1$ and always negative for $y_2 > a_1/b_1$.

[9] It turns out that

$$\frac{\mathrm{d}^2X_1}{\mathrm{d}y_1^2} = X_1 \left[(a_2+1)\, a_2 y_1^{-2} - 2a_2 b_2 y_1^{-1} + b_2^2\right].$$

Consider the inequality $(\mathrm{d}^2X_1/\mathrm{d}y_1^2) \gtreqless 0$. Since $X_1 > 0$ for $y_1 \geq 0$, we can consider only the inequality

$$(a_2+1)\, a_2 y_1^{-2} - 2a_2 b_2 y_1^{-1} + b_2^2 \gtreqless 0.$$

For $y_1 = 0$, it certainly holds with the $>$ sign; so let us multiply through by $y_1^2, (y_1 > 0)$ and consider the inequality

$$f(y_1) = b_2^2 y_1^2 - 2a_2 b_2 y_1 + (a_2+1)\, a_2 \gtreqless 0.$$

It can be checked that the equation $f(y_1) = 0$ has no real roots, so that the inequality is always satisfied with the $>$ sign for $y_1 > 0$.

Therefore the shape of X_2 is as shown in Fig. 24.5. Actually the curve has inflection points[10], but we have neglected them for graphical simplicity[11].

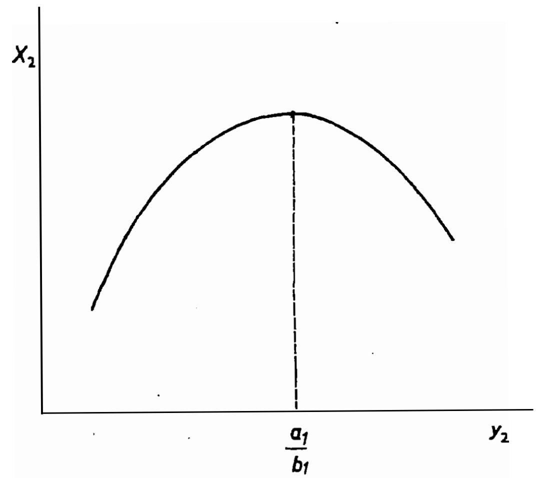


Figure 24.5: Lotka-Volterra equations, diagram 2

[10] It turns out that

$$\frac{\mathrm{d}^2X_2}{\mathrm{d}y_2^2} = X_2 \left[a_1 (a_1-1)\, y_2^{-2} - 2a_1 b_1 y_2^{-1} + b_1^2\right].$$

Consider the inequality $\mathrm{d}^2X_2/\mathrm{d}y_2^2 \gtreqless 0$. For $y_2 = 0$ it is not possible to know what happens without knowing the magnitude of a_1. Let us consider the range $y_2 > 0$, where $X_2 > 0$, so that we may examine only the inequality $a_1(a_1-1)\, y_2^{-2} - 2a_1 b_1 y_2^{-1} + b_1^2 \gtreqless 0$, that is

$$h(y_2) = b_1^2 y_2^2 - 2a_1 b_1 y_2 + a_1(a_1-1) \gtreqless 0.$$

The equation $h(y_2) = 0$ has two real roots, $a_1/b_1 \pm \sqrt{a_1/b_1}$, which determine the inflection points, that are symmetrical with respect to the maximum point of the curve X_2. The one to the right is certainly relevant; regarding the one to the left, since—as we said at the beggining of this section—only the range $y_2 \geq 0$ is considered, its relevance or not depends on the magnitude of a_1.

[11] In Andronov et al. (1966, pp. 144-5) the inflection points are not even mentioned and the curve is drawn as in Fig. 24.5 after consideration of $\mathrm{d}X_2/\mathrm{d}y_2$ only.

24.4.1 Construction of the integral curves

We can now construct the integral curves by means of Fig. 24.6[12]. In the second and fourth quadrant the curves X_2 and X_1 found above are drawn; in the third quadrant the straight line represents Eq. (24.56). Let us take an arbitrary point P_0 on the line OK. Draw from it two straight lines, one perpendicular to the OX_1 axis and the other perpendicular to the OX_2 axis. Call D, E, F, G the points of intersection of those lines with the X_1 and X_2 curves. From points D and E draw two straight lines parallel to the OX_1 axis and from F and G draw two straight lines parallel to the OX_2 axis. The four points of intersection of these four straight lines (points 1, 2, 3, 4) belong to the integral curve $X_1(y_1) = BX_2(y_2)$. In fact, each one of these points by construction is such as to equate $X_1(y_1)$ to $BX_2(y_2)$. The locus of these points when the point P slides along the OK line the interval P'P" is the required integral curve for the value tan α of the arbitrary constant B. To each value of B an integral curve corresponds and can be constructed in the same way. All such curves are closed (except one corresponding to the coordinate axes), so that the state of equilibrium (i.e. the singular point whose coordinates are $y_1 = a_2/b_2, y_2 = a_1/b_1$) is a *centre*. The origin too is a state of equilibrium, since $\mathrm{d}y_1/\mathrm{d}t = \mathrm{d}y_2/\mathrm{d}t = 0$ also for $y_1 = y_2 = 0$, but this point does not interest us. Note that the singular point in the origin is of the *saddle-point* type.

By inspection of the diagram and using Eqs. (24.47) it can be seen that the direction of the movement along the integral curve is that shown by the arrows (anticlockwise). Take, for example, point 2. There y_2 is greater than a_1/b_1, so that $a_2 - b_1y_2 < 0$ and $\mathrm{d}y_1/\mathrm{d}t < 0$ (y_1 decreases); y_1 is smaller than a_2/b_2, so that $a_2 - b_2y_1 > 0$ and $\mathrm{d}y_2/\mathrm{d}t < 0$ (y_2 decreases). Therefore the point travels in an anticlockwise direction. This of course depends on how we draw the diagram. If we had put y_2 on the horizontal, and y_1 on the vertical, axis, the resulting integral curves would have been traversed in a clockwise direction.

As the representative point travels around the integral curve, y_1 oscillates between points y_{11} and y_{12}, and y_2 oscillates between points y_{21} and y_{22}. Given the initial conditions, the slope of OK (and so the corresponding integral curve) is determined, as well as the point on the integral curve from which the movement starts.

It is also interesting to note that any external shock simply brings about a shift from one to another integral curve, where the system returns to its endless periodic motion.

We finally observe that the Lotka-Volterra system is *not* structurally sta-

[12] Andronov et al. (1966, p. 144), Volterra (1931, p. 16); we have changed the positions of the various curves so as to obtain the integral curves in the north-east quadrant.

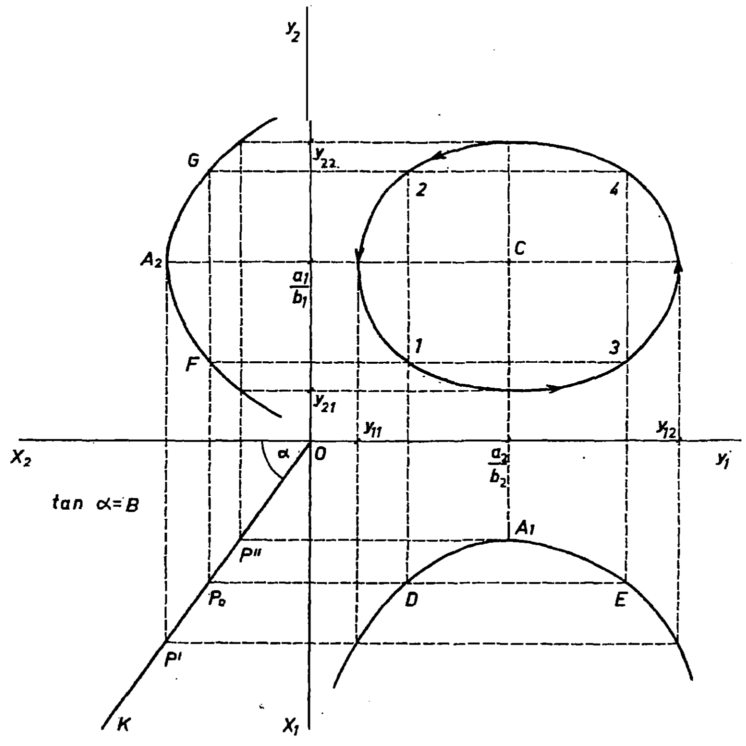


Figure 24.6: Lotka-Volterra equations, construction of the orbits

ble: consider, for example, the modified system

$$\begin{aligned}\frac{\mathrm{d}y_1}{\mathrm{d}t} &= (a_1 - b_1y_2 - py_1)\,y_1 = (a_1 - b_1y_2)\,y_1 - py_1^2, \\ \frac{\mathrm{d}y_2}{\mathrm{d}t} &= -(a_2 - b_2y_1 + qy_2)\,y_2 = -(a_2 - b_2y_1)\,y_2 - qy_2^2,\end{aligned} \tag{24.59}$$

where p, q are small positive constants. The additional terms are not arbitrary, since they have a biological interpretation (Hirsch and Smale, 1974, p. 263). But what interests us is to note that, from the mathematical point of view, for p, q sufficiently small we are in the context described in Chap. 21, Sect. 21.2.3, Eqs. (21.5). Now, system (24.59) possesses a stable equilibrium point at the origin, and a stable limit cycle in the remaining positive quadrant of the phase space (Hirsch and Smale, 1974, pp. 263-265), hence the original system (24.47) is not structurally stable. Generalizations of the Lotka-Volterra equations usually show stability, sometimes even global asymptotic stability (Hsu and Huang, 1995).

24.4.2 Conservative and dissipative systems, and irreversibility

Let us note that the Lotka-Volterra system that we have examined in the previous section is a non-linear 'conservative' system. This name is drawn from physical systems, which are called conservative when their total energy (potential plus kinetic) remains constant during the motion over time. A conservative system may be thought of as a system in which there is no friction (which would imply dissipation of energy). For non-physical systems the 'conservative' character is attributed on purely formal grounds.
From the formal point of view, a system $\mathrm{d}y_i/\mathrm{d}t = \varphi_i(y_1, y_2)$, $i = 1, 2$, is *conservative* in a domain D if it has a first integral in D, and D has the property that every trajectory having one point in D lies entirely in D for $t \to +\infty$ or for $t \to -\infty$.
A *first integral* is a single-valued differentiable function $F(y_1, y_2)$ defined on D and not identically constant, such that $F(y_1, y_2) =$ constant when $y_1 = y_1(t)$, $y_2 = y_2(t)$, $y_i(t)$ being a solution of the differential system.

By contrast, *dissipative* systems are those where total energy does not remain constant but is dissipated during the motion. For a general treatment of non-linear conservative and dissipative systems from both the physical and the purely formal point of view, see Andronov et al. (1966, Chaps. II and III) and Minorsky (1962, Chap. 2).

An alternative definition of a conservative system is sometimes given, based on the notion of *divergence* or *Lie derivative*, which is simply the trace of the Jacobian of the functions φ_i. More precisely, a general $n \times n$ system is said to be conservative if

$$\sum_{i=1}^{n} \frac{\partial \varphi_i}{\partial y_i} = 0$$

in the whole state space, while it is said to be *dissipative* if

$$\sum_{i=1}^{n} \frac{\partial \varphi_i}{\partial y_i} < 0$$

in the whole state space. This definition, however, cannot be accepted as equivalent to the one based on the first integral.

In fact, it can easily be checked that the Lotka-Volterra equations have a constant first integral in the whole phase space (excluding the origin and the axes): a first integral is, for example, the function on the left-hand side of (24.53). However, it is easy to see that these equations have a Jacobian whose trace is zero at the equilibrium point (see Exercise 5), but not elsewhere.

Hence the definition based on the Lie derivative must be taken with caution: it is true that all conservative systems have a Jacobian with a zero trace at the equilibrium point, but not necessarily elsewhere.

By contrast, dissipative systems possessing a limit cycle must have a Jacobian whose trace changes sign in the region where the limit cycle lies. This has been explained above, Sect. 23.3, in relation to the Bendixson criterion.

Let us conclude this section with a few words on reversibility and irreversibility. Given a differential equation system and its solution, we can compute the time path of the system not only for $t \to +\infty$ (i.e., for the future) but also for $t \to -\infty$ (i.e., for the past). If we know where the system is now, we can calculate not only where it will be at any point in the future, but also—by simply reversing the direction of the time variable—also where it was at any point in the past. Mathematically, time is *reversible.*

This reasoning is however fully warranted only in *conservative* systems. Since all paths of a conservative system are closed orbits, even if we do not know the initial point exactly, the calculated trajectory starting at a slightly different point in the phase space will be on an nearby orbit having the same nature. Hence both the future and the past can be calculated. The typical case where all this is true is celestial mechanics, where many phenomena can be described by conservative dynamical systems, and where predictions on the position of the planets in the solar system in the future and in the past are remarkably precise.

However, when a system is *dissipative*, the situation is completely different. Suppose for example that the equilibrium point of the system is a stable focus (see Chap. 21, Sect. 21.3.2.2), and that the system is now in a neighbourhood of the equilibrium point. If we want to predict the future path of the system, an imprecision on the initial (current) situation has no serious consequence: the 'wrong' path will stay close to (and approach more and more) the 'true' one, since both are convergent. On the contrary, if we want to answer the question, where was the system in the past, we should run time backwards on the appropriate spiral. But it is evident from inspection of Fig. 21.7 that even a slight imprecision in determining the initial point leads to widely divergent positions in the past. Time is generally *not reversible* in stable dissipative systems in the sense that, unlike astronomers, we are not able to determine the past of the system (unless we know the current situation with absolute precision).

Since most systems found in economic modelling are dissipative rather than conservative, and absolute precision on the current situation is far from what economists can achieve, time irreversibility seems to be a connatural property of economic systems.

24.4.3 Goodwin's growth cycle

Goodwin is one of the first and most strenuous advocates of the need for non-linear analysis in economics, and has built several non-linear cyclical and growth-cyclical models (a collection is contained in Goodwin, 1982). An interesting application of the Lotka-Volterra equations is Goodwin's growth cycle (Goodwin, 1965, 1967), which has given rise to a copious offspring (for these developments see Goodwin et al. eds., 1984, Part I; Bródy and Farkas, 1987; Gabisch and Lorenz, 1989, Sect. 4.3.3; Thygesen et al. eds., 1991, Part I; Goodwin, 1991).

The interest derives from the fact that it is a model of cycles in growth rates and hence represents a step towards more realistic interpretations of cycles and growth. In fact, after the Second World War in most Western countries there has been an almost continual growth in real national income, which has, however, occurred at different rates over time. Thus the cycle is actually a cycle in growth rates, not in absolute levels of national income, the latter being usually on the increase or at least not on the decrease. It should however be remembered that, as noted in the previous section, the Lotka-Volterra equations are not structurally stable. Hence it is no surprise that even slight modifications in Goodwin's original model lead to different results, as shown in the literature cited above.

24.4.3.1 The model

The assumptions of the model are the following:

(1) steady technical progress, of the disembodied type. If we call a the labour productivity, that is, output per unit of labour, we can then write

$$\frac{Y}{L} = a = a_0 e^{\alpha t}, \tag{24.60}$$

where α is a positive constant.

(2) Steady growth in the labour force N:

$$N = N_0 e^{\beta t}, \tag{24.61}$$

where β is a positive constant. Note that the labour force, that is, the labour supply N, and the employment L in this model do *not* coincide, i.e. there is no full employment assumption.

(3) There are only two factors of production, labour and 'capital', both homogeneous and non-specific.

(4) All quantities are real and net.

(5) All wages are consumed; all profits are saved and automatically invested.

(6) The capital/output ratio $k = \frac{K}{Y}$ is constant.

(7) The real wage rate w rises in the neighbourhood of full employment.

Assumptions (1)-(5), as Goodwin says, are made for convenience, whereas the last two are of a more empirical, and disputable, sort.

Given the definition of a, we can write the workers' share of the product (call u such a share) as

$$u = \frac{wL}{Y} = \frac{w}{a}. \tag{24.62}$$

Consequently, the capitalists' share is[13]

$$1 - \frac{w}{a}. \tag{24.63}$$

It follows that the rate of profit is

$$\frac{\left(1 - \frac{w}{a}\right) Y}{K}, \tag{24.64}$$

that is, given assumption (4), (5) and (6),

$$\frac{\left(1 - \frac{w}{a}\right) Y}{K} = \frac{\left(1 - \frac{w}{a}\right)}{k} = \frac{\dot{K}}{K} = \frac{\dot{Y}}{Y}, \tag{24.65}$$

where a dot denotes the time derivative d/dt.

Consider now Eq. (24.60), take the logarithms to the base e and differentiate with respect to time; the result is

$$\frac{\dot{Y}}{Y} - \frac{\dot{L}}{L} = \alpha, \tag{24.66}$$

so that

$$\frac{\dot{L}}{L} = \frac{\dot{Y}}{Y} - \alpha, \tag{24.67}$$

is, given (24.65),

$$\frac{\dot{L}}{L} = \frac{\left(1 - \frac{w}{a}\right)}{k} - \alpha. \tag{24.68}$$

Define now the employment ratio v as

$$v = \frac{L}{N}. \tag{24.69}$$

Logarithmic differentiation yields

$$\frac{\dot{v}}{v} = \frac{\dot{L}}{L} - \frac{\dot{N}}{N}, \tag{24.70}$$

[13] There is a misprint in Goodwin's (1967) paper, where the capitalists' share is indicates as $l - w/a$ whereas, of course, it is $1 - w/a$. The misprint has, however, no effect on the final equations, which are correct.

that is, given (24.68) and since $\dot{N}/N = \beta$,

$$\frac{\dot{v}}{v} = \frac{\left(1 - \frac{w}{a}\right)}{k} - (\alpha + \beta) = \frac{1 - u}{k} - (\alpha + \beta). \tag{24.71}$$

Let us now consider assumption (7). It can be written as $\dot{w}/w = f(v)$,

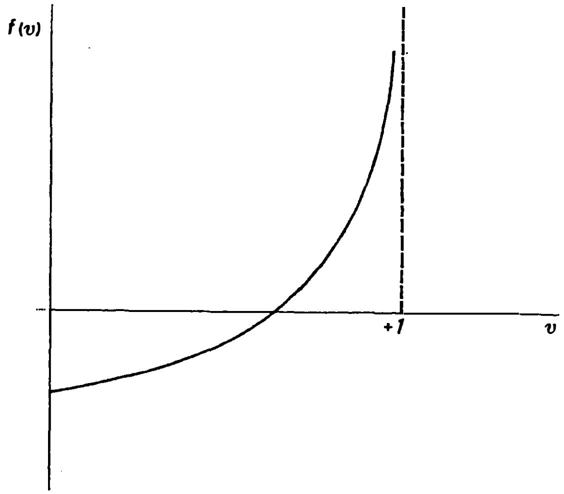


Figure 24.7: The real wage rate and employment

where f is an increasing function of the kind drawn in Fig. 24.7. Taking a linear approximation, we can write

$$\frac{\dot{w}}{w} = -\gamma + \rho v, \tag{24.72}$$

where γ and ρ are positive constants. Logarithmic differentiation of (24.62) yields

$$\frac{\dot{u}}{u} = \frac{\dot{w}}{w} - \alpha, \tag{24.73}$$

whence, given (24.72),

$$\frac{\dot{u}}{u} = -(\alpha + \gamma) + \rho v. \tag{24.74}$$

From Eqs. (24.71) and (24.74) we obtain the fundamental dynamic equations of the model:

$$\dot{v} = \left\{\left[\frac{1}{k} - (\alpha + \beta)\right] - \frac{1}{k}u\right\} v, \tag{24.75}$$

$$\dot{u} = [-(\alpha + \gamma) + \rho v]\, u. \tag{24.76}$$

Letting

$$\frac{1}{k} - (\alpha + \beta) = a_1, \quad \frac{1}{k} = b_1, \quad (\alpha + \gamma) = a_2, \quad \rho = b_2,$$

we have the Lotka-Volterra equations[14]

$$\begin{aligned} \dot{v} &= (a_1 - b_1 u)\, v \\ \dot{u} &= -(a_2 - b_2 v)\, u. \end{aligned} \tag{24.77}$$

24.4.3.2 The phase diagram of the model

Applying the procedure expounded in Sect. 24.4.1, we can draw the integral curves using the relation

$$\phi(v) = B\psi(u),$$

where $\phi(v) = v^{-a_2} \exp(b_2 v), \psi(u) = u^{a_1} \exp(-b_1 u)$; the shapes of these two functions have already been discussed in Sect. 24.4. The final result is shown in the diagram (Fig. 24.8), which is constructed as explained in Sect. 24.4.1 (the only difference is that the variable appearing in the first equation is now measured on the vertical axis instead of on the horizontal axis, so that the direction of the movement along the integral curve is clockwise). The variables by their definition are restricted in the interval from zero to $+1$[15].

A point on the u-axis will give us the distribution of income: workers' share is the segment from the origin to the point; capitalists' share is the segment from the point to $+1$. The latter share multiplied by the constant $1/k$ gives the profit rate of growth in output and in the capital stock (see Eqs. (24.64) and (24.65) above). From v we obtain the rate of growth of the wage rate (see Eq. (24.72) above). As the representative point travels around the closed curve, u vibrates between u_1 and u_2, and v between v_1 and v_2. A rough approximation to the typical time paths of u and v is illustrated in Fig. 24.9. Thus we have a cycle in the employment rate and in the growth rate of income. Whether the descending phase of the cycle implies also a fall in absolute values or means only that the latter increase less rapidly depends on the severity of the cycle. The same is true for real wages.

The economic mechanism underlying the motion of the representative point is, in Goodwin's (1967) words, the following: 'When profit is greatest, $u = u_1$, employment is average, $v = a_2/b_2$, and the high growth rate pushes employment to its maximum v_2, which squeezes the profit rate to its average value a_1/b_1. The deceleration in growth lowers employment (relative) to its

[14] The parameters b_1, b_2, a_2 are by definition positive. It seems realistic to assume that $(1/k) > \alpha + \beta$ (e.g., 0.2 can be taken as a safe lower limit for $1/k$, and 0.12 as a safe upper limit for $\alpha + \beta$), so that also $a_1 > 0$.

[15] It must be noted that u may exceptionally be greater than 1, since wages (= consumption) may exceed total product if there is disinvestment.

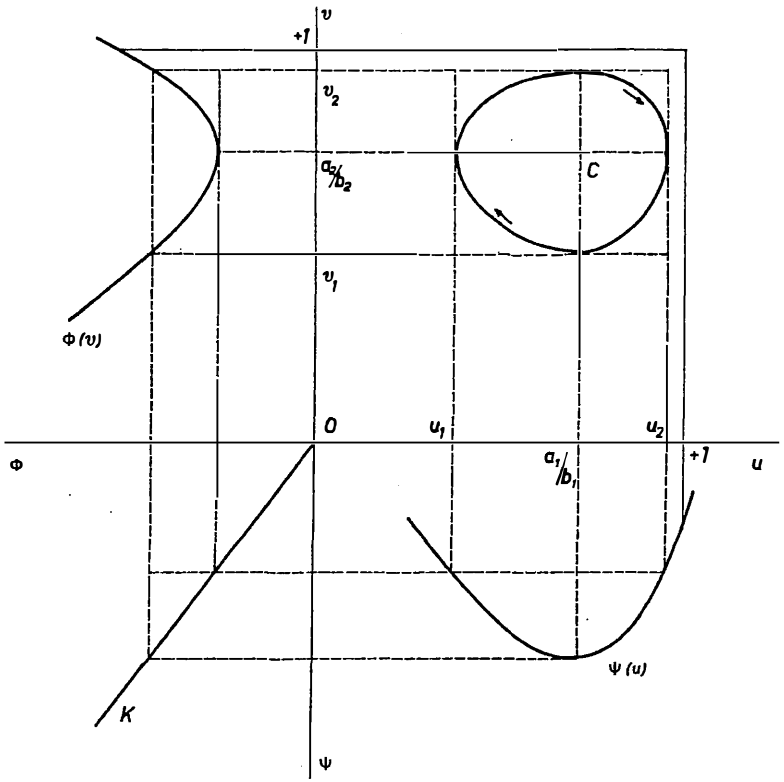


Figure 24.8: Goodwin's growth cycle

average value again, where profit and growth are again at their nadir u_2. This low growth rate leads to a fall in output and employment to well below full employment, thus restoring profitability to its average value because productivity is now rising faster than wage rates... .The improved profitability carries the seed of its own destruction by engendering a too vigorous expansion of output and employment, thus destroying the reserve army of labour and strengthening labour's bargaining power' (pp. 57-8; symbols have been changed in accordance with the notation used here). According to Goodwin, this is essentially Marx's idea of the 'contradictions' of capitalism; there is, however, a difference, since in the model the wage rate need not fall in absolute value. As we said above, in the descending phase the rate of growth of the wage rate falls, and whether or not this implies a fall in absolute value depends on the severity of the cycle.

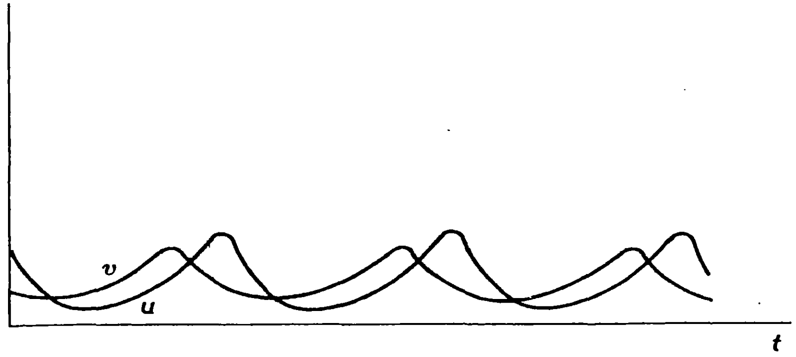


Figure 24.9: Path of u and v in Goodwin's growth cycle

As we know, external shocks will not alter the features of the (u, v) cycle, since they merely shift the representative point onto another integral curve, having the same shape and enclosing (or enclosed in) the previous one. In any case, that is for both undisturbed and disturbed systems, the very long-run average values of u and v, which can be taken as the coordinates of the singular point C (a centre) are independent of initial conditions and external shocks. This interesting property, just stated by Goodwin, can be proved as follows. Rewrite Eqs. (24.77) as

$$\begin{aligned} \frac{\mathrm{d}\log v}{\mathrm{d}t} &= a_1 - b_1 u, \\ \frac{\mathrm{d}\log u}{\mathrm{d}t} &= -a_2 + b_2 v, \end{aligned} \tag{24.78}$$

and integrate between two arbitrary values of t, say t' and t'', corresponding to which v and u take on, respectively, the values v', v'' and u', u''. The result is

$$\begin{aligned} \log\frac{v''}{v'} &= a_1(t''-t') - b_1 \int_{t'}^{t''} u \,\mathrm{d}t, \\ \log\frac{u''}{u'} &= -a_2(t''-t') + b_2 \int_{t'}^{t''} v \,\mathrm{d}t. \end{aligned} \tag{24.79}$$

Now call T the period of the oscillation, and let the limits of integration differ by T. It follows that $v' = v''$, $u' = u''$, $t'' - t' = T$, and so

$$0 = a_1 T - b_1 \int_0^T u \,\mathrm{d}t,$$

$$0 = -a_2 T + b_2 \int_0^T v \, dt,$$

that is

$$\frac{1}{T}\int_0^T u \, dt = \frac{a_1}{b_1}, \qquad (24.80)$$
$$\frac{1}{T}\int_0^T v \, dt = \frac{a_2}{b_2}.$$

Therefore the average values of u and v over a whole cycle are constant and respectively equal to $\frac{a_1}{b_1}$ and $\frac{a_2}{b_2}$, the coordinates of the centre of the fluctuation.

24.5 Exercises

1. Consider the neoclassical growth model (see Chap. 13, Sect. 13.2) and suppose that the production function is $Y = (aK^p + L^p)^{1/p}, 0 < p < 1, a > 0$, a case considered by Solow (1956). Take $p = 1/2$, and

(1.a) Show that the fundamental dynamic equation can be written as $r' = s(1 + A\sqrt{r})(1 + B\sqrt{r})$, where $A \equiv a - \sqrt{n/s}$, $B \equiv a + \sqrt{n/s}$.

(1.b) Show that this equation has separated variables, hence integrate it (Hint: before integrating multiply through by $(n/s)^{1/2}$).

2. Examine the phase diagram of the differential equation of the previous exercise in the initial form $r' = s[(a^2 - n/s)r + 2a\sqrt{r} + 1]$, and show that a necessary and sufficient condition for both the existence and the stability of equilibrium is $(a^2 - n/s) < 0$.

3. Consider the following multiplier-accelerator model by Goodwin (1951):

$$\begin{aligned} K^* &= kY, \\ C &= bY + a, \\ Y &= C + K', \end{aligned} \quad \text{and } K' = \begin{cases} K_1 & \text{if } K^* - K > 0 \\ 0 & \text{if } K^* - K = 0 \\ K_2 & \text{if } K^* - K < 0 \end{cases}$$

where K^* is the desired capital stock and the other symbols have the usual meaning. The assumption on investment behaviour is based on a non-linear form of the capital stock adjustment principle (that Goodwin calls the *flexible* or *non-linear accelerator*) and states that when $K^* \neq K$, entrepreneurs try to fill the gap as rapidly as possible. Therefore when $K^* > K$ investment is

carried out at the maximum rate allowed by the existing productive capacity, which in the short run is given (K_1). When $K^* < K$ disinvestment occurs at the maximum rate allowed by not replacing capital goods which are being scrapped for normal depreciation (K_2). Explain how to build the phase

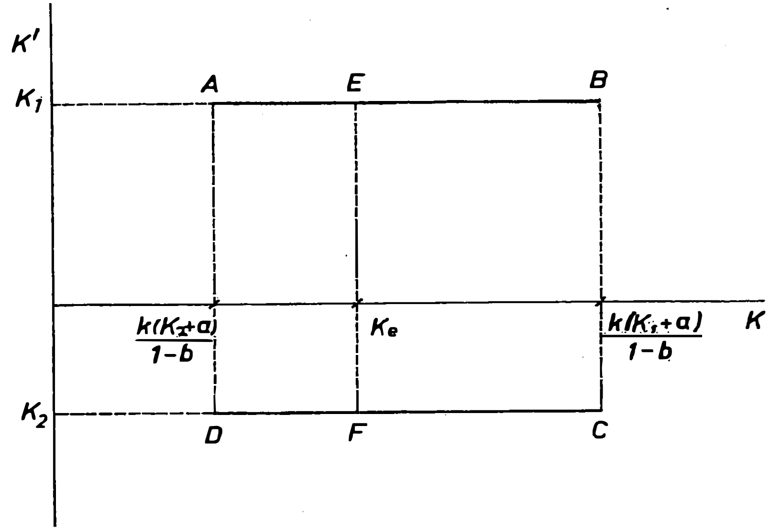


Figure 24.10: Goodwin's elementary non-linear cyclical model

diagram given in Fig. 24.10, and show that we have a continuous movement in the segments $\overline{AB}$ and $\overline{CD}$, with a discontinuous jump from B to C and from D to A.

This rather simple and admittedly crude model is sufficient to illustrate the basic features of non-linear cyclical models:

(i) the final result is independent of initial conditions (this is not true in linear models);

(ii) the oscillation is self-sustaining, without any need of exogenous factors nor of implausible assumptions about the coefficients (the latter, as we know, are necessary in linear models to obtain a constant-amplitude oscillations);

(iii) the equilibrium point is unstable, and therefore the slightest disturbance sets the mechanism into motion. However, notwithstanding such instability, the theory works, because the mechanism does not explode but is maintained within the limits of a constant-amplitude oscillation (this cannot occur in explosive linear models) thanks to its intrinsic non-linearities.

4. Consider a single market in disequilibrium. Then forces will operate on both price and quantity simultaneously. Assume that prices adjust according to the Walrasian hypothesis while quantities adjust according to the Marshallian hypothesis (see Chap. 13, Sect. 13.1). Thus we have the 2×2 system

$$p' = c[D(p) - q],$$

$$q' = k[p - p_s(q)],$$

where q is the quantity supplied and $p_s(q)$ is the supply price as defined in Sect. 13.1. Assume (Beckmann and Ryder, 1969) that $D(p) = (-1 + \varepsilon)p$, $p_s(q) = \delta q^3 - q$, where $0 < \varepsilon < 1, \delta > 0$, and that $c = k = 1$. Then

(4.a) show that the unique equilibrium point, which is the origin, is locally unstable (an unstable focus).

(4.b) apply the Poincaré-Bendixson theorem and show that the system has a stable limit cycle (Hint: as C_2 take the circle $p^2 + q^2 = \rho^2$, $\rho^2 > \rho_c \equiv \frac{1}{4}\{(5 - 4\varepsilon)/[\delta(1 - \varepsilon)]\}$. Then apply Liapunov's second method—see Chap. 23—and by taking $V = \frac{1}{2}(p^2 + q^2)$ show that $\mathrm{d}V/\mathrm{d}t < 0$ for $p^2 + q^2 > \rho_c$. Hence all trajectories coming from outside this circular region penetrate in it for t increasing, etc.).

(4.c) alternatively, show the existence of a stable limit cycle by using Nemitzky's theorem (hint: take any $\rho_1 < \frac{1}{2}\rho_c$ and any $\rho_2 > \rho_c$, where ρ_c is as defined in 4.b).

5. Calculate the linear approximation to the Lotka-Volterra equations and show that the singular point (other than the origin) is a centre.

24.6 References

Andronov, A.A., A.A. Vitt and S.E. Khaikin, 1966, *Theory of Oscillators*, Chaps. I-V.

Barnett, W.A., M. Salmon and A. Kirman (eds.), 1995, *Nonlinear Dynamics in Economics*.

Beckmann, M.J. and H.E. Ryder, 1969, Simultaneous Price and Quantity Adjustment in a Single Market.

Beltrami, E., 1987, *Mathematics for Dynamic Modeling*, Chap. 6.

Boole, G., 1960 (1872), *A Treatise on the Calculus of Finite Differences*, Chaps. IX, X, XII, XIII.

Bródy, A. and M. Farkas, 1987, Forms of Economic Motion.

Chang, W.W. and D.J. Smyth, 1971, The Existence and Persistence of Cycles in a Non-linear Model: Kaldor's 1940 Model Re-examined.

Chiarella, C., 1990, *The Elements of a Nonlinear Theory of Economic Dynamics*.

Dore, M.H.I., 1993, *The Macrodynamics of Business Cycles*.

Forsyth, A.R., 1890-1906, *Theory of Differential Equations*.

Franke, R. and T. Lux, 1993, Adaptive Expectations in a Nonlinear Metzlerian Model of the Inventory Cycle.

Gabisch, G. and H.-W. Lorenz, 1989, *Business Cycle Theory: A Survey of Methods and Concepts*.

Galeotti M. and F. Gori, 1989, Uniqueness of Periodic Orbits in Liénard-type Business Cycle Models.

Georgescu-Roegen, N., 1951, Relaxation Phenomena in Linear Dynamic Models.

Goodwin, R.M., 1951, The Non-linear Accelerator and the Persistence of Business Cycles.

Goodwin, R.M., 1967(1965), A Growth Cycle.

Goodwin, R.M., 1982, *Essays in Economic Dynamics*.

Goodwin, R.M., 1991, Non-linear Dynamics and Evolution.

Goodwin, R.M., M. Krüger and A. Vercelli (eds.), 1984, *Nonlinear Models of Fluctuating Growth*.

Grasman, J. and J.J. Wentzel, 1994, Co-existence of a Limit Cycle and an Equilibrium in Kaldor's Business Cycle Model and Its Consequences.

Hamburger, L., 1931, Analogie des fluctuations économiques et des oscillations de relaxation.

Hamburger, L., 1934, Note on Economic Cycles and Relaxation Oscillations.

Hirsch M.W. and S. Smale, 1974, *Differential Equations, Dynamical Systems, and Linear Algebra*, Chaps. 11-13.

Hsu, S.-B. and T.-W. Huang, 1995, Global Stability for a Class of Predator-Prey Systems.

Ichimura, S., 1955, Toward a General Nonlinear Macrodynamic Theory of Economic Fluctuations.

Ince, E.L., 1956 (1926), *Ordinary Differential Equations*.

Kaldor, N., 1940, A Model of the Trade Cycle.

Kaldor, N., 1971, A Comment.

Kosobud, R.F. and W.D. O'Neil, 1972, Stochastic Implications of Orbital Asymptotic Stability of a Nonlinear Trade Cycle Model.

Le Corbeiller, P., 1933, Les systèmes autoentretenus et les oscillations de relaxation.

Levinson, N. and O.K. Smith, 1942, A General Equation for Relaxation Oscillations.

Lorenz, H.-W., 1993, *Nonlinear Dynamical Economics and Chaotic Motion*, 2nd ed., Chap. 1.

Lotka, A.J., 1956 (1926), *Elements of Mathematical Biology*, pp. 88-92.

Marrama, V., 1946, Short Notes on a Model of the Trade Cycle.

Milne-Thomson, L.M., 1933, *The Calculus of Finite Differences*.

Minorsky, N., 1962, *Nonlinear Oscillations*, Chaps 1-4.

Sansone, G. and R. Conti, 1964, *Nonlinear Differential Equations*, Chaps I, II, IV, VI.

Solow, R.M., 1956, A Contribution to the Theory of Economic Growth.
Sordi, S., 1988 (1990), *Theories of the Business Cycle.*
Thygesen, N., K. Velupillai and S. Zambelli (eds.), 1991, *Business Cycles: Theories, Evidence and Analysis.*
Volterra, V., 1927 Variazioni e fluttuazioni del numero d'individui in specie animali conviventi.
Volterra, V., 1931, *Leçons sur la théorie mathématique de la lutte pour la vie.*
Zarnowitz, V., 1992, *Business Cycles: Theory, History, Indicators, and Forecasting.*

Chapter 25

Bifurcation Theory

25.1 Introduction

We have already met bifurcations in Chap. 21, Sect. 21.3.2.2. When a dynamic system involves a parameter, it may happen that the passage of the parameter through a critical value causes a qualitative change in the nature of the singular point(s) and of the trajectories. The value(s) of the parameter at which such a change occurs are called *bifurcation* or *branch* value(s) (Andronov et al., 1966, Chap. II, § 5; Minorsky, 1962, Chap. 7; a more recent treatment is Guckenheimer and Holmes, 1986, Chap. 3).

Some types of bifurcations (for example the Hopf bifurcation) involve a situation in which a system in (stable) equilibrium can lose its stability, giving rise to a limit cycle, as the value of a parameter is changed. This situation of loss of stability and the corresponding birth of a limit cycle is often the first step in a route to 'chaotic' motion (see the next chapter). Bifurcation theory has of course a great interest in itself, quite independently of its possible relation with chaos theory.

In the following sections we shall deal with *local* bifurcation theory; the global bifurcation behaviour of a system is best dealt with in the context of the subsequent treatment of chaotic dynamics.

25.2 Bifurcations in continuous time systems

The basic local bifurcation behaviour of a system can be simply illustrated by building on material treated at length in previous chapters. Consider the system

$$\mathbf{y}' = \varphi(\mathbf{y}, \alpha), \tag{25.1}$$

where α is a parameter. We know (Chap. 21, Sect. 21.2.1) that the null solution or equilibrium point $\mathbf{y}_e$ of this system is given by solving

$$\varphi(\mathbf{y}_e, \alpha) = \mathbf{0}. \tag{25.2}$$

Now, as α varies, the implicit function theorem (see Chap. 20, Sect. 20.2) implies that these equilibria are described by continuously differentiable functions of α *away* from those points at which the Jacobian determinant of (25.2) is zero:

$$\mathbf{y}_e = \mathbf{y}_e(\alpha). \tag{25.3}$$

The graph of each of these functions is called a *branch* of equilibria of Eq. (25.1).

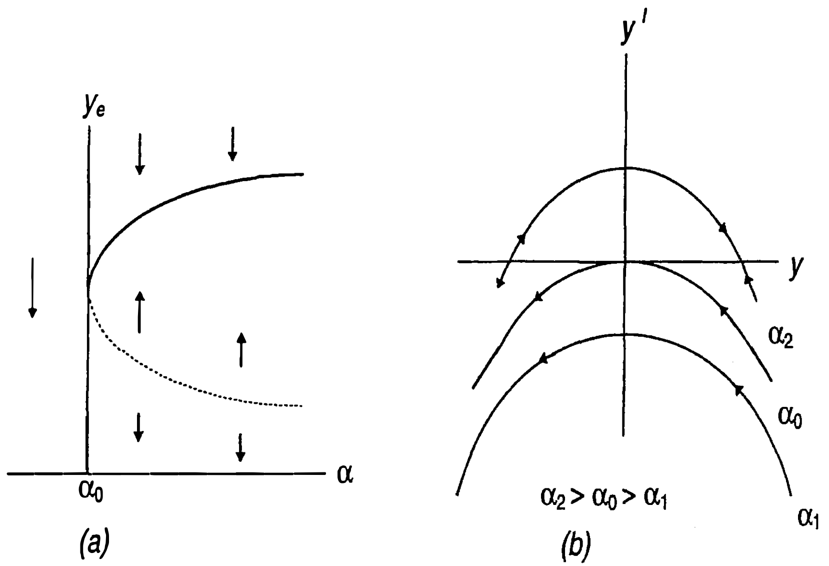


Figure 25.1: Saddle-node (or fold) bifurcation, $y' = \alpha - y^2$

We know that the Jacobian matrix of (25.1) is the matrix of the linear approximation to study local stability (Chap. 21, Sect. 21.4.1). A non-zero Jacobian means that there are no zero latent roots in this linear approximation (Chap. 18, Sect. 18.2.2). If, at an equilibrium point $(\mathbf{y}_{e,0}, \alpha_0)$ the Jacobian is zero and several branches of equilibria come together, one says that $(\mathbf{y}_{e,0}, \alpha_0)$ is a point of *bifurcation.* An equilibrium point at which no bifurcation occurs is called a *hyperbolic* fixed point.

We know that the local behaviour of the system depends on the latent roots of the Jacobian. Suppose, for example, that

(i) for a certain value of α, say α_0, the Jacobian is zero (there is a zero latent root), while it remains different from zero for $\alpha \neq \alpha_0$;

(ii) all the latent roots of the Jacobian are of a certain type for, say, $\alpha < \alpha_0$, while there is at least one latent root of a different type for $\alpha > \alpha_0$;

then a qualitative change occurs in the behaviour of the system as α passes through α_0 (which is called the *bifurcation value* of the parameter), since the system has a different phase portrait for $\alpha < \alpha_0$ and for $\alpha > \alpha_0$.

It should be stressed that a qualitative change is not only a change from stability to instability (or vice versa), which happens when for, say, $\alpha < \alpha_0$ all the latent roots have negative real parts while there is at least one root with positive real part for $\alpha > \alpha_0$.

The passage from stability to instability or vice versa is certainly the most striking qualitative change, but there may be a qualitative change in the behaviour of the system also without such a dramatic change in the stability properties. If, for example, the system remains stable but some real negative roots become complex with negative real part as α passes through α_0, we certainly have a change in the qualitative behaviour of the system, that from monotonic becomes oscillatory. In the case of a planar system, this would mean the passage from a stable node to a stable focus in the phase diagram (see Chap. 21, Table 21.1). In the case of a one-dimensional system, namely of a single non-linear differential equation, the qualitative change is only of the stability-instability type, since the Jacobian reduces to the partial derivative $\partial\varphi(y, \alpha)/\partial y$, which has just one (real) root.

It is clear that bifurcations are closely related to *structural stability* (see Chap. 21, Sect. 21.2.3). Usually, in fact, a system possessing bifurcations is not structurally stable, and this leads to an alternative definition of bifurcation, namely that a value α_0 of Eq. (25.1) for which the solution of (25.1) is not structurally stable is a bifurcation value of α (Guckenheimer and Holmes, 1986, p. 119).

25.2.1 Codimension-one bifurcations

In the case of one-dimensional systems bifurcations can be examined by means of a so-called *bifurcation diagram* (Andronov at al., 1966, Chap. III, § 5; Minorsky, 1962, Chap. 7, Sect. 8; Guckenheimer and Holmes, 1986, Chap. 3, Sect. 3.1), namely by a diagram in which the branches of equilibria are shown in (y_e, α) space. In Figs. 25.1-25.3 we illustrate three basic bifurcation points, namely the *saddle-node* or *fold* bifurcation, the *transcritical* bifurcation, and the *pitchfork* bifurcation. The fourth basic bifurcation, the *Hopf bifurcation*, will be treated in the next section since it involves at least a 2×2 system. Before commenting on the diagrams, let us give the conditions for the various bifurcations in one-dimensional systems, also called *codimension one bifurcations* (Guckenheimer and Holmes, 1986, Chap. 3, Sect. 3.4; Sotomayor, 1973).

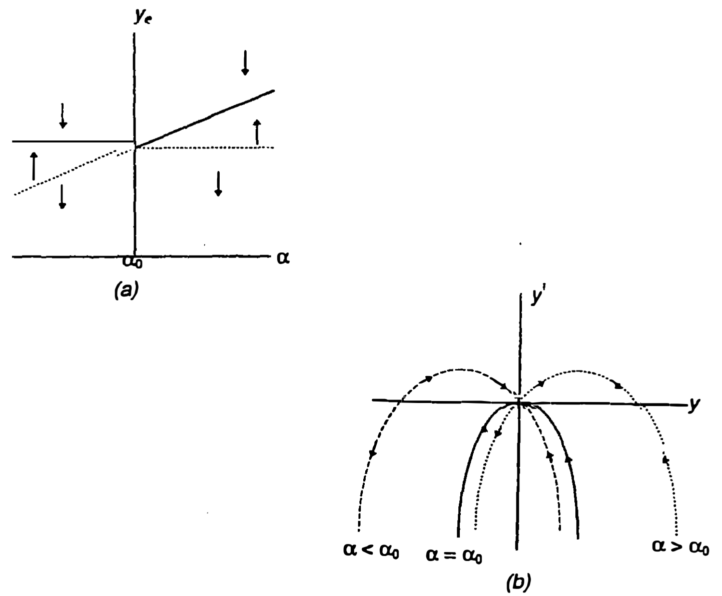


Figure 25.2: Transcritical bifurcation, $y' = \alpha y - y^2$

1. Saddle-node (or fold) bifurcation (prototype equation: $y' = \alpha - y^2$)

Consider the one-parameter differential equation

$$y' = f(y, \alpha)$$

and assume that when $\alpha = \alpha_0$ there is an equilibrium $y_{e,0}$ for which the following hypotheses are satisfied:

(1.a) $\partial f(y_{e,0}, \alpha_0)/\partial y = 0$,

(1.b) $\partial^2 f(y_{e,0}, \alpha_0)/\partial y^2 \neq 0$,

(1.c) $\partial f(y_{e,0}, \alpha_0)/\partial \alpha \neq 0$.

Then, depending on the signs of the expressions (1.b) and (1.c), there are

(i) no equilibria near $(y_{e,0}, \alpha_0)$ when $\alpha < \alpha_0$ ($\alpha > \alpha_0$);

(ii) two equilibria near $(y_{e,0}, \alpha_0)$ for each parameter value $\alpha > \alpha_0$ ($\alpha < \alpha_0$). These equilibria are hyperbolic; one of them is stable and the other unstable.

Conditions (1.b) and (1.c) are called transversality conditions in bifurcation theory. Figure 25.1a illustrates the bifurcation diagram for the case in

which (1.b) is negative and (1.c) positive. When the parameter α is lower than the bifurcation value α_0, no equilibria exist. When $\alpha > \alpha_0$, two branches of equilibria emerge, one being stable (continuous curve) and the other unstable (dashed curve). Other cases are possible with different signs of (1.b) and (1.c), but the nature of the diagram would be similar.

Figure 25.1b illustrates the phase diagram (see Chap. 21, Sect. 21.3.1) for the case under consideration. In addition to the critical phase curve for $\alpha = \alpha_0$, we have drawn two representative phase curves, one for $\alpha_1 < \alpha_0$ and the other for $\alpha_2 > \alpha_0$. It is easy to see that for $\alpha < \alpha_0$ the phase curve does not cross the y axis, hence there are no equilibria. On the contrary, when $\alpha > \alpha_0$, there are two equilibria, one stable and the other unstable.

The use of the phase diagram allows a simple derivation of the conditions (1.a)-(1.c). Consider the function $f(y, \alpha)$. Condition (1.a) means that the function has a stationary point with respect to y at $(y_{e,0}, \alpha_0)$; condition (1.b) means that this point is an extremum (a minimum if $\partial^2 f(y_{e,0}, \alpha_0)/\partial y^2 > 0$, a maximum if $\partial^2 f(y_{e,0}, \alpha_0)/\partial y^2 < 0$). Since this extremum point is an equilibrium point of the differential equation, namely $f(y_{e,0}, \alpha_0) = 0$, it must lie on the horizontal axis in the phase diagram, i.e. the function is tangent to this axis at the extremum. In the figure we have shown the case of a maximum.

Condition (1.c) means that the function is not stationary with respect to α at this point, hence the level curves in the (f, y) plane, that is in the phase diagram, will shift vertically as α changes, according to the sign of $\partial f(y_{e,0}, \alpha_0)/\partial \alpha$. If this sign is positive, an increase (decrease) in α will cause the phase curve to shift upwards (downwards), as in the figure. The other way round if $\partial f(y_{e,0}, \alpha_0)/\partial \alpha$ is negative.

2. Transcritical bifurcation (prototype equation: $y' = \alpha y - y^2$)

Consider the one-parameter differential equation

$$y' = f(y, \alpha),$$

and assume that when $\alpha = \alpha_0$ there is an equilibrium $y_{e,0}$ for which the following hypotheses are satisfied:

(2.a) $\partial f(y_{e,0}, \alpha_0)/\partial y = 0$,

(2.b) $\partial^2 f(y_{e,0}, \alpha_0)/\partial y^2 \neq 0$,

(2.c) $\partial^2 f(y_{e,0}, \alpha_0)/\partial \alpha \partial y \neq 0$.

Then, depending on the signs of the expressions (2.b) and (2.c),

(i) the equilibrium $y_{e,0}$ is stable (unstable) when $\alpha < \alpha_0$ ($\alpha > \alpha_0$);

(ii) the equilibrium $y_{e,0}$ becomes unstable (stable) for each parameter value $\alpha > \alpha_0$ ($\alpha < \alpha_0$), and a branch of additional stable (unstable) equilibria $y_e(\alpha)$ emerges.

Figure 25.2 shows this type of bifurcation. We have assumed that the transversality conditions (2.b) and (2.c) are respectively negative and positive.

The interpretation in terms of the phase diagram is, again, easier to understand. The function $f(y,\alpha)$ has an extremum, which we have supposed to be a maximum, at the equilibrium point. Condition (2.c), on the mixed second-order partial derivative of the function, means that a change in α shifts the phase curve vertically *and* horizontally, while letting it pass through $y_{e,0}$. If $\partial^2 f(y_{e,0},\alpha_0)/\partial\alpha\partial y$ is positive, as assumed in the diagram, an increase (decrease) in α will shift the phase curve upwards and to the right (left). Simple phase-diagram considerations show the nature of the various equilibrium points.

Thus the transcritical bifurcation is characterized by an *exchange of stability* at the initial equilibrium point.

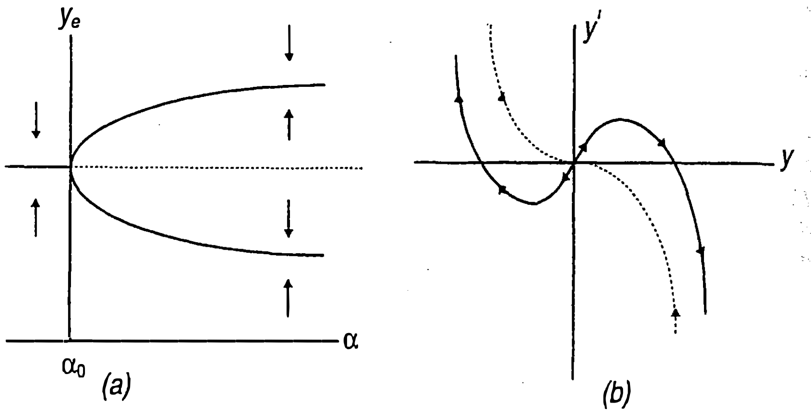


Figure 25.3: Pitchfork bifurcation, $y' = \alpha y - y^3$

We now pass to the pitchfork bifurcation, which occurs when f is an odd function with respect to y, namely $f(y,\alpha) = -f(-y,\alpha)$.

3. Pitchfork bifurcation (prototype equation: $y' = \alpha y - y^3$)
Consider the one-parameter differential equation

$$y' = f(y,\alpha),$$

and assume that when $\alpha = \alpha_0$ there is an equilibrium $y_{e,0}$ for which the following hypotheses are satisfied:

(3.a) $\partial f(y_{e,0},\alpha_0)/\partial y = 0$,
(3.b) $\partial^3 f(y_{e,0},\alpha_0)/\partial y^3 \neq 0$,
(3.c) $\partial^2 f(y_{e,0},\alpha_0)/\partial\alpha\partial y \neq 0$.

Then, depending on the signs of the expressions (3.b) and (3.c),

(i) the equilibrium $y_{e,0}$ is stable (unstable) when $\alpha < \alpha_0$ ($\alpha > \alpha_0$);

(ii) the equilibrium $y_{e,0}$ becomes unstable (stable) for each parameter value $\alpha > \alpha_0$ ($\alpha < \alpha_0$), and two branches of additional stable (unstable) equilibria $y_e(\alpha)$ emerge.

The pitchfork bifurcation is shown in Fig. 25.3. The fact that f is an odd function means that condition (b) in the previous forms will be violated for some y. This means that the function changes curvature. The third-order partial derivative being different from zero excludes the presence of a horizontal inflection point at $y_{e,0}$, a case in which the phase curve would never cross the y axis except at the origin. The effect of the mixed second-order partial derivative has the effect of shifting the phase curve as in the case of the transcritical bifurcation.

In the diagram we have shown the case in which (3.b) is negative and (3.c) is positive. In this case the pitchfork is called *supercritical,* and the two additional branches represent stable equilibria. In the case in which the two additional equilibria were unstable, the pitchfork would be called *subcritical.*

These three types of codimension one bifurcations can of course be generalized to higher dimensions. In this case, as we have already stated at the beginning of this section, the conditions on the single partial derivatives have to be replaced with conditions on the Jacobian and its latent roots; in particular, the condition $\partial f(y_{e,0},\alpha_0)/\partial y = 0$ is replaced by the condition that the Jacobian determinant is zero (see Guckenheimer and Holmes, 1986, Chap. 3, Sect. 3.4, and Sotomayor, 1973, for the general treatment).

25.2.2 The Hopf bifurcation

Unlike the three typical codimension one bifurcations, the Hopf bifurcation requires at least a 2×2 system to appear. It is a particularly interesting type of bifurcation because it shows the emergence of a limit cycle under certain conditions. For example, it shows how a system with a stable equilibrium point can lose its stability, giving rise to a (possibly) stable limit cycle, when the parameter on which the system depends is changed. Let us first consider the 2×2 system

$$\begin{aligned} y_1' &= \varphi_1(y_1,y_2,\alpha), \\ y_2' &= \varphi_2(y_1,y_2,\alpha), \end{aligned} \tag{25.4}$$

and assume that for each α in the relevant range this system has an isolated equilibrium point $\mathbf{y}_e = (y_{1e}, y_{2e})$ obtained by solving the system

$$\begin{aligned} \varphi_1(y_1, y_2, \alpha) &= 0, \\ \varphi_2(y_1, y_2, \alpha) &= 0. \end{aligned} \tag{25.5}$$

As already noted above (see Sect. 25.2), the solution of (25.5) will give y_{1e}, y_{2e} as continuously differentiable functions of the parameter, namely

$$\mathbf{y}_e = \mathbf{y}_e(\alpha), \tag{25.6}$$

if the Jacobian matrix of (25.5) is non singular at the equilibrium point. This Jacobian matrix, that coincides with the matrix of the linear approximation of system (25.6), is:

$$\mathbf{J}(\alpha) = \begin{bmatrix} \dfrac{\partial\varphi_1(y_{1e}, y_{2e}, \alpha)}{\partial y_1} & \dfrac{\partial\varphi_1(y_{1e}, y_{2e}, \alpha)}{\partial y_2} \\ \dfrac{\partial\varphi_2(y_{1e}, y_{2e}, \alpha)}{\partial y_1} & \dfrac{\partial\varphi_2(y_{1e}, y_{2e}, \alpha)}{\partial y_2} \end{bmatrix}, \tag{25.7}$$

and will in principle depend on α as well. Hence the latent roots of $\mathbf{J}(\alpha)$, on which the local behaviour of the system depends (since the system is autonomous, the conditions of the Poincaré-Liapunov-Perron theorem are satisfied: see Chap. 21, Sect. 21.4.1), are also continuously differentiable functions of α. Let us now suppose that both latent roots are complex—actually, a complex conjugate pair

$$\begin{aligned} &\lambda_{1,2} = \theta \pm \mathrm{i}\omega, \\ &\text{where} \\ &\theta = \theta(\alpha), \qquad \omega = \omega(\alpha). \end{aligned} \tag{25.8}$$

We know that when the real part is negative (positive) we have stability (instability), while when the real part is zero we cannot tell the local behaviour of the non-linear system from the linear approximation (see Chap. 21, Sect. 21.4.1). This is exactly where the Hopf bifurcation theorem comes in to help us. This theorem has an *existence part*, that shows the sufficient conditions for the emergence of a closed orbit, and a *stability* part showing the sufficient conditions for orbital stability of the cycle.

Unfortunately the stability part is exceedingly complicated. It involves the reduction of the system to its *centre manifold* before applying higher-order Taylor expansion at the equilibrium point and coordinate transformation so as to put the system in the *normal form*

$$\mathbf{y}' = \mathbf{N}\mathbf{y} + \mathbf{g}(\mathbf{y}), \tag{25.9}$$

where

$$\mathbf{N} \equiv \begin{bmatrix} 0 & -\omega \\ \omega & 0 \end{bmatrix},$$

and $\mathbf{g}(\mathbf{y})$ is a vector of non-linear terms. Then one has to calculate a certain expression b, that involves up to the third-order partial derivatives (direct and mixed) of the functions appearing in $\mathbf{g}(\mathbf{y})$. Finally, the emerging cycle is stable if $b < 0$, unstable if $b > 0$.

Apart from the mathematical difficulties (for details the reader can consult Guckenheimer and Holmes, 1986, Sects. 3.3 and 3.4), there is little hope to extricate an economic meaning from these stability conditions, especially insofar as third-order mixed partial derivatives have no clear economic interpretation.

Hence we just state the existence part (the following is a truncated version of the theorem given in Invernizzi and Medio, 1991, Appendix B).

Hopf bifurcation theorem (existence part)

Consider the general system (25.1) and suppose that for each α in the relevant interval it has an isolated equilibrium point $\mathbf{y}_e = \mathbf{y}_e(\alpha)$. Assume that the Jacobian matrix of φ with respect to $\mathbf{y}$, evaluated at $(\mathbf{y}_e(\alpha), \alpha)$ has the following properties:

(H1) it possesses a pair of simple complex conjugate eigenvalues $\theta(\alpha) \pm \mathrm{i}\omega(\alpha)$ that become pure imaginary at the critical value α_0 of the parameter—i.e., $\theta(\alpha_0) = 0$, while $\omega(\alpha_0) \neq 0$—and no other eigenvalues with zero real part exist at $(\mathbf{y}_e(\alpha_0), \alpha_0)$;

(H2) $\left.\dfrac{\mathrm{d}\theta(\alpha)}{\mathrm{d}\alpha}\right|_{\alpha=\alpha_0} \neq 0$;

THEN system (25.1) has a family of periodic solutions.

The critical value α_0 is called a *Hopf bifurcation point* of system (25.4). Note that by a suitable change of coordinates from $\mathbf{y}$ to $\mathbf{z} = \mathbf{y} - \mathbf{y}_e(\alpha)$ we can always make the critical value of α coincide with the origin, but this is not essential.

Since the existence part of the theorem leaves us in the dark as regards the nature of the cycle, there are in principle various possibilities. One possibility is that orbits spiral outward from $\mathbf{y}_e(\alpha)$ when $\alpha > \alpha_0$ toward a stable limit cycle (Fig. 25.4(i)). This is called a *supercritical Hopf bifurcation*. Another is that an unstable cycle exists for $\alpha < \alpha_0$, inside of which all orbits spiral in toward $\mathbf{y}_e(\alpha)$ (Fig. 25.4(ii)). This is called a *subcritical Hopf bifurcation*. Other possibilities may also exist.

In the 2×2 case, if we want to check the stability of the cycle (orbital stability), we can use the theorems on orbital stability given in the previous chapter, Sect. 24.3, where we have also given the basic theorems for the

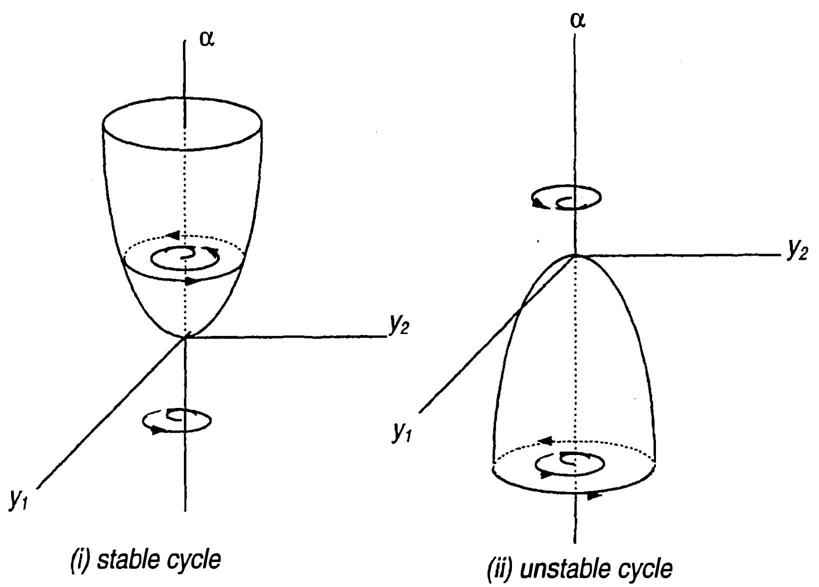


Figure 25.4: The Hopf bifurcation

existence of periodic motions. So—apart from pure mathematical interest—why do we need the Hopf bifurcation theorem?

The huge value-added of the Hopf bifurcation theorem over the standard planar theory of oscillations (see above, Sect. 24.3) lies in the fact that this theorem can be applied to higher-dimensional system, while the standard theory only holds for 2×2 systems.

However, also the application of the existence part of the Hopf bifurcation theorem often becomes analytically intractable for systems of dimension higher than the third. In the case of a three-dimensional system, the Jacobian is a 3×3 matrix, and we shall have to cope with a third-order polynomial of the type

$$\lambda^3 + a_1\lambda^2 + a_2\lambda + a_3 = 0, \tag{25.10}$$

where the coefficients a_i are related to the elements of the matrix by well-known relations (see Chap. 18, Sect. 18.2.1). Elementary theory of equations (Turnbull, 1957, Chap. IX, Sect. 52) tells us that, in general, the condition for this cubic equation to have one real root and a pair of complex conjugate roots is

$$\begin{aligned} &b^2/4 + a^3/27 > 0, \\ &\text{where} \\ &a \equiv a_2 - \tfrac{1}{3}a_1^2, \qquad b \equiv \tfrac{2}{27}a_1^3 - \tfrac{1}{3}a_1a_2 + a_3. \end{aligned} \tag{25.11}$$

Recalling that the coefficients a_i are functions of the elements of the Jacobian, condition (25.11) is difficult to check. Besides, this does not tell us anything on the nature of the roots. Since what we want is a *negative* real root and a pair of *pure imaginary* roots, we can apply the conditions shown in Chap. 16, Sect. 16.4 (page 221). Thus we get

$$\begin{aligned} &a_1 > 0, \quad a_2 > 0, \quad a_3 > 0, \\ &a_1a_2 - a_3 = 0, \end{aligned} \tag{25.12}$$

as the conditions for equation (25.10) to have one negative real root and a pair of complex roots with zero real part at α_0.

Conditions (25.12) are simpler than condition (25.11), but by no means easy to check considering again that the coefficients a_i are functions of the elements of the Jacobian. For fourth- and higher-order equations, the problem becomes practically intractable from the algebraic point of view except in particular cases.

Let us finally note that the Hopf bifurcation theorem has a *local* validity only, hence it cannot say anything on the global behaviour of the system. As regards this, it may happen that, after the system has undergone a Hopf bifurcation and is on the newly arising closed orbit, a new Hopf bifurcation arises (a case that will be used in the next chapter, see p. 525). A Hopf bifurcation for periodic orbits is called a *secondary Hopf bifurcation* or Neimark bifurcation. Conditions for this to occur are given in Medio (1992, Chap. 2, Sect. 2.7.2).

25.2.3 Sensitivity analysis and bifurcations: a reminder

As we have seen, bifurcation analysis involves the study of the behaviour of the latent roots of the Jacobian of the system when the parameter varies, and in particular of the derivative of the relevant latent roots with respect to the parameter, which is required to check (H.2) in the Hopf bifurcation theorem. In general a system may have several parameters, each of which has to be studied as a possible source of bifurcations.

We have already shown in the linear parts of this book how to compute the partial derivatives of the m-th latent root with respect of the s-th parameter in a first-order $n \times n$ linear system with n latent roots and p parameters. Hence we refer the reader to the treatment of sensitivity analysis given in Chap 18, Sect. 18.2.2.2.

25.2.4 Kaldor's non-linear cyclical model again

We shall apply the Hopf bifurcation theorem given above. Let us recall Kaldor's model from the previous chapter (Sect. 24.3.3.1):

$$\begin{aligned} Y' &= \alpha[I(Y,K) - S(Y,K)], \quad I_Y > S_Y > 0, I_K < 0, S_K > 0, \\ K' &= I(Y,K). \end{aligned} \tag{25.13}$$

To apply the Hopf bifurcation theorem we must parametrize the model. Although there may be various ways of doing this, the most obvious is to use the already present parameter, namely the adjustment speed α. Let us now consider the Jacobian matrix

$$\begin{bmatrix} \alpha(I_Y - S_Y) & \alpha(I_K - S_K) \\ I_Y & I_K \end{bmatrix}, \tag{25.14}$$

whose characteristic equation is

$$\begin{gathered} \lambda^2 + a_1\lambda + a_2 = 0, \\ \text{where} \\ a_1 \equiv -\left[\alpha(I_Y - S_Y) + I_K\right], \\ a_2 \equiv \alpha\left[(I_Y - S_Y)I_K - I_Y(I_K - S_K)\right]. \end{gathered} \tag{25.15}$$

A necessary condition for complex roots is $a_2 > 0$, which is certainly satisfied given the assumptions on the partial derivatives. As regards a_1, we see that

$$a_1 \gtreqless 0 \text{ according as } \alpha \lesseqgtr \alpha_0, \quad \alpha_0 \equiv \frac{I_K}{S_Y - I_Y} > 0\,. \tag{25.16}$$

Let us then consider the roots of the Jacobian,

$$\lambda_{1,2} = -\frac{1}{2}a_1 \pm \frac{1}{2}\sqrt{a_1^2 - 4a_2}. \tag{25.17}$$

Since $a_1 = 0$ for the critical value α_0 of the parameter, the Jacobian has a pair of pure imaginary roots there, $\lambda_{1,2} = \pm i\sqrt{a_2}$. The roots of (25.17) will remain complex conjugate for a_1 sufficiently small, namely for α sufficiently near to α_0. This satisfies assumption (H.1).

It is also easy to check that

$$\left.\frac{\mathrm{d}\left(-\frac{1}{2}a_1\right)}{\mathrm{d}\alpha}\right|_{\alpha=\alpha_0} = \frac{1}{2}(I_Y - S_Y) > 0, \tag{25.18}$$

which satisfies assumption (H.2). The existence of a closed orbit in the Kaldor model is proved.

Calculations to check the stability of this cycle via the determination of the normal form (25.9) and of the consequent expression b have been carried out by H.-W. Lorenz (1993, pp. 103-105), with the result that it is not possible to determine the sign of b on the basis of purely qualitative information.

25.2.5 Oscillations in optimal growth models

25.2.5.1 The model

We have seen that optimal growth models usually have the saddle-point property (Chap. 22, Sects. 22.2.1.3 and 22.2.2.3). This, however, is not the only kind of movement that can arise. Benhabib and Nishimura (1979), in fact, by using the Hopf bifurcation showed that in a multi-sector neoclassical optimal growth model solutions may exist in which the optimal path is a closed curve around the equilibrium point rather than a monotonic movement. Further study of the problem (always in the context of the Hopf bifurcation) and references are contained in Medio (1987), Dockner and Feichtinger (1991), and Cartigny and Venditti (1994).

The fact that a multi-sector neoclassical optimal growth model may not be (saddle-point) stable should come as no surprise, for the possible instability of multi-sector neoclassical *descriptive* growth models has been known for a long time under the name of 'Hahn problem' (Hahn, 1966; Gandolfo, 1971, App. III, § 7.6).

However, the possible emergence of *cycles* in *optimal* growth models has more far-reaching implications, since it amounts to saying that, under certain circumstances, 'cycles are optimal'. Fluctuations, that are usually considered as non desirable (hence non optimal), emerge as the result of an optimization process. It follows that in policy models, i.e., in models in which the optimizing agent is the government, the social welfare function should also take account of oscillations with a *negative* weight. But this lies outside the scope of the present book, hence let us turn back to our problem.

To illustrate the problem we shall build a simple model with one consumption good and two *heterogeneous* capital goods, where each good is produced (under no joint production) by using labour and the two capital goods with constant returns to scale technologies which are 'well behaved' (i.e., satisfy the Inada conditions–see Chap. 13, Sect. 13.2). It is further assumed that each factor of production is directly required in the production of all commodities. The properties of this model are well known in neoclassical descriptive growth theory, hence we shall just summarize them (for proofs the reader is referred to Burmeister and Dobell, 1970, Chap. 9).

The available technology can be summarized by the *production possibility frontier*

$$c = T(k_1, k_2, y_1, y_2), \tag{25.19}$$

where c is per capita consumption, k_1 per capita stock of the first capital good, k_2 per capita stock of the second capital good, and y_1, y_2 the current per capita outputs of the two capital goods. The production possibility frontier involves the solution to a previous optimal resource allocation problem, hence it tells us the (maximum possible) consumption at time t when the factors of production are optimally allocated to produce the three goods. Other things being equal, an increase in any input increases consumption, while an increase in the output of either capital good decreases consumption, i.e.

$$\partial T/\partial k_1 > 0,\ \partial T/\partial k_2 > 0,\ \partial T/\partial y_1 < 0,\ \partial T/\partial y_2 < 0. \tag{25.20}$$

Let us now recall for future reference the following properties (Burmeister and Dobell, 1970, Chap. 9):

$$\begin{aligned} p_i &= -\partial T/\partial y_i,\ i = 1,2, & (25.21)\\ y_i &= y_i(k_1, k_2, p_1, p_2), & (25.22)\\ w_i &= \partial T/\partial k_i = w_i(k_1, k_2, p_1, p_2), & (25.23)\end{aligned}$$

where p_i is the price of the i-th of capital good (*qua* commodity) and w_i is the rental rate (or factor reward) of the i-th capital good (*qua* factor of production). All prices are in terms of the consumption good taken as *numéraire*.

It is also known that, in these general equilibrium systems, the effect of an increase in the endowment of a factor on the output of a commodity, at unchanged prices and rentals, is exactly the same as the effect of an increase in the price of that commodity, other things being equal, on that factor's rental. These reciprocity relations are known in the theory of international trade as the duality between the Rybczynski and Stolper-Samuelson theorems (Gandolfo, 1994, Chap. 6, Sect. A.6.3). Thus we note

$$\frac{\partial y_i}{\partial k_j} = \frac{\partial w_j}{\partial p_i},\ i,j = 1,2. \tag{25.24}$$

We finally recall another property (also used in the theory of international trade in the so-called factor-price equalization theorem: see Gandolfo, 1994, Chap. 4, Sect. A.4.3), according to which, under the assumption of no factor-intensity reversal, factor rewards depend solely on goods prices, which means that

$$\frac{\partial w_i}{\partial k_j} = 0, \quad i,j = 1,2. \tag{25.25}$$

Let us now pass to the dynamics. The output of capital goods goes to increase the respective stocks, account being taken of depreciation and, since the variables are per capita, of the increase in the labour force. Hence we have

$$k_i' = y_i - \sigma k_i, \tag{25.26}$$

where σ is the sum of the rate of increase of the labour force and of the rate of depreciation, which is taken to be the same for both capital goods.

25.2.5.2 The optimality conditions

The problem is now to maximize the functional

$$\begin{aligned} \max_{k_1,k_2} W &= \int_0^\infty e^{-\delta t} U(c)\mathrm{d}t = \int_0^\infty e^{-\delta t} U(T(k_1, k_2, y_1, y_2))\mathrm{d}t & (25.27)\\ &\text{subject to}\\ k_i' &= y_i - \sigma k_i, \end{aligned}$$

where δ is the social discount rate. We can apply Pontryagin's maximum principle (see Chap. 22, Sect. 22.1.2) and start by writing the Hamiltonian

$$H = e^{-\delta t}\{U(T(k_1, k_2, y_1, y_2)) + q_1 (y_1 - \sigma k_1) + q_2 (y_2 - \sigma k_2)\}, \tag{25.28}$$

where q_1, q_2 are two costate variables.

Considering for simplicity only an interior maximum of the Hamiltonian we have

$$U_c \frac{\partial T}{\partial y_i} + q_i = 0,$$

from which

$$q_i = -U_c \frac{\partial T}{\partial y_i}. \tag{25.29}$$

If we take Eqs. (25.21) and (25.29) into account we see that

$$q_i = U_c p_i.$$

To simplify the analysis we follow Benhabib and Nishimura (1979) in assuming that, in the neighbourhood of the optimal solution, marginal utility of consumption is constant. We can then normalize U_c to unity, and write

$$q_i = p_i. \tag{25.30}$$

The canonical equations for the costate variables are

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\left(e^{-\delta t} q_i\right) &= -\frac{\partial H}{\partial k_i} = -e^{-\delta t}\left(U_c \frac{\partial T}{\partial k_i} - \sigma q_i\right), \\ &\text{namely} & (25.31)\\ -\delta e^{-\delta t} q_i + e^{-\delta t} q_i' &= e^{-\delta t}\left(\sigma q_i - \frac{\partial T}{\partial k_i}\right), \end{aligned}$$

from which, eliminating $e^{-\delta t}$ and rearranging terms, we have

$$q_i' = (\delta + \sigma) q_i - \frac{\partial T}{\partial k_i}. \tag{25.32}$$

These equations can be written, given Eqs. (25.30) and (25.23), as

$$p_i' = (\delta+\sigma)p_i - \frac{\partial T}{\partial k_i} = -w_i(k_1,k_2,p_1,p_2) + (\delta+\sigma)p_i. \tag{25.33}$$

The equations of motion (25.26), account being taken of Eqs. (25.22) can be rewritten as

$$k_i' = y_i(k_1,k_2,p_1,p_2) - \sigma k_i. \tag{25.34}$$

Hence the dynamics of the optimal path is described by the 4×4 differential system

$$\begin{aligned} k_i' &= y_i(k_1,k_2,p_1,p_2) - \sigma k_i,\\ p_i' &= -w_i(k_1,k_2,p_1,p_2) + (\delta+\sigma)p_i,\\ i &= 1,2. \end{aligned} \tag{25.35}$$

25.2.5.3 Emergence of a Hopf bifurcation

The procedure to follow is the usual one: we evaluate the Jacobian of system (25.35) at the equilibrium point $k_i' = p_i' = 0$, and check whether the conditions of the Hopf bifurcation theorem are satisfied. This Jacobian is

$$\mathbf{J} = \begin{bmatrix} \left(\frac{\partial y_1}{\partial k_1} - \sigma\right) & \frac{\partial y_1}{\partial k_2} & \frac{\partial y_1}{\partial p_1} & \frac{\partial y_1}{\partial p_2} \\ \frac{\partial y_2}{\partial k_1} & \left(\frac{\partial y_2}{\partial k_2} - \sigma\right) & \frac{\partial y_2}{\partial p_1} & \frac{\partial y_2}{\partial p_2} \\ -\frac{\partial w_1}{\partial k_1} & -\frac{\partial w_1}{\partial k_2} & -\left(\frac{\partial w_1}{\partial p_1} - \sigma\right) + \delta & -\frac{\partial w_1}{\partial p_2} \\ -\frac{\partial w_2}{\partial k_1} & -\frac{\partial w_2}{\partial k_2} & -\frac{\partial w_2}{\partial p_1} & -\left(\frac{\partial w_2}{\partial p_2} - \sigma\right) + \delta \end{bmatrix}, \tag{25.36}$$

where all derivatives are evaluated at the equilibrium point. Given Eqs. (25.24) and (25.25), this Jacobian becomes

$$\mathbf{J} = \begin{bmatrix} \left(\frac{\partial y_1}{\partial k_1} - \sigma\right) & \frac{\partial y_1}{\partial k_2} & \frac{\partial y_1}{\partial p_1} & \frac{\partial y_1}{\partial p_1} \\ \frac{\partial y_2}{\partial k_1} & \left(\frac{\partial y_2}{\partial k_2} - \sigma\right) & \frac{\partial y_2}{\partial p_1} & \frac{\partial y_2}{\partial p_2} \\ 0 & 0 & -\left(\frac{\partial y_1}{\partial k_1} - \sigma\right) - \delta & -\frac{\partial y_2}{\partial k_1} \\ 0 & 0 & -\frac{\partial y_1}{\partial k_2} & -\left(\frac{\partial y_2}{\partial k_2} - \sigma\right) - \delta \end{bmatrix}. \tag{25.37}$$

Given this triangular form, the latent roots of the Jacobian will be given by the roots of the upper left-hand and lower right-hand matrices, namely by

the roots of $\mathbf{J}_{11}$ and $\mathbf{J}_{22}$, where

$$\mathbf{J}_{11} \equiv \begin{bmatrix} \left(\frac{\partial y_1}{\partial k_1} - \sigma\right) & \frac{\partial y_1}{\partial k_2} \\ \frac{\partial y_2}{\partial k_1} & \left(\frac{\partial y_2}{\partial k_2} - \sigma\right) \end{bmatrix}, \tag{25.38}$$

and

$$\begin{aligned} \mathbf{J}_{22} &\equiv \begin{bmatrix} -\left(\frac{\partial y_1}{\partial k_1} - \sigma\right) + \delta & -\frac{\partial y_2}{\partial k_1} \\ -\frac{\partial y_1}{\partial k_2} & -\left(\frac{\partial y_2}{\partial k_2} - \sigma\right) + \delta \end{bmatrix} \\ &= \begin{bmatrix} -\left(\frac{\partial y_1}{\partial k_1} - \sigma\right) & -\frac{\partial y_2}{\partial k_1} \\ -\frac{\partial y_1}{\partial k_2} & -\left(\frac{\partial y_2}{\partial k_2} - \sigma\right) \end{bmatrix} + \delta\mathbf{I} \\ &= -\mathbf{J}_{11}^{\mathrm{T}} + \delta\mathbf{I}. \end{aligned} \tag{25.39}$$

Since a matrix and its transpose have the same roots, the roots of $\mathbf{J}_{22}$ are equal to δ minus the roots of $\mathbf{J}_{11}$.

The characteristic equation of $\mathbf{J}_{11}$ is

$$\begin{aligned} &\lambda^2 - \left[\left(\frac{\partial y_1}{\partial k_1} - \sigma\right) + \left(\frac{\partial y_2}{\partial k_2} - \sigma\right)\right]\lambda \\ &+ \left[\left(\frac{\partial y_1}{\partial k_1} - \sigma\right)\left(\frac{\partial y_2}{\partial k_2} - \sigma\right) - \frac{\partial y_1}{\partial k_2}\frac{\partial y_2}{\partial k_1}\right] = 0, \end{aligned} \tag{25.40}$$

which may well have a pair of complex conjugate roots (see exercise 3). Then the four roots of the Jacobian will be

$$\theta_1 \pm \mathrm{i}\omega,\ \theta_2 \pm \mathrm{i}\omega, \text{ where } \theta_2 \equiv \delta - \theta_1. \tag{25.41}$$

It is now easy to see that, for $\delta = \theta_1$, the Jacobian will have a pair of pure imaginary roots and no other root with zero real part. Furthermore, $\mathrm{d}\theta_2/\mathrm{d}\delta = 1 \neq 0$. The conditions of the Hopf bifurcation theorem are satisfied, which proves the existence of a closed orbit around the equilibrium point. Since this orbit satisfies the dynamic equations derived from the maximum principle, it is optimal.

Generalization of the above treatment to the case of n>2 capital goods is straightforward: see Benhabib and Nishimura, 1979. Further study and generalizations are contained in Medio (1987) and Cartigny and Venditti (1994). The possibility of oscillatory movement in an optimal growth multi-sector model is by no means a curiosum. Montrucchio (1992), by examining the inverse optimal problem has shown that, given an arbitrary differential equation system, it can be derived as the result of a well-behaved concave optimization problem. Hence in general any kind of movement (not only oscillatory, but also chaotic) can arise from a higher-dimensional optimal growth problem.

In conclusion we note that, as is clear from our simplified treatment, the parameter responsible for the emergence of cycles (this is also true for other types of movements, as shown by Montrucchio, 1992) is the social discount rate δ. Only if δ is 'sufficiently great' (i.e., equal to θ_1) can cycles emerge.

25.2.6 Cycles in an IS-LM model with pure money financing

The traditional dynamic IS-LM with pure money financing of the government budget deficit is given by

$$\begin{aligned}
Y' &= \alpha[I(Y,R)+G-S(Y^D)-T(Y)], \quad && I_R<0, S_{Y^D}>0, I_Y>0,\\
& && 0<T_Y<1,\\
R' &= \beta[L(Y,R)-M], \quad && L_Y>0, L_R<0,\\
M' &= G-T(Y),
\end{aligned} \tag{25.42}$$

where Y is income, I investment, R the rate of interest, S saving, $Y^D \equiv Y-T$, disposable income, T taxation, L money demand, M the money stock. The differential equations represent the usual assumptions of income reacting to excess demand with an adjustment speed $\alpha>0$, the rate of interest adjusting to excess demand for money with an adjustment speed $\beta>0$, and the money supply changing to finance the government deficit, which is assumed to be purely money financed. The partial derivatives have the standard signs.

The existence of oscillations in this model was studied by Schinasi (1981, 1982), who assumed $M=L(Y,R)$, namely an instantaneous adjustment in the money market, so as to reduce the model to a 2×2 system and hence apply the relevant mathematical tools for proving the existence of oscillations in planar systems (relaxation oscillations, and Poincaré-Bendixson theorem: see Chap. 24, Sect. 24.3.1). The full 3×3 dynamic model has been studied by Sasakura (1994) in the context of the Hopf bifurcation.

Under weak assumptions (Sasakura, 1994), the system has a unique positive equilibrium point, that we now examine for the possible emergence of cycles. The Jacobian is

$$\mathbf{J}=\begin{bmatrix} \alpha(I_Y-\sigma_Y) & \alpha I_R & 0\\ \beta L_Y & \beta L_R & -\beta\\ -T_Y & 0 & 0 \end{bmatrix}, \tag{25.43}$$

where $\sigma_Y \equiv S_{Y^D}(1-T_Y)+T_Y$, and all derivatives are taken at the equilibrium point. The characteristic equation of the Jacobian is

$$\lambda^3+a_1\lambda^2+a_2\lambda+a_3=0, \tag{25.44}$$

where

$$\begin{aligned}
a_1 &\equiv -[\alpha(I_Y-\sigma_Y)+\beta L_R],\\
a_2 &\equiv \alpha\beta[(I_Y-\sigma_Y)L_R-I_RL_Y],\\
a_3 &\equiv -\alpha\beta T_YI_R.
\end{aligned} \tag{25.45}$$

Under the assumed signs of the various derivatives, $a_2, a_3>0$, but the sign of a_1 remains uncertain if—following the Kaldorian tradition as Schinasi—we assume that $I_Y-\sigma_Y>0$. Since σ_Y is certainly greater than the simple marginal propensity to save, it seems safe to assume with Sasakura that $I_Y-\sigma_Y$, even if positive, is sufficiently small so that $a_1>0$.

Since we want to apply the Hopf bifurcation theorem, we need one negative real root and a pair of pure imaginary roots, hence we apply conditions (25.12) to the characteristic equation (25.44). Since all coefficients are positive, it remains to check that

$$a_1a_2-a_3=0, \tag{25.46}$$

namely

$$\alpha\beta\{-[\alpha(I_Y-\sigma_Y)+\beta L_R][(I_Y-\sigma_Y)L_R-I_RL_Y]+T_YI_R\}=0. \tag{25.47}$$

Here we have two parameters to play with—let us use β and consider α constant. Then from (25.47) we find β_0, the critical value of β:

$$\beta_0=\frac{1}{L_R}\left\{\frac{T_YI_R}{(I_Y-\sigma_Y)L_R-I_RL_Y}-\alpha(I_Y-\sigma_Y)\right\}>0. \tag{25.48}$$

Since $a_1a_2-a_3$ changes sign as β passes through β_0, the real part of the complex roots, $\theta(\beta)$, also changes sign as it passes through zero. This, unfortunately, is not sufficient to prove (H.2) in Hopf's theorem, which requires that $\theta(\beta)$ crosses the real axis at non-zero speed (for example, a horizontal inflection point of $\theta(\beta)$ at β_0 would not do). Hence we must check that $\mathrm{d}\theta(\beta)/\mathrm{d}\beta|_{\beta=\beta_0}$ is different from zero. This can be accomplished by using sensitivity analysis—see Chap. 18, Sect.18.2.2.2, in particular Eqs. (18.101), which yield

$$\begin{aligned}
-\frac{\partial\lambda_1}{\partial\beta} \quad & -2\frac{\partial\theta}{\partial\beta} && = \frac{\partial a_1}{\partial\beta},\\
2\theta\frac{\partial\lambda_1}{\partial\beta} \quad & +(2\lambda_1+\theta)\frac{\partial\theta}{\partial\beta} \quad +2\omega\frac{\partial\omega}{\partial\beta} && = \frac{\partial a_2}{\partial\beta},\\
-(\theta^2+\omega^2)\frac{\partial\lambda_1}{\partial\beta} \quad & -2\lambda_1\theta\frac{\partial\theta}{\partial\beta} \quad -2\lambda_1\omega\frac{\partial\omega}{\partial\beta} && = \frac{\partial a_3}{\partial\beta}.
\end{aligned} \tag{25.49}$$

The derivatives $\partial a_i/\partial\beta$ can be easily calculated from Eqs. (25.45), and turn out to be

$$\frac{\partial a_1}{\partial\beta} = -L_R,$$

$$\frac{\partial a_2}{\partial\beta} = \alpha[(I_Y - \sigma_Y)L_R - I_R L_Y] = \frac{a_2}{\beta},$$

$$\frac{\partial a_3}{\partial\beta} = -\alpha T_Y I_R = \frac{a_3}{\beta}.$$

Since we are interested in $\partial\theta(\beta)/\partial\beta|_{\beta=\beta_0}$, where $\theta = 0$, the system to solve is

$$\begin{aligned}
-\frac{\partial\lambda_1}{\partial\beta} - 2\frac{\partial\theta}{\partial\beta} \qquad\qquad &= -L_R,\\
2\lambda_1\frac{\partial\theta}{\partial\beta} \qquad +2\omega\frac{\partial\omega}{\partial\beta} &= \frac{a_2}{\beta_0},\\
-\omega^2\frac{\partial\lambda_1}{\partial\beta} \qquad -2\lambda_1\omega\frac{\partial\omega}{\partial\beta} &= \frac{a_3}{\beta_0}.
\end{aligned} \tag{25.50}$$

The determinant of this system turns out to be $4\omega(\lambda_1^2 + \omega^2) > 0$, and the numerator of the fraction that gives the solution for $\partial\theta/\partial\beta$ by Cramer's rule turns out to be

$$2\lambda_1\omega\frac{a_2}{\beta_0} + 2\omega\frac{a_3}{\beta_0} + 2\omega^3 L_R. \tag{25.51}$$

Since this expression contains two negative and one positive term, it might be zero. It can however be shown that it is negative. In fact, since $a_1 = -\sum_{i=1}^{3}\lambda_i$, and since we are carrying out the calculations at $\beta = \beta_0$ (hence $\theta = 0$ and $a_1a_2 - a_3 = 0$) we have $\lambda_1 = -a_1$. Thus Eq. (25.51) becomes

$$-\frac{2\omega}{\beta_0}(a_1a_2 - a_3) + 2\omega^3 L_R = 2\omega^3 L_R < 0, \tag{25.52}$$

which demonstrates that (H.2) in Hopf's theorem is satisfied. This completes the proof of the existence of an orbit for values of the parameter in the neighbourhood of β_0.

25.3 Bifurcations in discrete time systems

The concept of bifurcation is the same in discrete time systems as in continuous time systems. It should however be noted that in discrete time systems the place of the zero-real-part latent root of the Jacobian is taken by the root with unit modulus. This should be clear by our treatment of linear difference and differential equations in Parts I and II, and is due to the fact

that the solutions to the linearized system appear as $e^{\lambda_i t}$ in continuous time systems and as λ_i^t in discrete time systems, where λ_i denotes a latent root of the Jacobian.

25.3.1 Codimension-one bifurcations

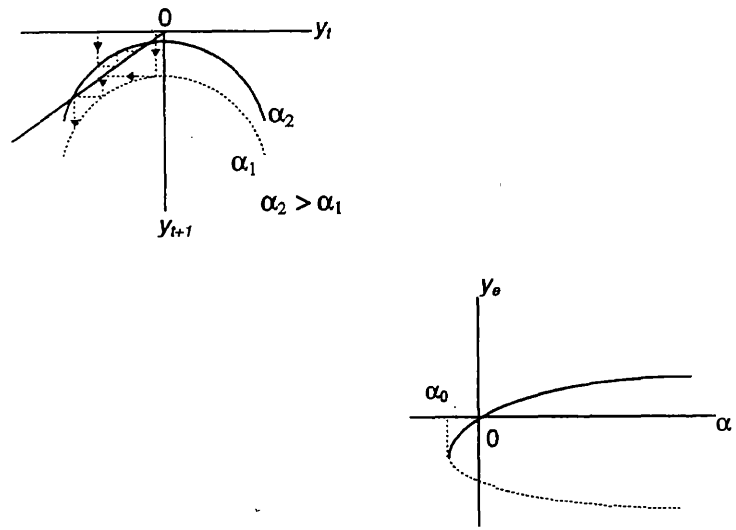


Figure 25.5: The fold bifurcation in discrete time, $y_{t+1} = \alpha - y_t^2$

The basic codimension one bifurcations present in continuous time systems (see Sect. 25.2.1) carry over to the discrete time system

$$y_{t+1} = f(y_t, \alpha) \tag{25.53}$$

by simply replacing the requirement $\partial f(y_{e,0}, \alpha_0)/\partial y = 0$ with the requirement $\partial f(y_{e,0}, \alpha_0)/\partial y = 1$. Given this, we have:

Saddle-node (or fold) bifurcation.
This bifurcation occurs when
$\partial^2 f(y_{e,0}, \alpha_0)/\partial y^2 \neq 0$,
and
$\partial f(y_{e,0}, \alpha_0)/\partial \alpha \neq 0$.

Prototype equations: $y_{t+1} = \alpha - y_t^2,\ y_{t+1} = y_t + \alpha - y_t^2$.

Transcritical bifurcation.
This occurs when
$\partial^2 f(y_{e,0}, \alpha_0)/\partial y^2 \neq 0$,
and
$\partial^2 f(y_{e,0}, \alpha_0)/\partial\alpha\partial y \neq 0$.
Prototype equations: $y_{t+1} = \alpha y_t - y_t^2,\ y_{t+1} = y_t + \alpha y_t - y_t^2$.

Pitchfork bifurcation.
This occurs when
$\partial^3 f(y_{e,0}, \alpha_0)/\partial y^3 \neq 0$,
and
$\partial^2 f(y_{e,0}, \alpha_0)/\partial\alpha\partial y \neq 0$.
Prototype equations: $y_{t+1} = \alpha y_t - y_t^3,\ y_{t+1} = y_t + \alpha y_t - y_t^3$.

Figure (25.5) shows a fold bifurcation. In interpreting the part where the phase diagram is illustrated, it should be kept in mind that the equilibrium points are the intersections of the phase curve with the 45° line (see Sect. 21.5.1).

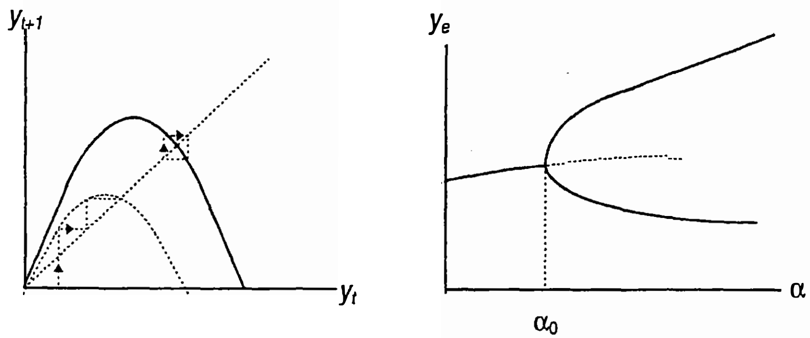


Figure 25.6: The supercritical flip bifurcation, $y_{t+1} = \alpha y_t - \alpha y_t^2$

A bifurcation that cannot arise in continuous time dynamical systems but only in discrete time dynamical systems is the *flip bifurcation,* an example of which is illustrated in Fig. 25.6. Its peculiarity derives from the fact that it arises when $\partial f(y_{e,0}, \alpha_0)/\partial y = -1$. In fact, while $\lambda^t = 1$ for $\lambda = 1$ is equivalent to $e^{\lambda t} = 1$ for $\lambda = 0$, the case $\lambda = -1$, which gives rise to alternations, has no equivalent in continuous time. Thus we have (Guckenheimer and Holmes, 1986, Chap. 3, Sect. 3.5):

Flip bifurcation (prototype equations: $y_{t+1} = \alpha y_t - \alpha y_t^2,\ y_{t+1} = -y_t - \alpha y_t - \alpha y_t^3$).

Consider the one-parameter difference equation (25.53) and assume that when $\alpha = \alpha_0$ there is an equilibrium $y_{e,0}$ for which the following hypotheses are satisfied:

(a) $\dfrac{\partial f}{\partial y} = -1$,

(b) $\left(\dfrac{\partial f}{\partial \alpha}\dfrac{\partial^2 f}{\partial y^2} + 2\dfrac{\partial^2 f}{\partial\alpha\partial y}\right) \neq 0$,

(c) $a \equiv \left[3\left(\dfrac{\partial^2 f}{\partial y^2}\right)^2 + 2\dfrac{\partial^3 f}{\partial y^3}\right] \neq 0$,

where all the partial derivatives are evaluated at $(y_{e,0}, \alpha_0)$. Then, depending on the signs of the expressions (b) and (c),

(i) the equilibrium $y_{e,0}$ is stable (unstable) when $\alpha < \alpha_0$ ($\alpha > \alpha_0$);

(ii) the equilibrium $y_{e,0}$ becomes unstable (stable) for each parameter value $\alpha > \alpha_0$ ($\alpha < \alpha_0$), and a branch of additional stable (unstable) equilibria $y_e(\alpha)$ *of order 2* emerges.

To clarify this definition we must explain what an equilibrium point of order 2 is. From the difference equation (25.53) we have

$$y_{t+2} = f(y_{t+1}, \alpha) = f(f(y_t, \alpha), \alpha) = f^{(2)}(y_t, \alpha), \tag{25.54}$$

where the notation $f^{(2)}$ simply means that the operator f has been applied twice. An equilibrium point of the difference equation (25.54) is called an equilibrium point (or fixed point) of order 2. While an equilibrium point of the difference equation (25.53) is a stationary point (also called an equilibrium point of order 1) by definition (when $y_e = f(y_e)$, then y equals y_e for any t), an equilibrium point of order 2 is a point that repeats itself every two periods, i.e. a constant-amplitude alternation (see Chap. 3, Sect. 3.1), also called an orbit of period two.

This explains why the flip bifurcation is also called a *period doubling* bifurcation. When the expression a in (b) is positive (negative), the emerging equilibrium points of order 2 are stable (unstable), and the flip bifurcation is said to be *supercritical (subcritical)* respectively.

25.3.2 The Hopf bifurcation in discrete time

We have seen the importance of the Hopf bifurcation in continuous time systems. A similar theorem exists for discrete time systems, but unfortunately only for 2×2 systems. As in the case of continuous-time systems, and for the same reasons, we shall give solely the existence part (for the full theorem see Guckenheimer and Holmes, 1986, Chap. 3, Sect. 3.5).

Hopf bifurcation theorem for discrete-time systems (existence part)
Consider the two-dimensional non-linear difference system with one parameter

$$\mathbf{y}_{t+1} = \varphi(\mathbf{y}_t, \alpha), \tag{25.55}$$

and suppose that for each α in the relevant interval it has a smooth family of equilibrium points $\mathbf{y}_e = \mathbf{y}_e(\alpha)$ at which the eigenvalues are complex conjugate, $\lambda_{1,2} = \theta(\alpha) \pm i\omega(\alpha)$. If there is a critical value α_0 of the parameter such that
(SH.1) the eigenvalues' modulus becomes unity at α_0, but the eigenvalues are not roots of unity (from the first up to the fourth), namely

$$|\lambda_{1,2}(\alpha_0)| = +\sqrt{\theta^2+\omega^2} = 1,\ \lambda^j_{1,2}(\alpha_0) \neq 1 \text{ for } j = 1,2,3,4;$$

(SH.2) $\left.\dfrac{\mathrm{d}\,|\lambda_{1,2}(\alpha)|}{\mathrm{d}\alpha}\right|_{\alpha=\alpha_0} \neq 0;$

THEN there is an invariant closed curve bifurcating from α_0.

The analogies with the Hopf bifurcation theorem in continuous time will be obvious if one recalls from Chaps. 5 (Sect. 5.1.3) and 14 (Sect. 14.1.3) that—when the roots are complex—a constant-amplitude trigonometric oscillation takes place when the real part is zero in the case of continuous time, and when the modulus of the roots is unity in the case of discrete time. In other words, the unit circle in the complex plane takes the place of the left-hand part of the complex plane, and the requirement that the roots cross the imaginary axis is replaced by the requirement that they cross the unit circle; in both cases this crossing must take place with non-zero speed, hence requirement (H.2).

The requirement $\lambda^j_{1,2}(\alpha_0) \neq 1$ (on the roots of unity see Chap. 18, Sect. 18.4.1) is related to the fact that bifurcation structures associated with fixed points that are third or fourth roots of unity are special, and hence must be excluded to proceed with the Hopf bifurcation analysis (Guckenheimer and Holmes, 1986, p. 162).

25.3.3 Kaldor's cyclical model in discrete time

We have applied continuous Hopf bifurcation theory to this model in Sect. 25.2.4. By reformulating the model in discrete time it is easy to apply the discrete time version of the Hopf bifurcation. If we replace the derivatives

with first differences, $\Delta y_t = y_{t+1} - y_t$, we can write

$$\begin{aligned} Y_{t+1} - Y_t &= \alpha[I(Y_t, K_t) - S(Y_t, K_t)],\quad I_Y > S_Y > 0,\ I_K < 0, S_K > 0, \\ K_{t+1} - K_t &= I(Y_t, K_t), \end{aligned} \tag{25.56}$$

namely

$$\begin{aligned} Y_{t+1} &= \alpha[I(Y_t, K_t) - S(Y_t, K_t)] + Y_t, \\ K_{t+1} &= I(Y_t, K_t) + K_t, \end{aligned} \tag{25.57}$$

whose Jacobian is

$$\mathbf{J} = \begin{bmatrix} \alpha(I_Y - S_Y) + 1 & \alpha(I_K - S_K) \\ I_Y & I_K + 1 \end{bmatrix}. \tag{25.58}$$

The characteristic equation is

$$\lambda^2 + a_1\lambda + a_2 = 0, \tag{25.59}$$

where

$$\begin{aligned} a_1 &= \alpha(S_Y - I_Y) - I_K - 2, \\ a_2 &= \alpha(I_Y - S_Y)(1 + I_K) + (I_K + 1) - \alpha(I_K - S_K)I_Y. \end{aligned} \tag{25.60}$$

For the roots to be complex conjugate and not pure imaginary we must have $a_1^2 - 4a_2 < 0$. With the given signs of the partial derivatives, a_2 may be either positive or negative. In the latter case there can be no complex roots. We assume that $a_2 > 0$, and that a_1^2 is sufficiently small so that the roots are complex.

We know (Chap. 5, Sect. 5.1.4) that the modulus of the complex roots equals $+\sqrt{a_2}$, hence the unit modulus condition is $a_2 = 1$, which gives the critical value of the parameter

$$\alpha_0 = \frac{-I_K}{(I_Y - S_Y)(1 + I_K) - (I_K - S_K)I_Y}. \tag{25.61}$$

This value is not acceptable when the denominator is negative, hence we assume that it is positive. We also assume that the condition on the roots of unity is satisfied.

We have now to calculate the derivative of the modulus with respect to α, i.e.

$$\frac{\mathrm{d}(a_2)^{\frac{1}{2}}}{\mathrm{d}\alpha} = \frac{1}{2}[\alpha(I_Y - S_Y)(1 + I_K) + (I_K + 1) - \alpha(I_K - S_K)I_Y]^{-\frac{1}{2}} \times$$

$$
\begin{aligned}
& \times[(I_Y - S_Y)(1+I_K) - (I_K - S_K)I_Y] \\
= & \frac{1}{2}(a_2)^{-\frac{1}{2}}[(I_Y - S_Y)(1+I_K) - (I_K - S_K)I_Y] \\
= & \frac{1}{2}[(I_Y - S_Y)(1+I_K) - (I_K - S_K)I_Y] \\
= & \frac{-I_K}{2\alpha_0} > 0. \qquad (25.62)
\end{aligned}
$$

Hence (SH.2) is satisfied, and there is a cycle.

Apart from the different critical value of the parameter, the reader should note the heavy amount of additional assumptions that we have had to introduce to obtain the desired result, a result that, on the contrary, was inherent in the continuous time formulation (see above, Sect. 25.2.4). This shows that replacing derivatives with differences, which might seem quite natural, is by no means an innocuous operation. Since Kaldor, in his presentation of the model, clearly reasoned in terms of continuous movements through time, the continuous time formulation is certainly more appropriate. For a general treatment of continuous vs discrete time in dynamic economic models see Chap. 27, Sect. 27.2.

25.3.4 Liquidity costs in the firm

In real-world economies money plays a central role. Firms hold money and use it to purchase means of production and hire labour. They can also use sales proceeds for the same purpose. The following model is due to Foley (1992).

25.3.4.1 The model

Let us consider a representative firm that holds money M_t and begins each period with a stock of output Q_t from its last period's production. Hence its liquidity is

$$L_tP_t = M_t + P_tQ_t, \qquad (25.63)$$

where the price level P_t is exogenously given. The firm makes real capital outlays C_t in the period to buy means of production and labour services, that are used in fixed proportions to produce a homogeneous output. The production process lasts one period; workers are paid their wages at the beginning of the period and spend them within the period on output.

The key assumption in this model is that the productivity of the inputs (means of production and labour) in producing output for the next period *depends on the level of real liquidity* ($L_t = M_t/P_t + Q_t$). This, as Foley notes, is a generalization of the traditional cash-in-advance model. This latter assumes that the marginal productivity of physical inputs is constant as long as their value is lower than liquidity, but drops to zero if the firm

tries to purchase an amount of inputs whose value is greater than liquidity. In Foley's model, on the contrary, such a purchase is possible: 'we could imagine that the firm has to divert workers and inputs from the production process in order to manage its liquidity position more closely as the value of its capital outlays, P_tC_t, rises in relation to its nominal liquidity' (Foley, 1992, p. 1074).

This gives rise to the production function

$$Q_{t+1} = F(C_t, L_t), \qquad (25.64)$$

which is assumed to be concave in C and L.

The firm uses its current liquidity (P_tL_t) for capital outlays (P_tC_t) and to distribute dividends (P_tD_t). The difference is carried over to the next period; this difference is clearly $P_tL_t - (P_tC_t + P_tD_t)$. Thus liquidity evolves according to the rule

$$
\begin{aligned}
P_{t+1}L_{t+1} &= P_t(L_t - C_t - D_t) + P_{t+1}Q_{t+1} \\
&= P_t(L_t - C_t - D_t) + P_{t+1}F(C_t, L_t), \qquad (25.65)
\end{aligned}
$$

since $Q_{t+1} = F(C_t, L_t)$ by Eq. (25.64). Thus we have

$$L_{t+1} = \frac{P_t}{P_{t+1}}(L_t - C_t - D_t) + F(C_t, L_t). \qquad (25.66)$$

Each firm is owned by a household that receives and consumes dividends. This simplifies the analysis, for, given a time-invariant utility function $u(D_t)$, the firm is assumed to maximize its owners' discounted utility subject to the budget constraint (25.66), that already incorporates the production function. The firm is also assumed to know the future path of the price level with certainty. Thus we have to solve

$$
\begin{aligned}
& \max \sum_{t=0}^{\infty} \beta^t u(D_t), \\
0 < & \ \beta < 1, \qquad (25.67) \\
& \text{sub } (25.66),
\end{aligned}
$$

where β is the discount factor. The first-order conditions for this maximum problem are (Foley, 1992, p. 1076)

$$
\begin{aligned}
\frac{u'(D_t)}{P_t} &= \beta\frac{u'(D_{t+1})}{P_{t+1}}\left(1 + \frac{F_L(C_{t+1}, L_{t+1})}{F_C(C_{t+1}, L_{t+1})}\right), \\
\frac{P_t}{P_{t+1}} &= F_C(C_t, L_t). \qquad (25.68)
\end{aligned}
$$

These equations incorporate familiar results in microeconomic theory. The capitalist equates the ratio of the marginal utility of consumption to the price level today to the discounted value of the same ratio tomorrow, account being taken of the liquidity loss from consuming. The firm equates marginal revenue P_{t+1} to marginal cost of capital outlays $P_t/F_C(C_t, L_t)$.

Let us now pass to macroeconomic equilibrium. Money is supplied by a monetary authority that injects it in the system by purchasing output, as there are no assets other than money and output. The authority follows a rule geared to the level of real money balances in the economy

$$\frac{M_{t+1}}{M_t} = \mu \frac{M_t}{P_t}. \tag{25.69}$$

From the definition of liquidity, Eq. (25.63), we have $M_t/P_t = L_t - Q_t$. Under perfect competition, output is entirely distributed to factors of production including capitalists, namely $Q_t = C_t + D_t$, hence

$$\frac{M_t}{P_t} = L_t - C_t - D_t. \tag{25.70}$$

In the goods market, the supply of output in period *t+1* is $Q_{t+1} = F(C_t, L_t)$, while the demand for output consists of three components:

i) capitalists consumption D_{t+1};

ii) capital outlays C_{t+1}, because it was assumed that workers spend their wages within the period in which they receive them;

iii) real monetary injections, $(M_{t+1} - M_t)/P_{t+1}$, because we have assumed that the monetary authorities can only supply money by purchasing output.

Thus the equilibrium condition on the goods market is

$$F(C_t, L_t) = D_{t+1} + C_{t+1} + (M_{t+1} - M_t)/P_{t+1}. \tag{25.71}$$

Since $(M_{t+1} - M_t) = [M_{t+1}/M_t - 1]M_t$, by using Eqs. (25.69) and (25.70) we have

$$\begin{aligned} \frac{M_{t+1} - M_t}{P_{t+1}} &= \frac{[M_{t+1}/M_t - 1]M_t}{P_{t+1}} = \frac{[\mu(M_t/P_t) - 1]M_t}{P_{t+1}} \\ &= \frac{[\mu(L_t - C_t - D_t) - 1]M_t}{P_{t+1}} \\ &= [\mu(L_t - C_t - D_t) - 1](L_t - C_t - D_t)\frac{P_t}{P_{t+1}}. \end{aligned} \tag{25.72}$$

Substitution of Eq. (25.72) into Eq. (25.71) yields, after rearranging terms and observing that $P_t/P_{t+1} = F_C(C_t, L_t)$ from the optimum conditions (25.68), the following equation

$$C_{t+1} = F(C_t, L_t) - D_{t+1} - [\mu(L_t - C_t - D_t) - 1](L_t - C_t - D_t)F_C(C_t, L_t). \tag{25.73}$$

The stationary equilibrium $[P, M, C, D, L]$ of the model satisfies the conditions

$$\begin{aligned} F(C, L) &= C + D, \\ 1/\beta &= 1 + F_L(C, L)/F_C(C, L), \\ 1 &= F_C(C, L), \\ \mu(M/P) &= 1, \\ P &= M/(L - C - D). \end{aligned} \tag{25.74}$$

The equations are self-explanatory; the only notation is that the last but one equation is a necessary condition for the monetary policy rule to permit a stationary equilibrium.

25.3.4.2 The dynamics

To simplify the analysis, Foley assumes that capitalist consumption D_t equals its stationary equilibrium level, D, under all circumstances. Hence we can neglect the first equation in (25.68)—the second has already been used—and consider the two dynamic non-linear equations (25.73) and (25.66), where of course $D_t = D$.

Thus we have

$$\begin{aligned} C_{t+1} &= F(C_t, L_t) - D - [\mu(L_t - C_t - D) - 1]\times \\ &\quad \times [L_t - C_t - D]F_C(C_t, L_t) = \varphi_1(C_t, L_t), \\ L_{t+1} &= [L_t - C_t - D]F_C(C_t, L_t) + F(C_t, L_t) = \varphi_2(C_t, L_t). \end{aligned} \tag{25.75}$$

The Jacobian of the system evaluated at the equilibrium point is

$$\mathbf{J} = \begin{bmatrix} \dfrac{\partial \varphi_1}{\partial C_t} & \dfrac{\partial \varphi_1}{\partial L_t} \\ \dfrac{\partial \varphi_2}{\partial C_t} & \dfrac{\partial \varphi_2}{\partial L_t} \end{bmatrix}. \tag{25.76}$$

Assuming for simplicity $F_{CL} = F_{LL} = 0$, we have

$$\begin{aligned} \frac{\partial \varphi_1}{\partial C_t} &= F_C + \mu_1(L - C - D)F_C \\ &\quad + [\mu_1(L - C - D) - 1][F_C - F_{CC}(L - C - D)], \\ \frac{\partial \varphi_1}{\partial L_t} &= F_L - \mu_1(L - C - D)F_C - [\mu_1(L - C - D) - 1]F_C, \\ \frac{\partial \varphi_2}{\partial C_t} &= [L - C - D]F_{CC}, \\ \frac{\partial \varphi_2}{\partial L_t} &= F_C + F_L, \end{aligned} \tag{25.77}$$

where the derivatives are evaluated at the equilibrium point and μ_1 is the value of μ at equilibrium. Further noting that—see Eqs. (25.74)—$F_C = 1$, $M/P = L - C - D$, $\mu_1(L - C - D) = 1$, we obtain

$$\mathbf{J} = \begin{bmatrix} 1 + \mu_1 \frac{M}{P} & F_L - \mu_1 \frac{M}{P} \\ F_{CC} \frac{M}{P} & 1 + F_L \end{bmatrix}. \tag{25.78}$$

The characteristic equation of the Jacobian is

$$\lambda^2 + a_1 \lambda + a_2 = 0, \tag{25.79}$$

where

$$\begin{aligned} a_1 &= -\left(2 + \mu_1 \frac{M}{P} + F_L\right), \\ a_2 &= 1 + F_L \left(1 - F_{CC} \frac{M}{P} + \mu_1 \frac{M}{P}\right) + \mu_1 \frac{M}{P}\left(1 + F_{CC} \frac{M}{P}\right) \\ &= 1 + \mu_1 \frac{M}{P}\left(1 + F_L + F_{CC} \frac{M}{P}\right) + F_L \left(1 - F_{CC} \frac{M}{P}\right). \end{aligned} \tag{25.80}$$

The model can be parametrized by μ_1. This requires $\left(1 + F_L + F_{CC}\frac{M}{P}\right) \neq 0$, for in the contrary case μ_1 would not appear in a_2. We assume that this condition is satisfied.

For the roots to be complex conjugate we must have $a_1^2 - 4a_2 < 0$. It can easily be checked that

$$a_1^2 - 4a_2 = \left(F_L - \mu_1 \frac{M}{P}\right)\left(F_L - \mu_1 \frac{M}{P} + 4F_{CC}\frac{M}{P}\right). \tag{25.81}$$

Since $\mu_1 < 0$, $F_{CC} < 0$, expression (25.81) will be negative for

$$F_L - \mu_1 \frac{M}{P} < -4F_{CC}\frac{M}{P}$$

or

$$\mu_1 \frac{M}{P} > F_L + 4F_{CC}\frac{M}{P}. \tag{25.82}$$

We know (Chap. 5, Sect. 5.1.4) that the modulus of the complex roots equals $+\sqrt{a_2}$, hence the unit modulus condition is $a_2 = 1$, which gives the critical value of the parameter. By setting $a_2 = 1$ it is easy to see that

$$\mu_1^* \frac{M}{P} = -\frac{F_L\left(1 - F_{CC}\frac{M}{P}\right)}{1 + F_L + F_{CC}\frac{M}{P}} \tag{25.83}$$

gives the critical value of the parameter. We can perturb the other parameters slightly so that the eigenvalues are not roots of unity. Hence (SH.1) of the Hopf bifurcation theorem is satisfied.

As regards (SH.2), the derivative of the modulus with respect to μ_1 is

$$\frac{\mathrm{d}(a_2)^{\frac{1}{2}}}{\mathrm{d}\mu_1} = \frac{1}{2} a_2^{-\frac{1}{2}} \frac{\mathrm{d}a_2}{\mathrm{d}\mu_1} = \frac{1}{2} a_2^{-\frac{1}{2}} \frac{M}{P}\left[1 + F_L + F_{CC}\frac{M}{P}\right] \neq 0. \tag{25.84}$$

Since also (SH.2) is satisfied, the model gives rise to a cycle when μ_1 passes through its critical value μ_1^*.

25.4 Exercises

1. Consider a foreign exchange market where the demand for and supply of foreign exchange come solely from traders and where the exchange rate is free to adjust according to excess demand:

$$r' = v[D(r, \alpha) - S(r)],$$

where r is the spot exchange rate defined as number of units of domestic currency per one unit of foreign currency, $v > 0$ is the adjustment speed, and α is a parameter representing exogenous factors that cause a shift in the demand curve such that $\partial D/\partial \alpha > 0$. It is known (Gandolfo, 1995) that a backward-bending supply curve is perfectly normal, hence two points of intersection between the (downward sloping) demand curve and the supply curve as well as a point of tangency between them may well exist on plausible economic grounds. If the supply curve is backward bending then clearly $\partial^2 S/\partial r^2 \neq 0$. Assume that the demand function is linear, hence $\partial^2 D/\partial r^2 = 0$.

Show that the point of tangency between the demand and supply schedules is a fold bifurcation.

2. Introduce a depreciation term in the investment function in the Kaldor model, so that the second dynamic equation of the model becomes $K' = I(Y, K) - \delta K$. Show that the depreciation coefficient δ must be lower than a critical value for a cycle to emerge.

3. Consider the multi-sector optimal growth model and show that the characteristic equation (25.40) will have a pair of complex conjugate roots provided that

$$4\frac{\partial y_1}{\partial k_2}\frac{\partial y_2}{\partial k_1} < -\left[\frac{\partial y_1}{\partial k_1} - \frac{\partial y_2}{\partial k_2}\right]^2.$$

4. The traditional dynamic IS-LM model is given by

$$\begin{aligned} Y' &= \alpha[I(Y,R) - S(Y,R)], \quad I_R < 0, 0 < S_Y < I_Y < 1, S_R > 0, \\ R' &= \beta[L(Y,R) - L_s], \qquad L_Y > 0, L_R < 0, \end{aligned}$$

where Y is income, I investment, R the rate of interest, S saving, L money demand, and L_s the money stock (exogenous). The differential equations represent the usual assumptions of income reacting to excess demand with an adjustment speed $\alpha > 0$ and the rate of interest adjusting to excess demand for money with an adjustment speed $\beta > 0$. The partial derivatives have the standard signs; note however that due to Kaldor's assumption (see Sect. ??) on the slopes of the saving and investment functions the equilibrium point is not necessarily stable. Show that this model has a limit cycle for appropriate values of the parameters α, β (Hint: use the Hopf bifurcation theorem. This exercise is drawn from the paper by Torre, 1977, who apparently was the first to introduce the Hopf bifurcation into economics).

5. The following advertising model is due to Feichtinger (1992):

$$\begin{aligned} X_1' &= k - a(t)X_1X_2 + \beta X_2, \\ X_2' &= a(t)X_1X_2 + \delta X_2, \end{aligned}$$

where X_1 is the number of potential buyers while X_2 is the number of brand users. The number of potential buyers who will purchase the commodity and thus become customers is proportional to the current number of potential buyers and to the number of brand users. The constant of proportionality, $a(t)$, is called the contact rate, hence the term $a(t)X_1X_2$. Current customers switch to a rival brand at a constant rate β, and some leave the market forever (due to migration or mortality) at the rate ϵ, where $\epsilon + \beta = \delta$. This explains the second equation.

It is also assumed that a continuous stream of new potential customers enters the market at the rate k, while those who become users of the brand leave the set of potential consumers. By noting that current customers who switch to a rival brand remain in the set of potential buyers, we have the first equation.

By assuming that the contact rate is an increasing function of advertising expenditure, which in turn is proportional to the number of buyers under certain rules of thumb, we have

$$\begin{aligned} X_1' &= k - \alpha X_1X_2^2 + \beta X_2, \\ X_2' &= \alpha X_1X_2^2 + \delta X_2. \end{aligned}$$

By a transformation of variables, $x_1 = (\alpha k/\delta\varepsilon)X_1$, $x_2 = (\varepsilon/k)X_2$, the system becomes

$$\begin{aligned} x_1' &= \gamma[1 - x_1x_2^2 + \phi(x_2 - 1)], \\ x_2' &= x_1x_2^2 - x_2, \end{aligned}$$

where $\gamma \equiv \alpha k^2/\delta\varepsilon$, $\phi \equiv \beta/\delta$. Parametrize this system by γ and show that there is a positive value of γ such that a Hopf bifurcation takes place (Hint: see Feichtinger, 1992).

25.5 References

Andronov, A.A., A.A. Vitt and S.E. Khaikin, 1966, *Theory of Oscillators,* Chap. II.

Barnett, W.A., M. Salmon and A. Kirman (eds.), 1995, *Nonlinear Dynamics in Economics.*

Beltrami, E., 1987, *Mathematics for Dynamic Modeling,* Chap. 6.

Benhabib, J. and K. Nishimura, 1979, The Hopf Bifurcation and Existence and Stability of Closed Orbits in Multisector Models of Optimal Economic Growth.

Boldrin, M. and M. Woodford, 1990, Equilibrium Models Displaying Endogenous Fluctuations and Chaos.

Burmeister, E. and A.R. Dobell, 1970, *Mathematical Theories of Economic Growth,* Chap. 9.

Cartigny, P. and A. Venditti, 1994, Turnpike Theory: Some New Results on the Saddle Point Property of Equilibria and on the Existence of Endogenous Cycles.

Chiarella, C., 1990, *The Elements of a Nonlinear Theory of Economic Dynamics.*

Dockner, E.J. and G. Feichtinger, 1991, On the Optimality of Limit Cycles in Dynamic Economic Systems.

Feichtinger, G., 1992, Hopf Bifurcation in an Advertising Diffusion Model.

Foley, D., 1992, A Contribution to the Theory of Business Cycles.

Gandolfo, G., 1971, *Mathematical Methods and Models in Economic Dynamics,* Appendix III, § 7.6.

Gandolfo, G., 1994, *International Economics, Vol. I: The Pure Theory of International Trade,* Chap. 4, Sect. A.4.3; Chap. 6, Sect. A.6.3.

Gandolfo, G., 1995, *International Economics, Vol. II: International Monetary Economics and Open-Economy Macroeconomics,* Chap. 12, 12.3.2.

Guckenheimer, J. and P. Holmes, 1986, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields,* Chap. 3.

Hahn, F.H., 1966, Equilibrium Dynamics with Heterogeneous Capital Goods.

Invernizzi, S. and A. Medio, 1991, On Lags and Chaos in Economic Dynamic Models.

Kaldor, N., 1940, A Model of the Trade Cycle.

Lorenz, H.-W., 1993, *Nonlinear Dynamical Economics and Chaotic Motion,* 2nd ed., Chap. 3.

Medio, A., 1987, Oscillations in Optimal Growth Models.

Medio, A., 1992, *Chaotic Dynamics: Theory and Applications to Economics,*

Chap. 2, Sect. 2.7.
Minorsky, N., 1962, *Nonlinear Oscillations,* Chap. 7.
Montrucchio, L., 1992, Dynamical Systems that Solve Continuous-Time Concave Optimization Problems: Anything Goes.
Sasakura, K., 1994, On the Dynamic Behaviour of Schinasi's Business Cycle Model.
Schinasi, G.J., 1981, A Nonlinear Dynamic Model of Short Run Fluctuations.
Schinasi, G.J., 1982, Fluctuations in a Dynamic, Intermediate-run IS-LM Model: Applications of Poincaré-Bendixson Theorem.
Sotomayor, J., 1973, Generic Bifurcations of Dynamical Systems.
Torre, V., 1977, Existence of Limit Cycles and Control in Complete Keynesian System by Theory of Bifurcations.
Turnbull, H.W., 1957, *Theory of Equations.*

Chapter 26

Complex Dynamics

26.1 Introduction

Complex dynamics, according to the definition given by Day (1994, Chap. 1), is dynamic behaviour that is not periodic and not balanced (such as steady state, balanced growth or decline, etc.) and does not converge to a periodic or balanced pattern. Complex dynamics (Day, 1994, p. 4) includes processes that involve

- nonperiodic fluctuations,
- overlapping waves,
- switches in regime or structural change.

Complex dynamics is inherently non-linear, in the sense that it can arise only in non-linear systems: for example, fluctuations arising from linear systems can only be periodic (i.e., regularly cyclical, in which distinct situations are repeated at fixed intervals of time). But not all non-linear systems are complex (for example, fluctuations taking place in non-linear systems may quite well be periodic, as we have seen in the previous chapters). Hence non-linearity is a necessary, but by no means a sufficient condition for complex dynamics.

One of the main instances of complex dynamic behaviour (though not the only one) is the presence of nonperiodic random fluctuations which are non-stochastic, but come instead from a deterministic dynamic system: the so-called *chaos* (sometimes also called *erratic dynamics* or *exotic dynamics*).

Chaos theory is a rapidly growing field. To adequately deal with it would require a book on its own. Hence we can do no more than offer an introduction. Books written for economists, but lacking no mathematical rigour, are H.-W. Lorenz (1993), Medio (1992), Day (1994). Non-technical introductory surveys for economists are Kelsey (1988) and Baumol and Benhabib (1989); Boldrin and Woodford (1990) and Grandmont (1986) are more technical, in increasing order of difficulty. An idiosyncratic but instructive survey is Mirowski (1990). Further references are given below, Sect. 26.4.

To introduce the topic, the definition given above will do: *chaos is apparently stochastic behaviour generated by a dynamic deterministic system.*

By 'apparently stochastic' we mean a random path that at first sight cannot be distinguished from the path generated by a stochastic variable. Since it is presumable that readers of this book own (or have access to) a PC, rather than using the traditional examples such as the Lorenz weather equations etcetera, we take from Brock et al. (1991) the simple example of a computer pseudo random number generator. The algorithm used by the computer is purely deterministic, but what comes out is a series of numbers that looks random, and that will fool any statistician in the sense that it passes all the standard tests of randomness. As a matter of fact, random numbers generated in this way are usually employed in statistical analysis.

A second feature that is often cited as typical of chaotic behaviour of deterministic systems is the *impossibility of predicting the future values of the variable(s) concerned.* This might at first sight seem a contradiction—if we have a dynamic deterministic system, even if we cannot solve it analytically we can simulate it numerically, hence we can compute the value(s) of the variable(s) for any future value of t. This is where another important feature of chaos comes in (as a matter of fact, some take it as defining chaos: see, for example, Brock et al., 1991, p. 9), that is *sensitive dependence on initial conditions.*

Sensitive dependence on initial conditions (henceforth SDIC) means that even very small differences in the initial conditions give rise to widely different paths. In a 'normal' deterministic system, all nearby paths starting very close to one another remain very close in the future. Hence a sufficiently small measurement error in the initial conditions will not affect our deterministic forecasts. On the contrary, in deterministic systems with SDIC, prediction of the future values of the variable(s) would be possible only if the initial conditions could be measured with infinite precision. This is certainly not the case.

All this had already been noted by Poincaré (1908, p. 68), whose often cited sentence runs as follows:

> 'If we knew exactly the laws of nature and the situation of the universe at the initial moment, we could predict exactly the situation of that same universe at a succeeding moment. But even if it were the case that the natural laws had no longer any secret for us, we could still only know the initial situation approximately.[...]It may happen that small differences in the initial conditions produce very great ones in the final phenomena. A small error in the former will produce an enormous error in the latter. Prediction becomes impossible, and we have the fortuitous phenomenon'.

Poincaré's work was apparently unknown to the mathematician who is said to have claimed (in the ante-chaos era) that, given a megacomputer and sufficient funds to collect data, weather could be forecast with accuracy. If the weather equations are chaotic, in the sense that they exhibit SDIC (and they do), no amount of funds and no supercomputer will ever yield the required infinite precision. It was in fact after the meteorologist E.N. Lorenz found SDIC in a system of three differential equations emerging in the theory of turbulence in fluids, that the mathematical study of chaos blossomed (Guckenheimer and Holmes, 1986, Chap. 2, Sect. 2.3).

A generally accepted mathematical definition of chaos does not yet exist. Some, as we have seen, take SDIC as the hallmark of chaos. Others argue that chaotic dynamics depends on the existence of a strange attractor (Guckenheimer and Holmes, 1986; see also below, Sect. 26.3.1), that Mandelbrot (1983) calls a fractal attractor (on these distinctions which are not only terminological see Rosser, 1991, Sect. 2.4, and Mirowski, 1990). Other definitions also exist (Rosser, 1991, Sect. 2.3.2.2). In what follows we shall adopt a pragmatic approach and start with discrete-time systems. The reason is that chaos (in the sense of erratic behaviour and SDIC) can occur already in a single first-order difference equation, while for it to occur in differential equations we need at least a 3×3 system.

26.2 Discrete time systems and chaos

26.2.1 The logistic map

The standard example to start with is the so-called *logistic (*or *Pearl-Verhulst) map*, namely the non-linear first-order difference equation

$$y_{t+1} = f(y_t, \alpha) = \alpha y_t(1 - y_t), \; y_t \in [0,1], \alpha \in [0,4]. \tag{26.1}$$

The function f gives rise to a *unimodal curve* in the (y_{t+1}, y_t) plane, since
$\partial f/\partial y_t = \alpha(1 - 2y_t) = 0$ for $y_t = 1/2$, and
$\partial^2 f/\partial y_t^2 = -2\alpha < 0$.

When $y_t = 1/2$, we have $f = \alpha/4$, which shows that the maximum of the curve shifts upwards when α increases in its interval of definition. The curve intersects the horizontal axis at the origin and at $y_t = 1$.

Figure 26.1 shows the logistic curve for different values of the parameter α.

Apart from the trivial point $y_e = 0$, this difference equation has another equilibrium point at $y_e = \alpha y_e(1 - y_e)$, namely

$$y_e = \frac{\alpha - 1}{\alpha}, \tag{26.2}$$

which is negative for $\alpha < 1$ and becomes zero for $\alpha = 1$.

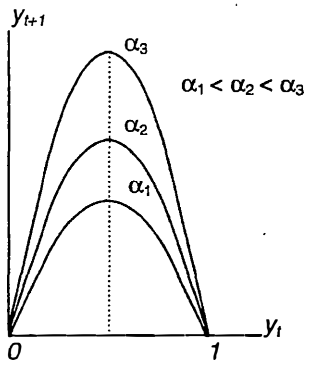


Figure 26.1: The logistic curve $y_{t+1} = \alpha y_t(1 - y_t)$

This fairly simple non-linear difference equation has a bewildering number of possible behaviours, including various types of bifurcations, and chaos. No wonder that it is the pet equation of all those wishing to illustrate complex dynamics in a simple way. Let us first note for future reference that

$$\frac{\partial f}{\partial y_t} = \alpha(1 - 2y_t) = \begin{cases} \alpha & \text{at } y_e = 0, \\ 2 - \alpha & \text{at } y_e = (\alpha - 1)/\alpha. \end{cases} \tag{26.3}$$

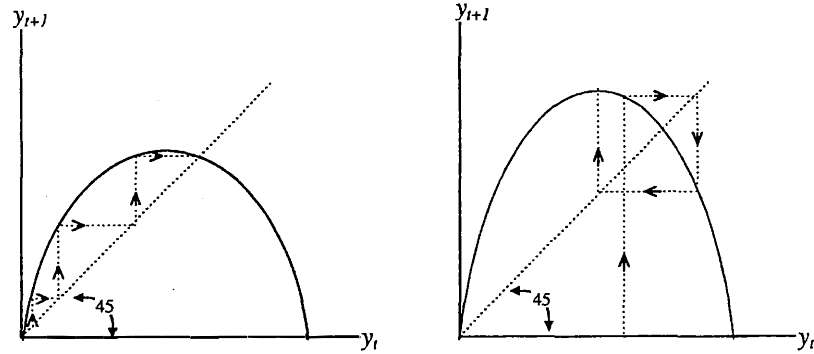
The main types of behaviour can be conveniently classified according to the values of the parameter α.

I) $0 < \alpha < 1$.
In this interval we have $|\partial f/\partial y_t| < 1$ at $y_e = 0$ and $\partial f/\partial y_t > 1$ at $y_e = 1 - 1/\alpha$. Thus the origin is stable while the second (negative) fixed point is unstable.

II) $\alpha = 1$.
At $\alpha = 1$ we have a *transcritical bifurcation* (see Chap. 25, Sect. 25.3), with exchange of stability. The two equilibrium points merge together at the origin and, while the origin is stable for $\alpha < 1$, it becomes unstable for $\alpha > 1$, and a new (positive) stable equilibrium point emerges.

III) $1 < \alpha < 3$.
In this interval nothing new happens. The equilibrium point $1 - 1/\alpha$ remains stable.

IV) $\alpha = 3$.
At $\alpha = 3$ a *flip bifurcation* (see Chap. 25, Sect. 25.3) occurs. Since for this value of α the derivative $\partial f/\partial y_t$ becomes equal to -1, the formerly stable equilibrium point becomes unstable and a new stable equilibrium state of period 2 emerges. This is shown in Fig. 26.2.



(a) stable *(b) unstable*

Figure 26.2: Flip bifurcation in the logistic equation

V) $3 < \alpha \leq \alpha_\infty \approx 3.5699$.
As α increases, the initially stable 2-period cycle loses stability at $\alpha = 1 + \sqrt{6} \approx 3.44949$, where a new, stable 4-period cycle is created through a new flip bifurcation. This sequence of events is repeated again and again and leads to an infinite sequence of flip bifurcations and *period-doublings* (see Fig. 26.3).
The accumulation point of this sequence is $\alpha_\infty \approx 3.5699$ (see Medio, 1992, Sect. 9.4), and the sequence is governed by the 'Feigenbaum constant' $\delta \approx 4.669202$ (this is a constant that appears in all equations with period-doubling, not only in the logistic). This constant is defined as

$$\delta = \lim_{n\to\infty} \frac{\alpha_n - \alpha_{n-1}}{\alpha_{n+1} - \alpha_n},$$

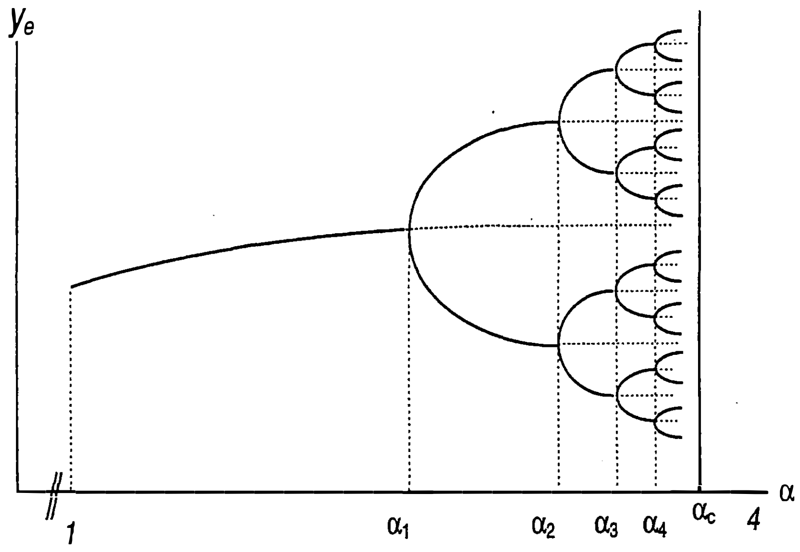


Figure 26.3: Stylized period doubling in the logistic equation

where $\alpha_{n-1}, \alpha_n, \alpha_{n+1}$ denote three successive values of α at which period doubling occurs.

Although period doubling is still in the domain of 'simple' (i.e., non-chaotic) dynamics, chaos begins to occur for values of $\alpha > \alpha_\infty$, hence period doubling is said to be a route to chaos.

VI) $\alpha_\infty < \alpha < 4$.

In this region phenomena other than period doubling can be observed, such as cycles with odd periods, aperiodic oscillations, and so on. Moreover, there is an infinite number of intervals of α-values where low-order periodic cycles (i.e., non-chaotic motion) prevail. These intervals are called *windows* (an inspection of Fig. 26.4 will make this terminology intuitively clear). The region under consideration is usually called the 'chaotic region'.

VII) $\alpha = 4$.

Although $\alpha = 4$ belongs to the chaotic region (VI), we have left this case on its own because of its special importance. In fact, when $\alpha = 4$, it is possible to give a closed-form solution to the logistic equation, and hence to prove the existence of chaos analytically. The solution turns out to be

$$y_t = \frac{1}{2}\left[1 - \cos(2^t u)\right], \tag{26.4}$$

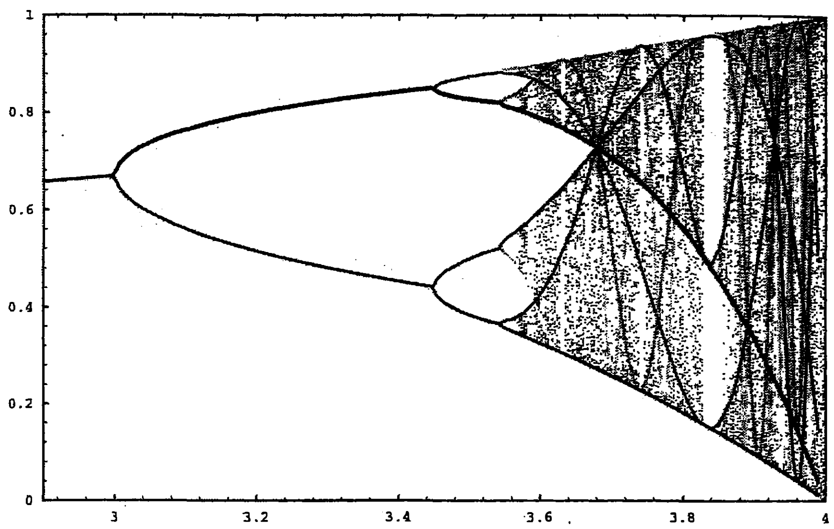


Figure 26.4: Windows in the logistic equation

where the arbitrary constant u depends on the initial condition at $t = 0$, and precisely

$$y_0 = \frac{1}{2}\left[1 - \cos u\right], \ u = \arccos\left(1 - 2y_0\right). \tag{26.5}$$

It can be checked by direct substitution that the function (26.4) identically satisfies Eq. (26.1). We have

$$\begin{aligned} \frac{1}{2}\left[1 - \cos(2^{t+1}u)\right] &= 4\frac{1}{2}\left[1 - \cos(2^t u)\right]\left\{1 - \frac{1}{2}\left[1 - \cos(2^t u)\right]\right\} \\ &= \left[1 - \cos(2^t u)\right]\left[1 + \cos(2^t u)\right] \\ &= 1 - \cos^2(2^t u). \end{aligned} \tag{26.6}$$

Since $\cos(2^{t+1}u) = \cos(2 \cdot 2^t u)$, we can use the elementary trigonometry formula $\cos 2x = 2\cos^2 x - 1$ and obtain

$$\cos(2^{t+1}u) = 2\cos^2(2^t u) - 1. \tag{26.7}$$

If we now substitute (26.7) into the left-hand side of (26.6) we get

$$\frac{1}{2}\left[1 - \cos(2^{t+1}u)\right] = \frac{1}{2}\left[1 - 2\cos(2^{t+1}u) + 1\right] = 1 - \cos^2(2^t u), \tag{26.8}$$

which reduces Eq. (26.6) to an identity. Hence the function (26.4) identically satisfies Eq. (26.1) and, since it contains one arbitrary constant, it is the general solution.

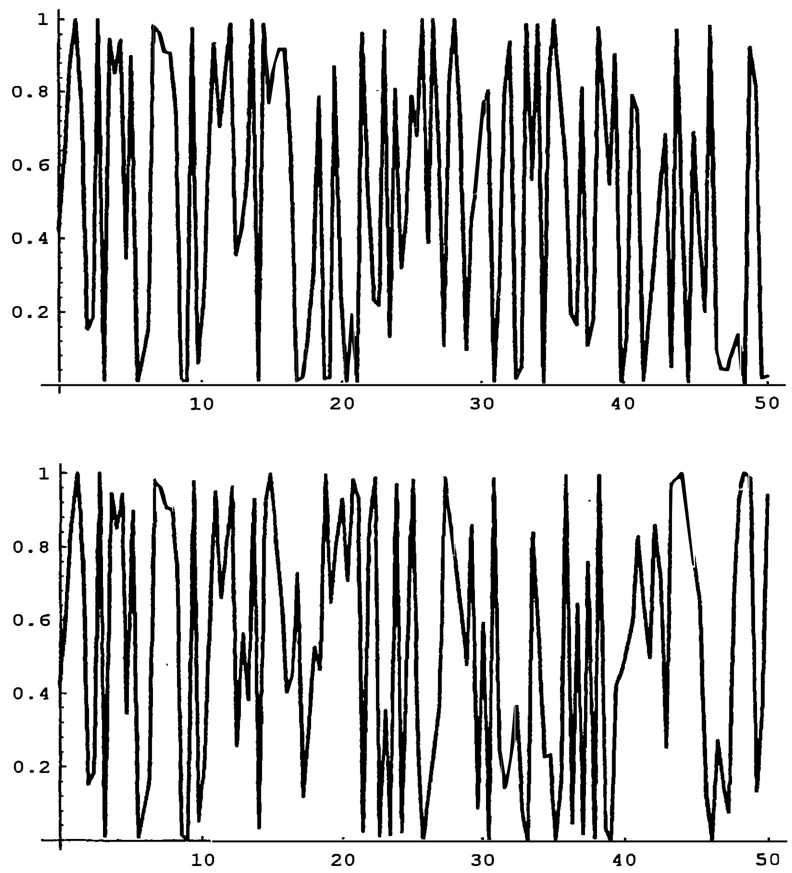


Figure 26.5: Chaos in the logistic difference equation

Equipped with this solution it is easy to show the presence of chaos (see Fig. 26.5).

First note that, if $u/2\pi$ is rational, the solution is periodic, i.e., regularly cyclical, in which distinct situations are repeated at fixed intervals of time. But if $u/2\pi$ is irrational, then the solution is aperiodic. Hence it is sufficient to take the initial condition such that $u/2\pi$ is irrational, and we have

aperiodic fluctuations, which is one of the hallmarks of chaos.

We also have sensitive dependence on initial conditions (SDIC is present in any case, independently of the aperiodicity or periodicity of the solution). Consider an initial state only slightly different from y_0, say a state such that instead of u we have $u + \epsilon$, where ϵ is an arbitrarily small number. The solution (26.4) becomes

$$y_t = \frac{1}{2}\left[1 - \cos(2^t u + 2^t \epsilon)\right], \tag{26.9}$$

where the term $2^t \epsilon$ becomes very large no matter how small ϵ is. Hence the two solutions will diverge, and this proves SDIC, which is another hallmark of chaos.

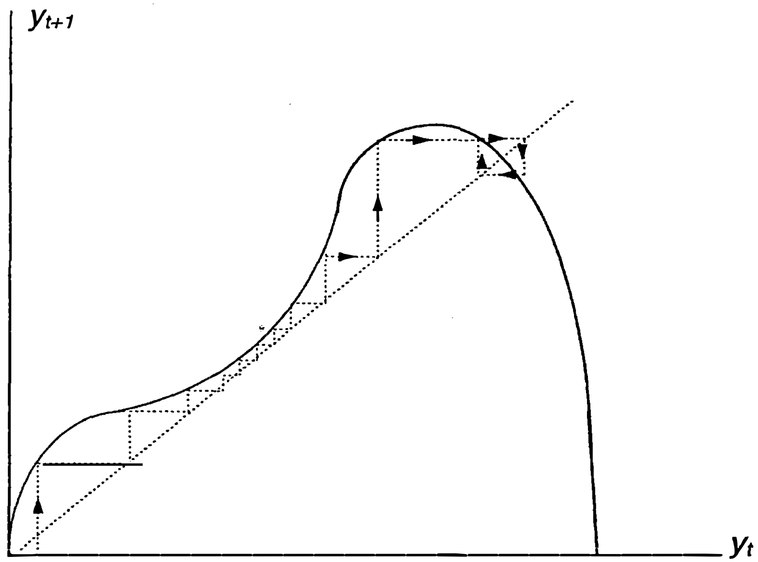


Figure 26.6: An example of intermittency

In Fig. 26.5, we have taken $u = \sqrt{2}$ in the upper panel and $u = \sqrt{2.0001}$ in the lower panel. These are irrational numbers that differ in the fifth decimal place. Aperiodic, random-like fluctuations are quite evident. SDIC makes itself felt already from $t = 13$ onwards.

The possible chaotic behaviour of the logistic map is not a peculiarity—in general, *all* one-parameter unimodal maps may give rise to chaos for appropriate values of the parameter. The logistic map is simply the best known and the easiest to analyse. Other well-known cases are the *tent map* (Medio, 1992, p. 164) and the map

$$y_{t+1} = \alpha y_t e^{-y_t}, \quad y_t \geq 0, \tag{26.10}$$

that is known to give rise to chaos for $\alpha > 14.675$ (Marotto, 1978).

26.2.2 Intermittency

Period-doubling is not the only one route to chaos. Others exist, amongst which the so-called *intermittency*. The behaviour of a variable is said to be *intermittent* when it is subject to large but infrequent variations. Figure (26.6) illustrates a possible case. We see that the phase curve comes close to the 45° line. In that region the sequence $\{y_t\}$ will show only minor changes—if the curve is sufficiently close (but doesn't touch!) the 45° line, the sequence will appear almost constant in the $\{y_t, t\}$ space. After a certain point, however, the sequence will start sharply to increase and, after a while, it will settle down to a period-2 cycle. Other more complicated behaviours can easily be constructed.

We have so far examined one-dimensional systems. In the case of 2×2 systems, a map in which chaos is known to exist is the Henon map (see Guckenheimer and Holmes, 1986, Chap. 5, Sect 5.6).

26.2.3 The basic theorems

What we have heuristically shown in the previous sections can be given a rigorous aspect through several fundamental theorems. We shall state the most important of them without proof (see, for example Guckenheimer and Holmes, 1986, Chap. 6, Sects. 6.3 and 6.8). For future reference let us first define the *Schwarzian derivative* of a function $f(y)$ which is three times differentiable:

$$f^S(y) = \frac{f'''(y)}{f'(y)} - \frac{3}{2}\left(\frac{f''(y)}{f'(y)}\right)^2, \tag{26.11}$$

where of course $f'(y) \neq 0$.

Šarkovskii theorem
Consider the following ordering of all positive integers:

$$\begin{array}{l}
1 \prec 2 \prec 4 \prec 8 \prec 16 \prec \ldots \prec 2^k \prec 2^{k+1} \prec \ldots \\
\qquad \ldots \\
\ldots \prec 2^{k+1}(2n+1) \prec 2^{k+1}(2n-1) \prec \ldots \prec 2^{k+1}5 \prec 2^{k+1}3 \prec \ldots \\
\ldots \prec 2^k(2n+1) \prec 2^k(2n-1) \prec \ldots \prec 2^k5 \prec 2^k3 \prec \ldots \\
\qquad \ldots \\
\ldots \prec 2(2n+1) \prec 2(2n-1) \prec \ldots \prec 2\cdot 5 \prec 2\cdot 3 \prec \ldots \\
\ldots \prec (2n+1) \prec (2n-1) \prec \ldots \prec 9 \prec 7 \prec 5 \prec 3.
\end{array} \tag{26.12}$$

If f is a continuous map of an interval into itself with a period p, and $q \prec p$ in this ordering, then f has a periodic point of period q.

Note that in the above ordering k can take on any value from zero to ∞. Hence, once a period-three cycle has been detected, it follows that there are infinitely many periodic orbits. This is often condensed in the statement of Li and Yorke that 'period three implies chaos'. This statement should however be interpreted with care (Medio, 1992, pp. 156-157). Let us now come to the Li and Yorke theorem proper (Casti, 1989, p. 235). In what follows $f^{(n)}(y)$ denotes the n-th iterate of f, namely, given $y_{t+1} = f(y_t)$, the second iterate is $y_{t+2} = f(y_{t+1}) = f(f(y_t)) = f^{(2)}(y_t)$ and so on up to $y_{t+n} = f^{(n)}(y_t)$.

Li-Yorke theorem
Let $f : I \to I, I = [0,1]$ be a continuous map. Assume that there exists a point $\hat{x} \in I$ such that the first three iterates of $\hat{x}$ are given by $f(\hat{x}) = b, f^{(2)}(\hat{x}) = c, f^{(3)}(\hat{x}) = d$. Assume further that these iterates satisfy the inequality

$$a \leq \hat{x} < b < c \quad (\text{or } d \geq \hat{x} > b > c).$$

Then for every $k = 1, 2, \ldots$, there is a periodic point in I having period k, and there is an uncountable set $U \subset I$ containing no periodic points, which satisfies the following conditions:

(1) for every $x, y \in U, x \neq y$

(i) $\limsup_{n\to\infty} \left| f^{(n)}(x) - f^{(n)}(y) \right| > 0,$

and

(ii) $\liminf_{n\to\infty} \left| f^{(n)}(x) - f^{(n)}(y) \right| = 0,$

(2) for every point $x \in U$ and periodic point $x_0 \in I$

$\limsup_{n\to\infty} \left| f^{(n)}(x) - f^{(n)}(x_0) \right| > 0.$

Statement (1) means that, in the aperiodic set U (also called the scrambled set), aperiodic trajectories starting from two different (but arbitrarily close) points will infinitely often become finitely separated from one another, and will equally infinitely often become very close to one another, but they will never intersect. This is practically equivalent to SDIC. Statement (2) means that if an aperiodic cycle approximates a periodic cycle for a while, it must move away from that cycle, namely the aperiodic orbits will have no point of contact with any of the periodic orbits.

Since we have aperiodic orbits and SDIC, we are in the presence of *chaos in the Li-Yorke sense* (also called topological chaos).

The conditions of the theorem are satisfied by any function f that has a point of period-3 (hence the statement 'period-three implies chaos'), but it is not necessary for f to possess such a point for chaos to emerge.

Let us now pass to n-dimensional systems. To consider an n-dimensional system let us first introduce the notion of *snap-back repeller*. A snap-back repeller is a fixed point around which there is a region such that any trajectory starting in it (even if arbitrarily close to the fixed point), at first moves *away* from the fixed point (i.e., is repelled by it), but after having left the region suddenly jumps back onto the fixed point. The mathematical conditions for a snap back repeller are the following:

1) the difference equation system

$$\mathbf{y}_{t+1} = \mathbf{f}(\mathbf{y}_t) \tag{26.13}$$

has a fixed point $\mathbf{y}_e = \mathbf{f}(\mathbf{y}_e)$ around which there is a ball $B_r(\mathbf{y}_e)$ having radius r and such that all the eigenvalues of the Jacobian of $\mathbf{f}(\mathbf{y})$ are greater than one in modulus at any point lying in the ball. This point is called an *expanding fixed point,* and is obviously unstable: any initial point lying in the ball will sooner or later go out of it. However, points outside the ball may be stable (the eigenvalues are less than unity in modulus).

2) Given an expanding fixed point, in the ball there is a point $\mathbf{y}_0 \neq \mathbf{y}_e$ such that for some positive integer k we have $\mathbf{f}^{(k)}(\mathbf{y}_0) = \mathbf{y}_e$, and the Jacobian of $\mathbf{f}^{(k)}(\mathbf{y}_0)$ is non-singular. Then $\mathbf{y}_e$ is called a snap-back repeller of $\mathbf{f}$.

We then have the following

Marotto theorem
If $\mathbf{f}$ has a snap-back repeller, then (26.13) is chaotic.

We conclude our overview of the basic theorems with Singer's theorem, whose interest lies in the fact that it under its conditions it is possible to exclude the presence of more than one stable periodic orbit in one-dimensional discrete-time maps.

Singer theorem
Let $f : I \xrightarrow{f} I$ be a three times differentiable map with the following properties:

1) f has one critical point c, namely a point such that $f'(c) = 0$, with $f'(x) > 0, x < c$, $f'(x) < 0, x > c$;

2) the origin is a repelling fixed point, namely $f(0) = 0$ and $f'(0) > 1$;

3) the Schwarzian derivative is negative at all points in I (excluding of course c).

THEN f has at most one stable periodic orbit in I.

26.2.4 Discrete time chaos in economics

There are by now numerous applications of discrete-time chaos to economic models, as shown for example by the papers collected in Anderson et al. eds., 1988, Barnett et al. eds., 1989, Benhabib ed., 1992, Creedy and Martin eds., 1994. Here we shall examine one of the pioneering contributions, showing the occurrence of chaos in descriptive growth theory (Day, 1982), and an application to exchange rate dynamics (DeGrauwe et al., 1993).

26.2.4.1 Chaos in growth theory

The neoclassical growth model has been treated at length in Chap. 13, Sect. 13.2, hence we can directly go on to its reformulation in discrete time. First, note that in discrete time the assumption of a constant rate of growth of the labour force becomes

$$L_t = L_0(1+n)^t.$$

Second, it is assumed that the capital stock lasts for exactly one period, which means a depreciation rate of 100% per period. From the definition

$$K_{t+1} - K_t = I_t - \delta K_t,$$

we have, by letting $\delta = 1$,

$$K_{t+1} = I_t.$$

Since $I_t = S_t = sY_t = sF(K_t, L_t)$, and the production function is homogeneous of the first degree, we have

$$K_{t+1} = L_t s f(k_t).$$

We now divide through by L_t and observe that $L_t = L_{t+1}/(1+n)$, so that we finally have the fundamental dynamic equation

$$k_{t+1} = \frac{s}{1+n} f(k_t). \tag{26.14}$$

In the case of the traditional production function satisfying the Inada conditions, it is easy to show that there is a unique positive equilibrium point which is asymptotically stable (see exercise 1).

Let us now modify the production technology by assuming that per capita output is produced according to the function

$$\frac{Y_t}{L_t} = Ak_t^\beta (m - k_t)^\gamma, \quad k_t \leq m, \tag{26.15}$$

where the basic term Ak_t^β is clearly a Cobb-Douglas-like component, and the term $(m - k_t)^\gamma$ was introduced by Day (1982) to represent the (negative) influence of pollution caused by an increasing k. Thus the dynamic equation (26.14) becomes

$$k_{t+1} = \frac{sA}{1+n} k_t^\beta (m - k_t)^\gamma. \tag{26.16}$$

It is then easy to show the emergence of chaos for appropriate values of the parameters. Let us, for example, set $\beta = \gamma = 1$, which yields the production function (in levels) $Y = AK(m - k)$. This should not look implausible or arbitrary. As regards the basic term, we have already met the AK production function in endogenous growth theory (see Chap. 22, Sect. 22.2.2.1). The influence of pollution can be taken linear to a first approximation.

Now, with $\beta = \gamma = 1$, Eq. (26.16) becomes a logistic map—see above, Eq. (26.1)—and hence capable of giving rise to chaos for $[sA/(1+n)] > \alpha_\infty$. Chaos is also possible when β and γ are not equal to one, and other routes to chaos exist (for example a variable saving rate): see Day (1982).

26.2.4.2 Exchange rate dynamics and chaos

We have already examined exchange-rate dynamics in the context of the rational expectations overshooting model (see Chap. 22, Sect. 22.3.2). The approach taken by DeGrauwe et al. (1993) is completely different. They start by observing that exchange rate movements are mainly determined by speculators, who—following earlier contributions by Frankel and Froot (1988)—can be divided into two categories: 'chartists' and 'fundamentalists'.

Chartists are those who base their forecasts of the future spot exchange rate on technical analysis, which is mainly a study of past exchange rates to detect patterns that can be projected in the future. Fundamentalists are those who look for fundamental determinants of the exchange rate, calculate an equilibrium exchange rate consistent with these fundamentals, and expect that the current exchange rate will gradually move towards its equilibrium value. Speculators, by buying and selling foreign exchange in the current period in relation to their expectations of next period's exchange rate, determine today's exchange rate.

Let us then start from the basic equation determining the exchange rate

$$S_t = X_t E_t(S_{t+1})^b, \tag{26.17}$$

where S_t is the current exchange rate, $E_t(S_{t+1})$ is the exchange rate that speculators expect to obtain in the next period, $0 < b < 1$ is the discount factor that speculators use to discount the expected exchange rate, and X_t can be thought of as a reduced-form equation describing the structure of the model and the exogenous variables (DeGrauwe et al., 1993, p. 73).

Given the distinction between chartists and fundamentalists, the change in the exchange rate expected by the market can be written as a weighted

average of the two groups' expectations;

$$E_t(S_{t+1})/S_{t-1} = (E_{ct}(S_{t+1})/S_{t-1})^{m_t} (E_{ft}(S_{t+1})/S_{t-1})^{1-m_t}, \tag{26.18}$$

where the subscripts c and f denote the chartists and fundamentalists respectively, and m_t is the weight given to the chartists in period t. The reason for using S_{t-1} is that, since expectations are formed at period t, the information on S_t is not yet known, hence S_{t-1} has to be used.

Chartists in practice use various extrapolative models, that can be subsumed under the general formulation

$$E_{ct}(S_{t+1})/S_{t-1} = f(S_{t-1}, S_{t-2}, ..., S_{t-N}). \tag{26.19}$$

Fundamentalists are assumed to calculate the equilibrium exchange rate S_t^* through the model. In equilibrium, $E_t(S_{t+1}) = S_t = S_{t-1}$. By imposing this in Eq. (26.17) we get

$$S_t^* = (X_t)^{1/(1-b)}. \tag{26.20}$$

Then the gradual adjustment of the current exchange rate towards its equilibrium value implies

$$E_{ft}(S_{t+1})/S_{t-1} = (S_{t-1}^*/S_{t-1})^\alpha. \tag{26.21}$$

In order to simplify the analysis, the authors (DeGrauwe at al., 1993, p. 75) set X_t as a constant and normalize it to 1. Thus we have the equation

$$S_t = \left[S_{t-1} f(S_{t-1}, S_{t-2}, ..., S_{t-N})^{m_t} (S_{t-1}^*/S_{t-1})^{\alpha(1-m_t)}\right]^b. \tag{26.22}$$

It remains to determine the weight m_t. The basic idea is that the fundamentalists have heterogeneous expectations that are normally distributed around the true equilibrium value S_{t-1}^*. This means that when the market exchange rate equals the true equilibrium value, half the fundamentalists will find that the market rate is too low, while the other half will find it too high, compared to their own estimates of the equilibrium rate. Assuming that fundamentalists have the same degree of risk aversion and the same wealth, the demand for foreign exchange by the first half will be exactly matched by the supply of foreign exchange by the second half. Hence the fundamentalists do not influence the market: their weight is zero, and $m_t = 1$.

When the market exchange rate deviates from the true equilibrium value in one sense or the other, the two groups will be numerically different (the more so the farther away is the market exchange rate from the true equilibrium exchange rate) and their demands for, and supplies of, foreign exchange will not match. Since this divergence grows greater as the market exchange rate deviates more and more from the true equilibrium exchange rate, the weight of fundamentalists is an increasing function of the deviation of the market exchange rate from the true equilibrium value.

This reasoning can be summarized in the following function

$$m_t = 1/[1 + \beta(S_{t-1} - S^*_{t-1})^2], \tag{26.23}$$

which equals one for $S_{t-1} = S^*_{t-1}$, and decreases toward zero as $(S_{t-1} - S^*_{t-1})$ increases.

The model is thus made up of the two non-linear, higher-order difference equations (26.22) and (26.23). The function $f(S_{t-1}, S_{t-2}, ..., S_{t-N})$ is so far unspecified even qualitatively. By looking at forecasting rules used in practice by chartists, the authors arrive at specific functional forms, which however are not of much help for the analytical study of the model, that remains intractable. Hence they resort to numerical simulations, that do show the onset of chaos for plausible values of the parameters.

For another analysis of exchange rate markets, where a cause of possible chaotic dynamics is seen to be the monetary authorities' intervention to stabilize the market, see Szpiro (1994).

26.3 Continuous time systems and chaos

In discrete time the points in the sequence (at equally spaced values of t) that represents the solution can jump irregularly in the phase space. In the case of differential equations the solution is represented by a continuous trajectory in the phase space, and the typical jumps of discrete time equations cannot occur. Since in planar systems a trajectory representing the solution cannot intersect itself, the behaviour of a 2×2 differential equation system is regular, and the most complex behaviour that can occur consists of limit cycles (Guckenheimer and Holmes, 1986, pp. 50 ff).

It follows that, while chaos can occur already in a single first-order non-linear difference equation, it takes a system of at least three first-order non-linear differential equations for chaos to arise.

26.3.1 The Lorenz equations, strange attractors, and chaos

It was in fact by studying the solution to a three differential equation system representing the turbulence in fluids that the meteorologist E.N. Lorenz discovered the existence of chaos. The Lorenz equations are

$$\begin{aligned} x' &= \sigma(-x + y), \\ y' &= \rho x - y - zx, \qquad \sigma, \rho, \beta > 0,\ \sigma > 1 + \beta. \\ z' &= -\beta z + xy, \end{aligned} \tag{26.24}$$

The equilibrium points are found by setting $x' = y' = z' = 0$. The resulting

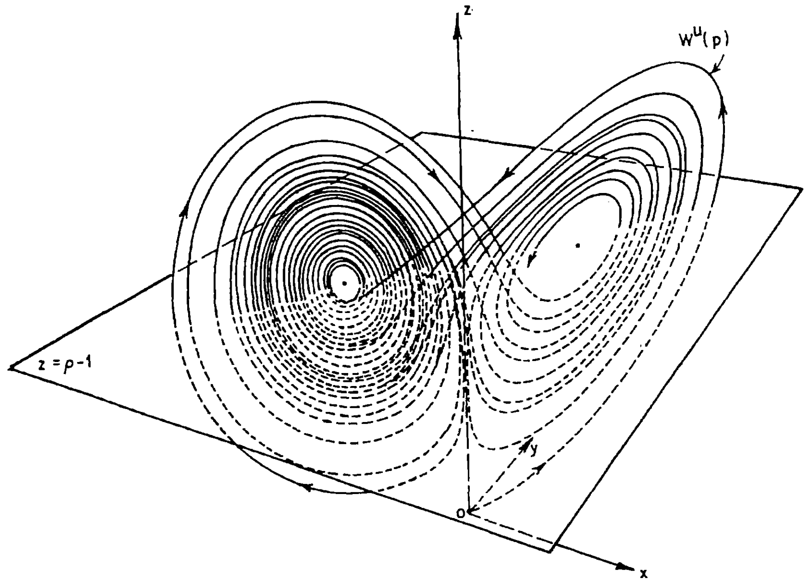


Figure 26.7: The Lorenz equations (Guckenheimer and Holmes)

system can easily be solved, and shows that the origin is the only equilibrium point for $\rho \leq 1$, while for $\rho > 1$ two additional non-trivial fixed points emerge, at

$$(x_e, y_e, z_e) = (\pm\sqrt{\beta(\rho - 1)}, \pm\sqrt{\beta(\rho - 1)}, \rho - 1). \tag{26.25}$$

It is fairly simple to examine the local behaviour of system (26.24) by using methods treated at length in previous chapters (see Chap. 21, Sect. 21.4.1, and Chap. 25, Sect. 25.2). The Jacobian of the system is

$$\begin{bmatrix} -\sigma & \sigma & 0 \\ \rho - z & -1 & -x \\ y & x & -\beta \end{bmatrix}. \tag{26.26}$$

When $\rho < 1$, the only equilibrium point (the origin) is locally asymptotically stable (see exercise 2). At $\rho = 1$ the Jacobian becomes zero at the origin, hence nothing can be said from the linear approximation. For $\rho > 1$ the two additional equilibria emerge. The origin becomes a saddle point (see exercise 3) while the two other equilibria are locally stable (see exercise 4) as long as

$$1 < \rho < \rho_h, \quad \rho_h \equiv \sigma(\sigma + \beta + 3)/(\sigma - \beta - 1). \tag{26.27}$$

At $\rho = \rho_h$, a Hopf bifurcation occurs at the non-trivial equilibrium points, where the Jacobian has one negative real root and two pure imaginary roots (see exercise 5). For $\rho > \rho_h$ the non-trivial equilibrium points are saddle points (see exercise 6), hence now all three equilibria are unstable.

One might think that the Hopf bifurcation occurring as ρ passes through ρ_h could give rise to stable periodic orbits, but this is not the case. Numerical integration of the system by E.N. Lorenz and subsequent analytical studies (see Guckenheimer and Holmes, 1986, Chap. 2, Sect. 2.3, from whom we have taken Fig. 26.7) have shown that for $\rho > \rho_h$ the solutions wander erratically in a set (in three-dimensional space) without leaving it: this set contains a *strange attractor.*

Chaotic behaviour in differential equation systems is in fact intimately related to the notion of strange attractor. We have already met attractors in the study of stability (see Chap. 21). A stable equilibrium point (for example either non-trivial equilibrium for $1 < \rho < \rho_h$) is a simple attractor. A less simple (but still 'normal') attractor is a closed curve to which all trajectories converge, for example a limit cycle possessing orbital stability (Chap. 21, Sect. 21.3.2.4). A strange attractor is a more complicated form of attractor, consisting of a set A that attracts all trajectories starting in an appropriate neighbourhood U: however, any trajectory that approaches A wanders around in the neighbourhood of A in an irregular way, giving rise to chaotic motion. More precisely, the definitions that we adopt, due to Ruelle (1979), are the following:

Strange attractor

Consider the one-parameter vector-differential equation

$$\mathbf{y}' = \mathbf{f}(\mathbf{y},\alpha), \quad \mathbf{y} \in \mathbf{R}^n. \tag{26.28}$$

A bounded set $A \subset \mathbf{R}^n$, invariant under the flow of system (26.28), is a *strange attractor* for the system if there is a set U with the following properties:

1. U is an n-dimensional neighbourhood of A;
2. for any initial point $\mathbf{y}(0) \in U$, the trajectory representing the solution $\mathbf{y}(t)$ remains in U for all $t > 0$, and $\mathbf{y}(t) \underset{t\to\infty}{\longrightarrow} A$, namely A is an invariant attracting set in the sense that any trajectory starting in U approaches and remains arbitrarily close to A for t sufficiently great ;
3. there is sensitive dependence on initial conditions when $\mathbf{y}(0) \in U$, namely two trajectories starting arbitrarily close in U will become essentially different as t increases;
4. the attractor is simple, namely it cannot be split into two or more separate attractors.

Chaos

The dynamical system (26.28) is chaotic if it possesses a strange attractor.

In the following sections we shall examine a few instances that are known to give rise to continuous-time chaos besides the Lorenz equations.

26.3.2 Other routes to continuous time chaos

26.3.2.1 The Rössler attractor

Unlike the discrete time case, in continuous time there are unfortunately no general criteria that allow to establish the presence of chaos. There are of course dynamical systems where a strange attractor is known to exist, one of which is the Rössler attractor, arising from the following system

$$\begin{aligned} x' &= -(y+z), \\ y' &= x + ay, \qquad , \quad a,b,c > 0. \\ z' &= b + z(x-c), \end{aligned} \tag{26.29}$$

Several examples of possible economic applications of the Rössler equations are given by Goodwin (1990). See also below, Sect. 26.3.4.

26.3.2.2 The Shil'nikov scenario

Transient chaos (i.e., the motions are chaotic for a while but eventually leave the chaotic region, which means that the invariant set is not attracting) is known to occur in Smale's *horseshoe map* (Guckenheimer and Holmes, 1986, Chap. 5, Sect. 5.1), whose existence can be proved for example through a theorem due to Shil'nikov (Guckenheimer and Holmes, 1986, Chap. 6, Sect. 6.5), which involves the study of homoclinic orbits (for the notion of homoclinic orbits see above, p. 349). A system which is known to satisfy the Shil'nikov scenario is the following dynamical system, where $f(y;\alpha)$ is a one-parameter family of functions:

$$\begin{aligned} y'' + ay' + y &= z, \\ z' &= f(y;\alpha). \end{aligned} \tag{26.30}$$

By differentiating the first equation with respect to time and substituting z' from the second, system (26.30) can be written as a third-order differential equation,

$$y''' + ay'' + y' = f(y;\alpha). \tag{26.31}$$

Equation (26.31) gives rise to chaotic motion for appropriate forms of the family of functions f; one of these is our old acquaintance, the logistic equation $f(y;\alpha) = \alpha y(1-y)$. H.-W. Lorenz (1989, pp. 169-174; 1993, pp. 194-200) was the first to call the economists' attention to a series of papers by

Arneodo et al. (see, for example, 1982) where Eq. (26.31) was studied, and to apply it to show the emergence of chaos by introducing non-linearities in a continuous time linear inventory cycle model of Gandolfo's (1971, Part II, Chap. 7, § 2; see also Chap. 17 in this volume, Sect. 17.2, exercise 5, and Medio, 1992, Chap. 13).

26.3.2.3 The forced oscillator

Transient chaos may also arise in the case of a forced oscillator. A forced oscillator is simply a relaxation oscillation equation of the van der Pol-Liénard type (see Chap. 24, Sect. 24.3) with a forcing term added, namely

$$y'' + f(y)\,y' + g(y) = h(t), \tag{26.32}$$

where $h(t)$ is a periodic function. When there is no forcing term, a relaxation oscillation equation can give rise to limit cycles (see Chap. 24, Sect. 24.3), but not to chaos, for the simple reason that a second-order autonomous differential equation is equivalent to a 2×2 first-order system, while we need at least a 3×3 system. With a term explicitly containing t, the equation is *non-autonomous* but can be transformed into an equivalent *3×3 autonomous* system, namely

$$\begin{aligned} y' &= x - F(y), \qquad F(y) = \int_0^y f(v)\mathrm{d}v, \\ x' &= -y + h(z), \\ z' &= 1. \end{aligned} \tag{26.33}$$

This equation may possess transient chaotic behaviour under appropriate forms of the functions f, g, h (Guckenheimer and Holmes, 1986, Chap. 2, Sect. 2.1, and Chap. 5, Sect. 5.3).

Although in economics the addition of an exogenous periodic forcing function to the Liénard equation may seem a bit *ad hoc*, Goodwin (1990, Chap. 10) strongly argues in favour of this procedure, concluding (p. 130) that '... the analysis of the consequences of forcing functions (that is, all relevant events beyond the control of the producer) on the behaviour of production is an important addition to our understanding of the extremely complex interdependence of the modern economy'. As a matter of fact, H.-W. Lorenz (1987a), by adding a trigonometric forcing function to Goodwin's (1951) nonlinear (but non-chaotic) business cycle model, shows the possible emergence of chaos.

26.3.2.4 The coupled oscillator

Finally, another dynamical system that is known possibly to produce chaos is the *coupled oscillator*. Dynamical coupling and de-coupling was already introduced in economics by Goodwin (1947), but it was only much later, after the necessary mathematical theory had been developed in the 1970s, that it was used to show the possible occurrence of chaos (H.-W. Lorenz, 1987b).

The basic idea behind dynamical coupling is fairly simple. Let us first consider two distinct 2×2 dynamical systems, say

$$\begin{aligned} \mathbf{x}' &= \mathbf{g}(\mathbf{x}), \\ \mathbf{z}' &= \mathbf{h}(\mathbf{z}), \end{aligned} \tag{26.34}$$

where both $\mathbf{x}$ and $\mathbf{z}$ have two elements, say (x_1, x_2) and (z_1, z_2). Also suppose that each system gives rise to a closed orbit in the phase plane, hence a constant oscillation. This may for example arise through a Hopf bifurcation (see Chap. 25, Sect. 25.2.2)

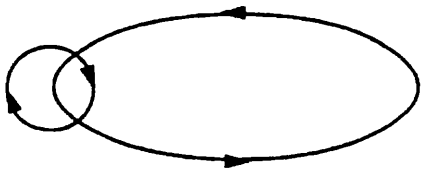
Although the two oscillators are by definition decoupled, i.e. distinct, nothing prevents us from putting the four variables together and describe the behaviour of the 4×4 system as the movement of a point $[\mathbf{x}(t), \mathbf{z}(\mathbf{t})]$ in four-dimensional space. We know that a 2×2 system cannot give rise to chaotic motion, and neither can the decoupled 4×4 system. This implies that the trajectory of the representative point may be very complicated, but will not exhibit sensitive dependence on initial conditions (hence the position of the point can be quite well predicted).

The two basic oscillations may have the same frequency or different frequencies. In the second case a further distinction is made according as to whether the ratio of the frequencies is a rational or an irrational number. When this ratio is an irrational number, the overall oscillation is called a *quasi-periodic* oscillation.

Without loss of generality we can take the ratio such that the greater frequency is at the numerator. Then, when the result is a rational number, the overall oscillation is periodic, namely the system will pass again through its initial point after a period of time equal to this number. On the contrary, when the oscillation is quasi-periodic the system will never meet its initial point again (oscillations which are neither periodic nor quasi-periodic are called nonperiodic or aperiodic; these cannot occur in the case under consideration).

All orbits describing the overall oscillation will lie on a geometric object that is called a two-dimensional *torus* or $\mathbf{T}^2$. The simplest way to imagine a torus is an inflated inner tube of a tyre (see Fig. 26.8). Take one of the two 2×2 systems, which by assumption gives rise to a closed orbit in the plane, for example a circle. The second oscillation (due to the other 2×2 system) rotates the circle around the z-axis to form the torus.

Note that the dimension of the torus is equal to the number of frequencies involved. In our case we have two frequencies, hence the torus is two-dimensional (with three frequencies—this requires three 2×2 systems—we would have a three-dimensional torus $\mathbf{T}^3$, and so on). Notwithstanding its



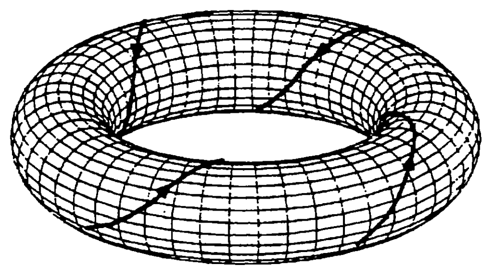


Figure 26.8: A two-dimensional torus in three-dimensional space

3-D representation in Fig. 26.8, it is two-dimensional in the sense that it can in principle be built by appropriately bending and gluing a two-dimensional object (a sheet of rubber in our example). Any trajectory on this torus is given by the composition of the oscillation in the vertical direction (due to one of the two 2×2 systems) and the oscillation in the horizontal direction (due to the other 2×2 system).

Let us now suppose that we couple the oscillators, namely we introduce the assumption that $\mathbf{x}'$ also depends on $\mathbf{z}$, and $\mathbf{z}'$ also depends on $\mathbf{x}$. This can be done in several ways, for example by adding appropriate functions:

$$\begin{aligned} \mathbf{x}' &= \mathbf{g}(\mathbf{x}) + \mathbf{P}(\mathbf{z}), \\ \mathbf{z}' &= \mathbf{h}(\mathbf{z}) + \mathbf{Q}(\mathbf{x}), \end{aligned} \tag{26.35}$$

where $\mathbf{P}$ and $\mathbf{Q}$ are the coupling elements. More generally, we can write

$$\begin{aligned} \mathbf{x}' &= \boldsymbol{\gamma}(\mathbf{x}, \mathbf{z}), \\ \mathbf{z}' &= \boldsymbol{\chi}(\mathbf{z}, \mathbf{x}). \end{aligned} \tag{26.36}$$

The motion of system (26.36) will still take place on $\mathbf{T}^2$.

Dynamical de-coupling is the opposite operation, that is to say we eliminate the dependence of $\mathbf{x}'$ on $\mathbf{z}$ and the dependence of $\mathbf{z}'$ on $\mathbf{x}$ from system

(26.35) or (26.36) as the case may be. Partial decoupling may also be carried out, by eliminating *either* the dependence of $\mathbf{x}'$ on $\mathbf{z}$ *or* the dependence of $\mathbf{z}'$ on $\mathbf{x}$. This is equivalent to unilateral coupling, i.e., we start from (26.34) and introduce either the dependence of $\mathbf{z}'$ on $\mathbf{x}$ *or* the dependence of $\mathbf{x}'$ on $\mathbf{z}$.

We can have as many coupled systems as we wish, each of order two or higher (so as to be able to give rise to Hopf bifurcations leading to closed orbits), namely

$$\begin{aligned} \mathbf{y}_1' &= \mathbf{f}_1(\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_m, \alpha), \\ \mathbf{y}_2' &= \mathbf{f}_2(\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_m, \alpha), \\ \dots & \quad \dots \\ \mathbf{y}_m' &= \mathbf{f}_m(\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_m, \alpha), \end{aligned} \tag{26.37}$$

where α is a parameter. Suppose also that for α in a certain range system (26.37) possesses a fixed point.

System (26.37) can give rise to the following scenario, in which as α increases there is a Hopf bifurcation (for this to occur we must of course assume that the conditions of the Hopf bifurcation theorem are satisfied, see Chap. 25, Sect. 25.2.2). Further increases in α may cause a second Hopf bifurcation, with the appearance of a second cycle. Note that the bifurcation now does not take place from a fixed point, but from a closed orbit (see Chap. 25, p. 479). Let us assume that there is indeed a second Hopf bifurcation, which puts the system on a two-dimensional torus.

As α continues to increase there may be a third Hopf bifurcation, and so on. If system (26.37) is sufficiently high dimensional, we may have a great number of Hopf bifurcations, say n, that put the system on an n-dimensional torus. This kind of scenario[1] (called the *Landau* scenario) can give rise to turbulence and complicated motions on $\mathbf{T}^n$ but not to chaos; in particular, there will be no SDIC.

Another more interesting scenario is the *Ruelle-Takens scenario,* in which already after the third Hopf bifurcation, that puts the system on a $\mathbf{T}^3$ torus, there may be the onset of chaos. More precisely, we state the Newhouse-Ruelle-Takens (1978) theorem:

Newhouse-Ruelle-Takens theorem

Let $\mathbf{y} = (y_1, y_2, ..., y_m)$ be a constant vector field on the torus $\mathbf{T}^m$.

If $m = 3$, in every C^2 neighbourhood of $\mathbf{y}$ there exists an open vector field with a strange attractor.

If $m \geq 4$, in every C^∞ neighbourhood of $\mathbf{y}$ there exists an open vector field with a strange attractor.

[1]When a control parameter is changed, a system may pass through several instabilities. Such a sequence is called a *route*, especially when it leads to turbulence or chaos. One then speaks of a route to turbulence or route to chaos. The theoretical discussion of a route is usually referred to as a *scenario* or picture.

Since a three-dimensional torus requires three frequencies, we must have at least three coupled two-dimensional oscillators for chaos possibly to arise in this context. Note that two coupled three-dimensional systems will not do, for a three-dimensional system cannot have more than one closed orbit.

26.3.3 International trade as the source of chaos.

One of the most obvious ways of coupling dynamical economic systems is through international trade. The possible onset of chaos in international trade models was shown by Puu (1987, 1993) by using the forced oscillator (see the previous section), but his analysis was limited by the necessity of assuming one-directional trade to remain within the mathematical context chosen.

Hence we follow the more general analysis of Lorenz (1987b), who takes into consideration three two-dimensional economic systems represented by simple IS-LM dynamic equations, which are

$$\begin{aligned} Y_i' &= \alpha_i[I_i(Y_i,r_i) - S_i(Y_i,r_i)], \quad \alpha_i > 0, \\ & \qquad\qquad\qquad\qquad\qquad\quad i=1,2,3 \\ r_i' &= \beta_i[L_i(Y_i,r_i) - \overline{M}_i], \qquad\quad \beta_i > 0, \end{aligned} \tag{26.38}$$

where Y is real income, I real investment, S real saving, i the interest rate, L demand for money, and $\overline{M}$ the exogenously given real money supply. Prices are fixed.

The dynamic equations express the usual adjustment of output to excess demand for goods, and of the interest rate to excess demand for money.

Let us suppose that each economy has an equilibrium point (Y_{ei}, r_{ei}), which is unstable and such as to give rise to oscillations. We know from our previous analysis of Kaldor's business cycle model (see Chap. 25, Sect. 25.2.4) that it is easy to plausibly specify non-linear functions for I and S so that a Hopf bifurcation occurs. We thus have three independent oscillators, whose overall motion is a (non-chaotic) motion on a three-dimensional torus $\mathbf{T}^3$.

International trade gives us the required coupling. We assume fixed exchange rates and no capital movements, so that we can concentrate on the trade balance. Let $IMP_{ji}(Y_i), j \neq i$, be the imports of country i from country j, and $IMP_i(Y_i)$ be total imports of country i, where

$$IMP_i(Y_i) = \sum_{\substack{j=1 \\ j\neq i}}^{3} IMP_{ji}(Y_i). \tag{26.39}$$

In the short run we may take the shares $IMP_{ji}(Y_i)/IMP_i(Y_i)$ as given. Since in this context the exports of a country are determined by the other countries' import demand for its goods (see Chap. 10, Sect. 10.2), we have

$$X_i = \sum_{\substack{k=1 \\ k\neq i}}^{3} IMP_{ik}(Y_k) \equiv X_i(Y_h, Y_k), \qquad i,h,k = 1,2,3, \quad h,k \neq i. \tag{26.40}$$

By using Eqs. (26.39) and (26.40), system (26.38) becomes

$$\begin{aligned} Y_i' &= \alpha_i[I_i(Y_i,r_i) - S_i(Y_i,r_i) + X_i(Y_h,Y_k) - IMP_i(Y_i)], \quad \alpha_i > 0, \\ & \quad i=1,2,3 \\ r_i' &= \beta_i[L_i(Y_i,r_i) - \overline{M}_i], \qquad \beta_i > 0, \end{aligned} \tag{26.41}$$

where the monetary authorities are assumed to sterilize the balance of payments, so that the money supply remains exogenous (this would be untenable in the presence of capital movements and perfect capital mobility, but in our context it is perfectly plausible).

System (26.41) is a system of three coupled non-linear oscillators that can be taken as the perturbation of the motion of the three independent economies on $\mathbf{T}^3$. Application of the Newhouse-Ruelle-Takens theorem immediately shows the possible presence of chaos. We say 'possible' because nothing ensures the transition of the system to the newly arising chaotic region.

26.3.4 A chaotic growth cycle

Goodwin (1990, Chap. 8; 1991) revisited his famous growth cycle model (Goodwin, 1965) and modified it by considering a linear approximation at the equilibrium point and by introducing a non-linear compensatory fiscal policy. He also introduced an additional term to avoid the criticism that the original model (as all Lotka-Volterra systems) is structurally unstable, but we shall neglect this because it looks a bit *ad hoc* (and in any case it is inessential for the onset of chaos).

Since the 1965 model has been treated at length in a previous chapter (see Chap. 24, Sect. 24.4.3), we can briefly recall the basic equations, which are

$$\begin{aligned} v' &= (a_1 - b_1 u)\, v, \\ u' &= -(a_2 - b_2 v)\, u, \end{aligned} \tag{26.42}$$

where v is the ratio of employment to the labour force, and u is the share of wages in gross product. If we take a linear approximation (see Chap. 24, Sect 24.5, exercise 5) we have

$$\begin{aligned} v' &= -du, \\ u' &= hv, \end{aligned} \tag{26.43}$$

where d, u are positive constants, and both variables are now measured in deviations from the fixed point equilibrium, that Goodwin takes, for illustration, as 90% of full employment.

The third variable to be introduced is a dynamical control variable z, representing compensatory fiscal policy (the government budget surplus). In a dynamic context, policy instruments are changed in response to disequilibria in the targets (see Chap. 19, Sect. 19.5, exercise 5, for a linear formulation). Here the target is employment. Letting c denote the target level of the deviations of employment and adopting a non-linear formulation, we have

$$z' = b + gz(v - c), \tag{26.44}$$

where b, g are positive constants (b can be taken for example as an inherent trend in the budget deficit), and z being negatively proportional to output and employment (remember that z is the budget surplus). Equation (26.44) shows that when employment falls short of its target level c, the budget surplus is decreased and vice versa.

The presence of the government budget surplus has of course an effect on output and employment, hence an appropriate term has to be added to the first equation in (26.43).

Thus we end up with the system

$$\begin{aligned} v' &= -du - ez, \quad e > 0, \\ u' &= hv, \\ z' &= b + gz(v - c). \end{aligned} \tag{26.45}$$

By an appropriate choice of units we can write this system as

$$\begin{aligned} v' &= -(u + z), \\ u' &= hv, \\ z' &= b + z(v - c), \end{aligned} \tag{26.46}$$

that is clearly a case of the Rössler attractor (see above, Sect. 26.3.2.1). As Goodwin (1991, p. 427) points out, 'The consequence of dynamical as opposed to statical control is dramatic: a stable, bounded region is defined in which chaotic motion occurs [...] It is important to note that there is only *one* non-linearity and that it is defined by full employment, thus embodying the original Harrod problem'.

Further elaboration of the model by introducing structural change and growth is contained in the cited works of Goodwin's.

26.4 Significance and detection of chaos: Stochastic dynamics or chaos?

This book is exclusively concerned with deterministic dynamics. Stochastic dynamics would require a book on its own (see, for example, Malliaris and Brock, 1982), and is mentioned here only because of chaos.

One sometimes reads expressions like 'stochastic chaos and deterministic chaos', where 'deterministic chaos' is simply chaos, while 'stochastic chaos' is standard stochastic dynamics, namely random behaviour coming from a non-chaotic but stochastic system. One reason for this terminology is that both may be characterised by their correlation dimension (which is a measure of complexity: see Brock et al, 1991, pp. 15-17). Now, the correlation dimension is a finite number in the case of (deterministic) chaos, while an independent and identically distributed stochastic process (with a non-degenerate distribution function) has correlation dimension equal to infinity. Thus in principle the correlation dimension can be used to distinguish true stochastic processes from deterministic chaos (which in turn may be low-dimensional or high-dimensional).

Since the terminology 'stochastic chaos vs deterministic chaos' may be confusing when one is not in the context of correlation dimension analysis, we shall speak of stochastic dynamics vs chaos.

Chaos theory has generated a lot of excitement and important results in physics and some other fields. Can we say the same as regards economics? Economists are well known for drawing heavily from physics (Mirowski, 1990, 1992) but, in general, a successful theory in one field is not automatically a serious theory in another field. The excitement does exist in economics, as shown by the numerous papers (collections are contained for example in Anderson et al. eds., 1988, Grandmont ed., 1988, Barnett et al. eds., 1989, 1996, Benhabib ed., 1992, Day and Chen eds., 1993, Creedy and Martin eds., 1994, Leydesdorff and Van den Besselaar eds., 1994) and the several books (see, e.g., Brock et al., 1991, Chiarella, 1990, Day, 1994, H.-W. Lorenz, 1993, Medio, 1992, Peters, 1991, 1994, Rosser, 1991) on chaotic economic dynamics.

These show that chaos theory in economics is not a fad; they also show that, from the theoretical point of view, plausible economic models can be built, and old economic models can be revisited, in which chaotic behaviour is present, although sometimes the assumptions may look a bit *ad hoc* (on this point see for example Sordi, 1993, and Nusse and Hommes, 1990). We know that *periodic* oscillations can arise in standard non-linear models (see the previous two chapters), but before the advent of chaos to explain the aperiodicity often observed in actual economic variables it was necessary to rely on an unexplained *exogenous* random variable. Chaos, on the contrary, gives us an *endogenous* explanation of erraticity. As Goodwin (1991, p. 425) aptly put it, 'Poincaré generalised an equilibrium point to an equilibrium motion; a chaotic attractor generalises the motion to a bounded equilibrium region towards which all motions tend, or within which all motions remain; the conception of equilibrium is more or less lost since all degrees of aperiodic, or erratic fluctuations can occur within the region. *The special relevance of this to economics is that it offers not* one *but* two *types of explanation of the pervasive irregularity of economic time-series—an endogenous one in addition to the conventional exogenous shock*' (emphasis added).

Let us now come to three other major theoretical implications of chaos. One is that the rational expectations hypothesis is untenable in the face of chaos (Chiarella, 1990, pp. 124-125; Kelsey, 1988, pp. 682-683; Medio, 1993, pp. 17-18). It should be stressed that this is *not* a criticism to rational expectations of the type: it is practically impossible that all economic agents have perfect information etc. as required by REH, hence they must rely on other processes such as bounded rationality (Sargent, 1993). Such a criticism would be an *external criticism,* i.e., the assumptions cannot obtain in reality, hence we must drop REH (but if the assumptions obtained, REH would be all right). The criticism coming from chaos theory is an *internal* criticism, all the more destructive because it shows that REH is untenable even when its assumptions obtain.

In fact, if the true model is chaotic, economic agents—assumed to have perfect information including knowledge of the model (exactly as a physicist knows the equations governing a certain phenomenon)—cannot conceivably achieve the infinite precision required to avoid the devastating effects of sensitive dependence on initial conditions. Perfect deterministic foresight *out of steady states* (Grandmont, 1985, Sect. 3) would be impossible in economics as it is in physics when the model is chaotic. The situation would become worse if, in addition to chaos, stochastic elements were also present. Even if economic agents knew the stochastic process driving the exogenous shocks, the presence of SDIC on the deterministic part would make it impossible for them to calculate the objective probability distribution of outcomes. Hence it would not be possible to verify the essential properties (see Chap. 22, Sect. 22.3.1) of rational expectations in a stochastic context. In both cases (deterministic and stochastic), the economic theorist would have to abandon the assumption of rational expectations and rely on other rules for expectation formation. Heiner (1989) has suggested a form of adaptive expectations. From the theoretical point of view, rules for expectation formation *should be consistent with the underlying chaotic model,* but *general* rules of this type have not yet been devised (Medio, 1992, p. 18). In the meantime, rules of thumb used by practical agents might have to be taken into account, with the proviso that these rules have a sense only in the very short run.

A second implication of chaos concerns the use of econometric models in forecasting (Baumol and Quandt, 1985). Estimated parameters in econometric models have a confidence interval—which means that the 'true' value may be anywhere within this interval with the assumed probability. But even if this confidence interval could be shrunk almost to zero (which is practically impossible), it takes no econometrician to understand that—if the 'true' model is chaotic—the presence of SDIC implies the impossibility of forecasting except for maybe the very short run (on short-run predictability in chaotic models see also H.-W. Lorenz, 1993, Chap. 6, Sect. 6.4).

A third implication of chaos is the *irreversibility of time* **in theory**. This can easily be seen by considering unimodal maps (see Sect. 26.1.1). As pointed out by Barnett and Chen (1988, p. 203), the existence of a turning point in the map f makes f non-invertible because the inverse of f is set-valued. While f is a function, the inverse of f is a correspondence. This means that, while 'normal' equations can in principle be integrated either forward or backward in time, only forward integration is possible here.

We would like again to stress that this is time irreversibility *in theory,* unlike time irreversibility 'in practice', which occurs in 'normal' dissipative systems that can be in principle integrated either forward or backward in time but in practice do not allow a correct 'prediction' of the past (on this point see Chap. 24, Sect. 24.4.2).

This as regards the theory. In physics and related disciplines, apart from purely theoretical interest, important results in the study and explanation of real phenomena have been obtained. Hence the economist with a more applied bent would undoubtedly ask to be shown that chaotic models give *better* explanations of *real* economic phenomena than *non-linear non-chaotic stochastic* models.

Let us stress that we are *not* suggesting to compare chaotic models (which are necessarily non-linear) with linear stochastic models, but with non-linear dynamic models that are non-chaotic but stochastic. There are by now quite a few models of this last type around: for one such model specified as a set of 24 non-linear stochastic differential equations and estimated in continuous time, see Gandolfo and Padoan (1990) and Gandolfo et al. (1996); see also Wymer (1996). This model turns out to possess a structurally stable steady state, and to give good explanation and forecasts of the Italian economy.

However, the distinction between *endogenous* aperiodic behaviour (coming from a non-stochastic chaotic model) and *exogenous* aperiodic behaviour (coming from a non-chaotic model with exogenous stochastic disturbances) is slippery in applied work, since—generally speaking—it depends on the size of the model. A small model takes all the rest as exogenous, and it might be unwarranted to assume that the exogenous rest has no influence on the endogenous variables. Since there are a large number of exogenous factors at work, to a first approximation it may be reasonable to assume a stochastic influence (which is of course the standard justification for adding a stochastic disturbance term in econometric models). When the small model is considered as part of a larger model, some of the exogenous influences may be taken in as endogenous variables. This has induced some to argue that the distinction is not only slippery, but meaningless: 'Whether fluctuations are endogenously or exogenously generated, stochastic or deterministic, is a property of a model, not of the real world. Only if there were a true model in much more precise correspondence with the real world than are macroeconomic models might it be a useful shorthand to speak of the actual business cycle as being "stochastic" or "deterministic" (Sims, 1994, p. 1886)'.

This shows that the main problem in applied economics is one of empirical *detection*, namely whether actual economic data show *evidence of chaos as* *distinct from the behaviour deriving from a non-linear non-chaotic stochastic system.* Various tests have been developed (the basic references are Brock et al., 1991, and Pesaran and Potter eds., 1992), but it should be emphasized that it is not enough to show the presence of non-linearity in the data. Non-linearity is a necessary, but by no means a sufficient condition for chaos, and can at best show that *linear* stochastic models are not suitable: it cannot discriminate between non-linear stochastic and non-linear deterministic random behaviour, which is what we are concerned with.

Specific tests to discriminate between chaotic systems and non-linear non-chaotic stochastic systems, such as the correlation integral, do exist, but they do not work very well in uncovering low-dimensional chaos when stochastic white noise is present (Mirowski, 1990; Liu et al., 1992; Chen, 1993; Granger, 1994). Real economic data usually show non-linearity but generally fail to exhibit low-dimensional chaos (Day, 1994; Liu et al., 1992; Granger, 1994); this is the case also when one has enough observations (one of the problems with these tests is that they require several thousand observations to be reliable) as when testing chaos on daily exchange rates (DeGrauwe et al., 1993).

Thus, for the moment the empirical evidence seems to point to the abandonment of *linear* stochastic models in favour of *non-linear* (but non-chaotic) stochastic models rather than in favour of chaotic models. Hence, until new cogent empirical evidence is presented, we feel that—though chaotic dynamics is an important and welcome addition to the dynamic economist's tool kit—it is a little soon to declare its undisputed pre-eminence in economics. On this, the jury is still out.

26.5 Other approaches

26.5.1 Introduction

Chaotic dynamics, though certainly the most important subset of complex economic dynamics, does not exhaust it. Among the various other approaches we mention just two: catastrophe theory and synergetics. The former because in the past it had a brief spurt of popularity among economists. The latter because, though yet scarcely employed in economics, in the future might offer useful insight to economists. On the basis of Occam's razor (*pluralitas non est ponenda sine necessitate ponendi,* or, one must not introduce more concepts than are strictly required) we shall only briefly deal with these two approaches.

26.5.2 Fast and slow, and synergetics

By a 'fast' variable we mean a variable that adjusts very rapidly to its (partial) equilibrium value following any disturbance. By contrast, 'slow' variables adjust only slowly.

This can be easily understood in terms of the notion of partial adjustment equations introduced in Chap. 12, Sect. 12.4. Let us recall that, in the simplest case, we have

$$y'(t) = \alpha[x(t) - y(t)], \tag{26.47}$$

where $x(t)$ can be interpreted as the desired or potential value of $y(t)$, and $\alpha > 0$ is the speed of adjustment of the actual to the desired value. Since α can be interpreted as the reciprocal of the mean time-lag, as $\alpha \to \infty$ the mean time-lag tends to zero, which means an instantaneous adjustment of $y(t)$ to $x(t)$ (for details see Sect. 12.4).

The distinction between 'fast' and 'slow' variables is, explicitly or implicitly, at the base of much theorizing and debate in economics. An orthodox Keynesian, for example, would say that quantities adjust much more rapidly than prices. An orthodox monetarist would say the contrary. An orthodox new classical macroeconomist would say that all the relevant variables adjust very rapidly, so that markets can be taken as being in equilibrium. In international economics, the free-market exchange rate is determined by the demand for and supply of foreign exchange, which come from both commercial and financial transactions. In a free market with perfect capital mobility, financial assets and related prices adjust much more rapidly than commodity trade, hence the exchange rate is principally determined by excess demand for (net) foreign assets. In general, variables which adjust much more rapidly than others can be taken as being practically always in equilibrium, and this is not irrelevant from the point of view of the stability of the system (for a treatment of this point in the context of linear systems see Schoonbeek, 1995).

What is important to observe in general, is that the evolution over time of a system where both fast and slow variables are present, is mainly determined by the *slow* variables. Slow variables are, in a sense, more fundamental than fast variables. This is the essence of Haken's (1983a, 1983b) *slaving principle.* The slow variables are the slaving or *order* variables (also called *order parameters* by Haken), the fast variables are the slaved variables. Variables are also spoken of as 'modes', a terminology adapted from physics.

The workings of the slaving principle can be shown quite simply by the following elementary example. Consider the system

$$\begin{aligned} y_1' &= \alpha_1(\hat{y}_1 - y_1), \\ y_2' &= \alpha_2(\hat{y}_2 - y_2), \\ &\alpha_1, \alpha_2 > 0. \end{aligned} \tag{26.48}$$

The two partial adjustment equations are apparently distinct, but this is not so if we introduce the dependence of the desired values, $\hat{y}_1$ and $\hat{y}_2$, on the actual values, namely

$$\begin{aligned}\hat{y}_1 &= g_1(y_1, y_2),\\ \hat{y}_2 &= g_2(y_1, y_2).\end{aligned}$$

Thus we have

$$\begin{aligned}y_1' &= \alpha_1[g_1(y_1, y_2) - y_1] = -\alpha_1 y_1 + \alpha_1 g_1(y_1, y_2),\\ y_2' &= \alpha_2[g_2(y_1, y_2) - y_2] = -\alpha_2 y_2 + \alpha_2 g_2(y_1, y_2).\end{aligned} \tag{26.49}$$

To link the two equations it is of course sufficient that $\hat{y}_1$ depends on y_2 and that $\hat{y}_2$ depends on y_1. For example, y_1 might be the quantity actually demanded, and $\hat{y}_1$ the desired or optimal demand (which depends on price through the demand schedule); the first equation would describe the partial adjustment of consumers toward the demand schedule. Then y_2 might be the supply price, that adjusts to its desired value which depends on the actual quantity sold.

Be it as it may, suppose that one of the two variables is fast and the other slow. Let for example $\alpha_1 \to \infty$. Then the adjustment is almost instantaneous (see Sect. 12.4), which means that we can take y_1 as practically equal to $\hat{y}_1$ at all times. Hence we have $y_1 - g_1(y_1, y_2) = 0$ from which—assuming that the conditions of the implicit function theorem are satisfied—we get

$$y_1 = h_1(y_2), \tag{26.50}$$

where h_1 is a continuously differentiable function. Substitution of (26.50) into the second equation of system (26.49) leaves us with the single differential equation in the unknown function y_2:

$$y_2' = \alpha_2[g_2(h_1(y_2), y_2) - y_2] = \alpha_2\varphi(y_2). \tag{26.51}$$

Thus the dynamics of the system is entirely determined by y_2, the slow variable; by Eq. (26.50), the dynamic behaviour of y_1 is entirely determined by that of y_2. This latter variable 'slaves' y_1, the 'slaved' variable.

The same result could be obtained if α_1 is a positive finite number and $\alpha_2 \to 0$.

This simple example can be extended to n-dimensional systems of the type

$$\begin{aligned}y_1' &= -\alpha_1 y_1 + g_1(y_1, y_2, ..., y_n),\\ y_2' &= -\alpha_2 y_2 + g_2(y_1, y_2, ..., y_n),\\ \dots & \quad \dots\\ y_n' &= -\alpha_n y_n + g_n(y_1, y_2, ..., y_n),\end{aligned} \tag{26.52}$$

where the speeds of adjustment can be arranged in two subsets, $\alpha_i, i = 1, 2, ..., m$ and $\alpha_s, s = m+1, ..., n$, such that

$$\alpha_i \to 0, \quad \alpha_s > 0 \text{ but finite.} \tag{26.53}$$

It is understood that the functions $g_j, j = 1, 2, ..., n$, are non-linear functions of $y_1, y_2, ..., y_n$ only, with no constant or linear terms. Thus they can be neglected to a first approximation and, by considering the linear part of system (26.52), we see that the system is locally stable since all the α's are positive. More generally, it is sufficient that $\alpha_s > 0$, while α_i might also be non positive (Haken, 1983a, pp. 195-196), but this does not interest us since we are dealing with adjustment speeds that are by definition positive.

In such systems we can apply Haken's *adiabatic approximation*, that consists in treating the faster variables as approximately in equilibrium, solving for them in terms of the slower variables, and substituting the result in the differential equations that determine the movement of the slower variables, thus greatly reducing the dimensionality of the system. Given the difficulty of analysing non-linear systems, this is an important simplification.

As shown by Haken, the adiabatic approximation can also be applied to non-linear systems that are not in the form (26.52) but in the general form

$$\begin{aligned}y_1' &= f_1(y_1, y_2, ..., y_n),\\ y_2' &= f_2(y_1, y_2, ..., y_n),\\ \dots & \quad \dots\\ y_n' &= f_n(y_1, y_2, ..., y_n).\end{aligned} \tag{26.54}$$

Unfortunately in this general case the adiabatic approximation procedure is valid only under certain conditions, which require that most real eigenvalues of the Jacobian of system (26.54) should be negative (with only a few positive), and that all complex eigenvalue with positive real part should have imaginary parts much smaller than the real parts of the stable eigenvalues. These conditions cannot possibly be checked algebraically when the order of the system is high, hence one has to work with numerical simulations.

Once we have reduced the dimensionality of the system through the adiabatic approximation, we can analyse the resulting low-dimensional system through the standard techniques for non-linear analysis. What is important to stress is that *the dynamic behaviour of the overall system will be in any case governed by the behaviour of the order variables.*

Suppose that changes in external conditions, for example a change in a parameter (in the strict sense), cause a qualitative change in an eigenvalue of the reduced system which undergoes a bifurcation of some type. The order variables will move into the new dynamic state *and drag the slaved variables with them into this new state.*

It would, however, be reductive to identify synergetics with the adiabatic approximation. Haken (1983a, 1983b) proposed synergetics as a general theory of dynamic systems composed of many subsystems in which non-deliberate cooperative interaction takes place so as to produce a macroscopic behaviour of a self-organized nature (including nonequilibrium phase transitions and instability hierarchies). In this direction synergetics might offer an answer to the moot question of the microfoundations of macroeconomics (Koblo, 1991). It is also important to note (Haken, 1983a, p. 343) that *chaotic motion occurs when the slaving principle fails and the formerly stable mode can no longer be slaved but is destabilized.* One may think of this as a situation in which a formerly slaved mode 'revolts' and becomes an order parameter (Diener and Poston, 1984). Thus chaotic dynamics may be associated with a self-organised restructuring of the system and the emergence of a new order. In this sense it is a view similar to Prigogine's 'order through chaos' (Prigogine, 1980; for a survey of the possible economic applications see Foster, 1994). Haken's synergetics, like Prigogine's order through chaos, purports to be an encompassing scientific approach, whose examination lies outside the scope of the present book.

The adiabatic approximation was used by Medio (1984) to analyse a dynamic input-output model and by Silverberg (1984) to study embodied technical progress in a dynamic model. More ambitious attempts to apply the whole *corpus* of synergetics to economics are Koblo (1991) and Zhang (1991). Related approaches are the approach suggested by Aoki (1994), who uses the idea of hierarchical dynamics, and that suggested by Scheinkman and Woodford (1994), who use the idea of self-organisation.

26.5.3 Catastrophe theory

Scientists, in addition to their specialistic jargon, sometimes take words from the ordinary language and use them in a technical meaning that does no longer reflect the ordinary meaning (or reflects it only in part). When the layman reads these words, confusion is bound to arise. No layman, for example, would be confused by the expression *homoclinic orbit*: he would simply not understand it. But he would be confused by words like *chaos* or *catastrophes* if he interprets them in the ordinary meaning.

Thus catastrophe theory is not necessarily concerned with catastrophic events—it simply deals with situations in which a variable suddenly jumps from one state to another. In this sense it is closely related to bifurcation theory (see the previous chapter) so much so that after its inception, due to Thom (1972), a lot of controversy took place among mathematicians on whether the label *theory* was appropriate at all.

One of the appealing features of catastrophe theory is that, while the mathematics is quite complicated, simple applications can be made with no mathematics at all by relatively simple diagrams, as shown by the inflation model of Woodcock and Davis (1978, pp. 117-119), that is one of the first applications of catastrophe theory to economics (the first in absolute seems to be a model of stock market crashes by Zeeman, 1974). In the study of the inflationary process, the *expected* inflation rate plays a prominent role together with the unemployment rate. The actual inflation rate depends directly on the expected inflation rate, for expectations of higher inflation in

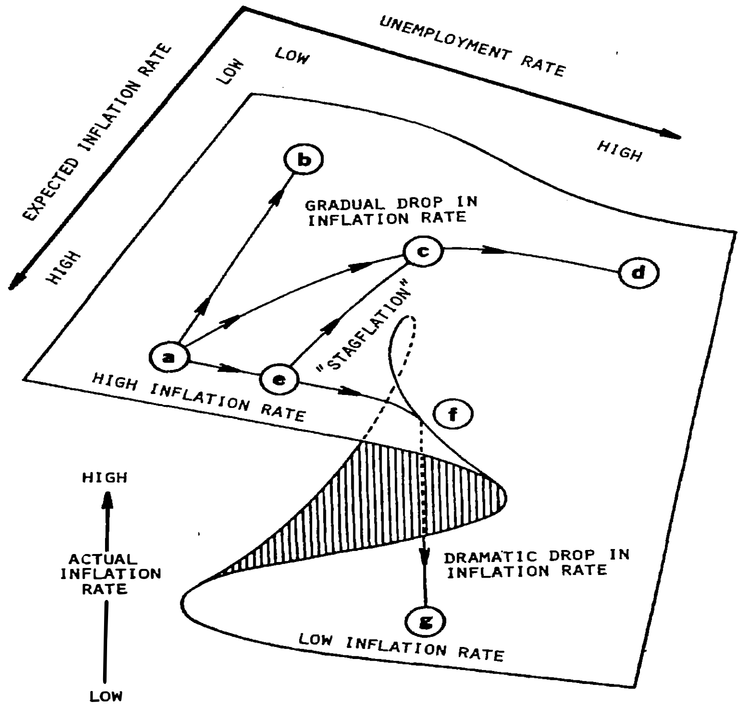


Figure 26.9: A catastrophe in inflation (Woodcock and Davis)

the future will lead to a current behaviour that increases the actual inflation rate. The unemployment rate has an inverse effect, for a higher rate of unemployment means lower wage pressure and hence lower current inflation.

The worst situation (as regards inflation) is undoubtedly one of high expected inflation rate and low unemployment rate (point **a** in Fig. 26.9, taken from Woodcock and Davis, 1978, p. 118). It may be improved in several ways:

1) by acting on expectations. For example, the government may adopt an aggressive policy of discouraging price increases. This will bring the system to point **b**, at no cost in terms of additional unemployment.

2) by acting on both expectations and the unemployment rate it will be possible to achieve a higher reduction in inflation (path **a-c-d**), at some cost in terms of additional unemployment.

3) by acting on the unemployment rate alone (path **a-e**). This, however, produces only a slight improvement in the inflation rate and the economy

falls into stagflation. To move from this region to one of lower inflation it is also necessary to act on inflationary expectations and, simultaneously, cause a further slight increase in the unemployment rate (path **e-c-d**).

4) by a drastic increase in the unemployment rate. This brings the system from **e** towards the critical point **f**, where the discontinuity in the dynamics will cause the point to jump to **g**, with a dramatic drop in the inflation rate.

This diagram, with the sudden transition from point **f** to point **g**, is an example of one of Thom's seven 'elementary' catastrophes, the *cusp catastrophe*. The cusp catastrophe has indeed been the one generally used in subsequent economic applications, including a revisitation of Kaldor's trade cycle model by Varian (1979).

Catastrophe theory had a fad in economics in the late 1970s and early 1980s, after which interest and research in this field dwindled away. One reason for the disappearance of catastrophe theory from the dynamic economists' tool kit might be that, as soon as one leaves popular presentations and goes into the mathematics, one sees that the notion of potential, which catastrophe theory starts from, is fundamental in physics but usually has no economic meaning. If a potential cannot be defined (a weaker condition is the existence of a stable Liapunov function), the application of catastrophe theory may be invalid. Another reason is that, once a complete dynamical system has been defined, catastrophe theory can give no better description of the dynamical behaviour of the system than other mathematical tools. For an appraisal of these points see Pol (1993).

Be it as it may, the additional number of pages that would be required to give an introduction to the mathematics is at the moment unwarranted. Hence we refer the reader to the relevant literature. For the mathematics see, for example, Saunders (1980) and Poston and Stewart (1978); for surveys of the main economic applications see Gabisch and Lorenz (1989, Chap. 5, Sect. 5.2.2), Rosser (1991) and Tu (1994, Chap. 10, Sect. 10.4).

26.6 Exercises

1. Consider the non-linear difference equation

$$k_{t+1} = \frac{s}{1+n} f(k_t),$$

where the production function is well behaved, and show that

(1.a) there exists a unique positive equilibrium point $(1+n)k_e = sf(k_e)$;

(1.b) the positive equilibrium point is locally stable;

(1.c) the positive equilibrium point is globally stable (Hint: use a phase diagram).

2. Examine the local stability at the origin of the Lorenz model 26.24, and show that it is stable for $\rho < 1$.

3. Show that for $\rho > 1$ the Jacobian of the Lorenz model 26.24 evaluated at the origin has one positive and two negative real roots.

4. Show that for $1 < \rho < \sigma(\sigma+\beta+3)/(\sigma-\beta-1)$ the roots of Jacobian of the Lorenz model 26.24 evaluated at the non-trivial equilibrium points are stable (Hint: use the conditions for a third-degree characteristic equation given in Chap. 16, Sect. 16.4).

5. Show that for $\rho = \sigma(\sigma+\beta+3)/(\sigma-\beta-1)$ the roots of Jacobian of the Lorenz model 26.24 evaluated at the non-trivial equilibrium points are $\lambda_1 = -(\sigma+\beta+1)$, $\lambda_{2,3} = \pm \mathrm{i}\sqrt{2\sigma(\sigma+1)/(\sigma-\beta-1)}$, hence there is a Hopf bifurcation.

6. Show that for $\rho > \sigma(\sigma+\beta+3)/(\sigma-\beta-1)$ the Jacobian of the Lorenz model 26.24 evaluated at the non-trivial equilibrium points has one negative real root and a pair of complex roots with positive real part (Hint: use the conditions for a third-degree characteristic equation given in Chap. 16, Sect. 16.4).

7. A consumer maximizes the utility function $U = x_1^a x_2^{1-a}$ subject to the budget constraint $p_1 x_1 + p_2 x_2 = M$, where p_1, p_2 are given and M is the given money income. Assume that the utility weight a is variable through time and depends on past choices according to the function $a_t = \alpha x_{1t-1}^* x_{2t-1}^*$, where α is a parameter and x_{1t-1}^*, x_{2t-1}^* are the previous period's optimal quantities. Show that chaos occurs for suitable values of α (hint: this is not an intertemporal maximization problem. Take a_t as given in period t and find x_{1t}^*, x_{2t}^* by the standard procedure of static maximization. Then replace a_t with $\alpha x_{1t-1}^* x_{2t-1}^*$ etcetera. This exercise is based on Benhabib and Day, 1981).

8. Consider an overlapping generations model in which it is assumed that a representative individual lives for two periods and, when young, determines a non-negative consumption for both his youth $c_0(t)$ and his old age $c_1(t+1)$. His preferences are represented by a utility function $U(c_0(t), c_1(t+1))$. In his youth and old age he receives endowments w_0, w_1 respectively. The interest rate at time t, ρ_t, defines the exchange rate between present and future consumption. The individual's budget constraint is

$$c_1(t+1) = w_1 + \rho_t[w_0 - c_0(t)], \quad c_0(t) \geq 0, \geq 0.$$

It is further assumed that the population and the aggregate endowment grow at the rate γ, hence the equilibrium (market-clearing) condition for the economy is $(1+\gamma)[w_0 - c_0(t)] + w_1 - c_1(t+1) = 0$.

The budget constraint and the market clearing conditions define the set of feasible programs for this economy. The representative individual of the t-th generation is assumed to maximize his utility function subject to the budget constraint.

It is known (Gale, 1973) that under weak assumptions this problem has the following simple solution

$$\rho_t = \frac{U_0(c_0(t), c_1(t+1))}{U_1(c_0(t), c_1(t+1))} = \frac{w_1 - c_1(t+1)}{c_0(t) - w_0},$$

where U_0, U_1 denote the partial derivatives of U with respect to its arguments. This solution can always be made explicit and gives rise to a first-order nonlinear difference equation

$$c_1(t+1) = f(c_0(t); w_0, w_1).$$

Suppose that the utility function is

$$U(c_0, c_1) = A - e^{k-c_0} + c_1 = A - e^{a\{1-[(c_0-w_0)/a]\}} + c_1,$$

where $k = a + w_0$ and A, a, k are positive constants, and assume that there is zero population growth ($\gamma = 0$). Then show that:

(8.a) the difference equation expressing the dynamic behaviour of the economy is

$$c_0(t+1) - w_0 = e^{a\{1-[(c_0(t)-w_0)/a]\}}[c_0(t) - w_0],$$

(8.b) there is chaos for appropriate values of the parameter a.

(Hint: set $c_0(t) - w_0 = y_t$ and use Eq. (26.10). See Benhabib and Day, 1982).

9. Show that the logistic map has at most one stable periodic orbit (Hint: apply Singer's theorem).

26.7 References

Anderson, P., K.J. Arrow and D. Pines (eds.), 1988, *The Economy as an Evolving Complex System.*

Aoki, M., 1994, New Macroeconomic Modeling Approaches: Hierarchical Dynamics and Mean-Field Approximation.

Arneodo, A., P. Coullet and C. Tresser, 1982, Oscillations with Chaotic Behavior: An Illustration of a Theorem by Shil'nikov.

Barnett, W.A. and P. Chen, 1988, The Aggregation-Theoretic Monetary

Aggregates are Chaotic and Have Strange Attractors: An Econometric Application of Mathematical Chaos.

Barnett, W.A., J. Geweke and K. Shell (eds.), 1989, *Economic Complexity: Chaos, Sunspots, Bubbles, and Nonlinearity.*

Barnett, W.A., M. Salmon and A. Kirman (eds.), 1996, *Nonlinear Dynamics in Economics.*

Baumol, W.J. and J. Benhabib, 1989, Chaos: Significance, Mechanism, and Economic Applications.

Baumol, W.J. and R.E. Quandt, 1985, Chaos Models and their Implications for Forecasting.

Beltrami, E., 1987, *Mathematics for Dynamic Modeling,* Chaps. 7-8.

Benhabib, J. (ed.), 1992, *Cycles and Chaos in Economic Equilibrium.*

Benhabib, J. and R. Day, 1981, Rational Choice and Erratic Behaviour.

Benhabib, J. and R. Day, 1982, A Characterization of Erratic Dynamics in the Overlapping Generations Model.

Boldrin, M. and M. Woodford, 1990, Equilibrium Models Displaying Endogenous Fluctuations and Chaos.

Brock, W.A., D.A. Hsieh, and B. LeBaron, 1991, *Nonlinear Dynamics, Chaos, and Instability: Statistical Theory and Economic Evidence.*

Cartigny, P. and A. Venditti, 1994, Turnpike Theory: Some New Results on the Saddle Point Property of Equilibria and on the Existence of Endogenous Cycles.

Casti, J., 1989, *Alternate Realities: Mathematical Models of Nature and Man,* Chap. 5.

Chen, P., 1988, Empirical and Theoretical Evidence of Economic Chaos.

Chen, P., 1993, Searching for Economic Chaos: A Challenge to Econometric Practice and Nonlinear Tests.

Chiarella, C., 1988, The Cobweb Model: Its Instability and the Onset of Chaos.

Chiarella, C., 1990, *The Elements of a Nonlinear Theory of Economic Dynamics.*

Creedy, J. and V.L. Martin (eds.), 1994, *Chaos and Non-linear Models in Economics: Theory and Applications.*

Day, R.H., 1982, Irregular Growth Cycles.

Day, R.H., 1994, *Complex Economic Dynamics,* Vol. I.

Day, R.H. and P. Chen (eds.), 1993, *Nonlinear Dynamics and Evolutionary Economics.*

DeGrauwe, P., H. Dewachter and M. Embrechts, 1993, *Exchange Rate Theory: Chaotic Models of Foreign Exchange Markets.*

Diener, M. and T. Poston, 1984, The Perfect Delay Convention, or The Revolt of the Slaved Variables.

Dore, M.H.I., 1993, *The Macrodynamics of Business Cycles.*

Foster, J., 1994, The Self-Organisation Approach in Economics.

Frankel, J.A. and K.A. Froot, 1988, Chartists, Fundamentalists, and the

Demand for Dollars.
Gabisch, G. and H.-W. Lorenz, 1989, *Business Cycle Theory: A Survey of Methods and Concepts.*
Gale, D., 1973, Pure Exchange Equilibrium of Dynamic Economic Models.
Gandolfo, G., 1971, *Mathematical Methods and Models in Economic Dynamics,* Appendix III, § 7.6.
Gandolfo, G., 1994, *International Economics, Vol. I: The Pure theory of International Trade,*
Gandolfo, G. and P.C. Padoan, 1990, The Italian Continuous Time Model: Theory and Empirical Results.
Gandolfo, G., P.C. Padoan, G. De Arcangelis and C.R. Wymer, 1996, Nonlinear Estimation of a Nonlinear Continuous Time Model.
Goodwin, R.M., 1947, Dynamical Coupling with Especial Reference to Markets having Production Lags.
Goodwin, R.M., 1951, The Nonlinear Accelerator and the Persistence of Business Cycles.
Goodwin, R.M., 1982, *Essays in Economic Dynamics.*
Goodwin, R.M., 1990, *Chaotic Economic Dynamics.*
Goodwin, R.M., 1991, Non-linear Dynamics and Economic Evolution.
Goodwin, R.M., M. Krüger and A. Vercelli (eds.), 1984, *Nonlinear Models of Fluctuating Growth.*
Grandmont, J.-M., 1985, On Endogenous Competitive Business Cycles.
Grandmont, J.-M., 1986, Periodic and Aperiodic Behaviour in Discrete One-Dimensional Dynamical Systems.
Grandmont, J.-M. (ed.), 1988, Nonlinear Economic Dynamics, New York, Academic Press.
Granger, C.W.J., 1994, Is Chaotic Economic Theory Relevant for Economics?
Guckenheimer, J. and P. Holmes, 1986, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields.*
Haken, H., 1983a, *Synergetics: An Introduction.*
Haken, H., 1983b, *Advanced Synergetics.*
Heiner, R.A., 1989, The Origin of Predictable Behavior.
Invernizzi, S. and A. Medio, 1991, On Lags and Chaos in Economic Dynamic Models.
Kaldor, N., 1940, A Model of the Trade Cycle.
Kelsey, D., 1988, The Economics of Chaos or the Chaos of Economics.
Koblo, R., 1991, *The Visible Hand: Synergetic Microfoundations of Macroeconomic Systems.*
Leydesdorff, L. and P. Van den Besselaar (eds.), 1994, *Evolutionary Economics and Chaos Theory: New Directions in Technology Studies.*
Liu, T., C.W.J. Granger and W.P. Heller, 1992, Using the Correlation Exponent to Decide whether an Economic Series is Chaotic.
Lorenz, H.-W., 1987a, Goodwin's Nonlinear Accelerator and Chaotic Motion.
Lorenz, H.-W., 1987b, International Trade and the Possible Occurrence of

Chaos.
Lorenz, H.-W., 1993, *Nonlinear Dynamical Economics and Chaotic Motion,* 2nd ed. (1st ed. 1989).
Malliaris, A.G. and W.A. Brock, 1982, *Stochastic Methods in Economics and Finance.*
Marotto, F.R., 1978, Snap-Back Repellers Imply Chaos in R^n.
Mandelbrot, B., 1983, *The Fractal Geometry of Nature.*
Medio, A., 1984, Synergetics and Dynamic Economic Models.
Medio, A., 1992, *Chaotic Dynamics: Theory and Applications to Economics.*
Mirowski, P., 1990, From Mandelbrot to Chaos in Economic Theory.
Mirowski, P., 1992, Do Economists Suffer from Physics Envy?
Montrucchio, L., 1992, Dynamical Systems that Solve Continuous-Time Concave Optimization Problems: Anything Goes.
Newhouse, S., D. Ruelle and F. Takens, 1978, Occurrence of Strange Axiom A Attractor near Quasi-Periodic Flows on $T^m, m \geq 3$.
Nusse, H.E. and C.H. Hommes, 1990, Resolution of Chaos with Application to a Modified Samuelson Model.
Pesaran, M.H. and S. Potter (eds.), 1992, Nonlinear Dynamics and Econometrics.
Peters, E.E., 1991, *Chaos and Order in the Capital Markets.*
Peters, E.E., 1994, *Fractal Market Analysis: Applying Chaos Theory to Investment and Economics.*
Poincaré, H., 1908, *Science et méthode,* p. 68.
Pol, E., 1993, Theoretical Economics and Catastrophe Theory: An Appraisal.
Poston, T. and I. Stewart, 1978, *Catastrophe Theory and its Applications.*
Prigogine, I., 1980, *From Being to Becoming.*
Puu, T., 1987, Complex Dynamics in Continuous Models of the Business Cycle.
Puu, T., 1993, *Nonlinear Economic Dynamics.*
Rand, D., 1978, Exotic Phenomena in Games and Duopoly Models.
Rosser, J.B., Jr., 1991, *From Catastrophe to Chaos: A General Theory of Economic Discontinuities.*
Ruelle, D., 1979, Strange Attractors.
Sargent, T.J., 1993, *Bounded Rationality in Macroeconomics.*
Saunders, P.T., 1980, *An Introduction to Catastrophe Theory.*
Scheinkman, J.A. and M. Woodford, 1994, Self-Organized Criticality and Economic Fluctuations.
Schoonbeek, L., 1995, The Instantaneous Speed of Adjustment Assumption and Stability of Economic Models.
Silverberg, G., 1984, Embodied Technical Progress in a Dynamic Economic Model: The Self-Organization Paradigm.
Sims, C., 1994, Review of Zarnowitz's book *Business Cycles.*
Sordi, S., 1993, Nonlinear Models of the Business Cycle and Chaos Theory:

A Critical Analysis.
Szpiro, G.G., 1994, Exchange Rate Speculation and Chaos-Inducing Intervention.
Thom, R., 1975 (1972), *Structural Stability and Morphogenesis.*
Thygesen, N., K. Velupillai and S. Zambelli (eds.), 1991, *Business Cycles: Theories, Evidence and Analysis.*
Tu, P.N.V., 1994, *Dynamical Systems*, Chap. 10.
Varian, H.R., 1979, Catastrophe Theory and the Business Cycle.
Woodcock, A. and M. Davis, 1978, *Catastrophe Theory.*
Wymer, C.R., 1996, Advances in the Estimation and Analysis of Non-Linear Differential Equation Models in Economics.
Zeeman, E.C., 1974, On the Unstable Behavior of Stock Exchanges.
Zhang, W.-B., 1991, *Synergetic Economics: Time and Change in Nonlinear Economics.*

Chapter 27

Mixed Differential-Difference Equations

27.1 General concepts

A mixed differential-difference equation (also called a *delay differential equation*) is a functional equation in an unknown function of time and certain of its derivatives, evaluated at different values of t[1]. If we recall the definitions of order of a differential equation and of order of a difference equation, it is clear that a mixed differential-difference equation has both a differential order and a difference order. The general form of an equation of differential order n and of difference order m is the following:

$$\begin{aligned} F[t, y(t), y(t-\omega_1), ..., y(t-\omega_m), y'(t), y'(t-\omega_1), ..., y'(t-\omega_m), ... \\ ..., y^{(n)}(t), y^{(n)}(t-\omega_1), ..., y^{(n)}(t-\omega_m)] = 0 , \end{aligned} \tag{27.1}$$

where F is a given function, and the positive numbers $\omega_i, \omega_1 < \omega_2 < ... < \omega_m$ (called retardations, lags, or delays) are also given. In the particular case in which the ω_i are equispaced, we can put $\omega_1 = 1, \omega_2 = 2, ..., \omega_m = m$.

Before going on to treat these equations, a few words are in order on the problem of the choice of 'time' in economic models.

27.2 Continuous vs discrete time in economic models

The choice of the kind of 'time' (continuous or discrete) to be used in the construction of dynamic models is a moot question. We know that such a

[1] We shall consider only functions having one independent variable, so that the derivatives and the differences which appear in the equation are ordinary, and not partial, derivatives and differences. It must be noted that both the function and its derivatives may be evaluated at different points of time.

choice implies the use of different analytical tools: differential equations in the continuous case, difference equations in the discrete one. This choice is not unbiased with respect to the results of the model, given the different behaviour of the solutions of these types of functional equations and the different nature of the stability conditions. A very simple example will suffice to suffragate this statement. In continuous time, the price-adjustment mechanism to excess demand (see Chap. 13, Sect. 13.1) is

$$p' = f[D(p) - S(p)], \tag{27.2}$$

where f is a sign-preserving function. After linearising and considering the deviations from equilibrium we have

$$\overline{p}' = c(b - b_1)\overline{p}, \tag{27.3}$$

where c is the adjustment speed and b, b_1 are the slope of the demand and supply curve respectively. The solution is $\overline{p} = Ae^{c(b-b_1)t}$, and the stability condition is

$$b - b_1 < 0. \tag{27.4}$$

Note that the parameter c plays no role in determining stability: it simply serves to determine the speed of the movement. In other words, once the movement has been found to be stable, a higher value of c means a faster convergence to equilibrium.

Let us now consider the discrete-time equivalent to (27.2), that is, after linearisation,

$$\Delta\overline{p}_t = \overline{p}_{t+1} - \overline{p}_t = c_\Delta(b - b_1)\overline{p}_t. \tag{27.5}$$

Note that we have used a different symbol for the adjustment speed—in fact, c_Δ is not independent of the time unit chosen.

The solution is

$$\overline{p}_t = A[c_\Delta(b - b_1) + 1]^t,$$

which gives rise to the stability condition $|c_\Delta(b - b_1) + 1| < 1$ or $-1 < c_\Delta(b - b_1) + 1 < 1$, from which

$$-2 < c_\Delta(b - b_1) < 0. \tag{27.6}$$

It is easy to see that condition (27.4), although necessary, is no longer sufficient. Condition (27.6) additionally requires the adjustment speed to be smaller than a critical value, which is easily found to be $2/(b_1 - b)$.

Therefore, the choice between the formulation in terms of differential equations and that in terms of difference equations is by no means an irrelevant matter. This choice is often made on the basis of practical considerations, such as the empirical analysis of the model under examination. Since statistical data are available as discrete-time measurements, it seems natural to start with a discrete time model with a view to empirical testing of the

theory. In fact, current econometric techniques for the estimation of complete models are based on discrete time: therefore, in order to estimate a model, it seems necessary to formulate it in discrete time. What if it has been conceived in continuous time for theoretical reasons? Well, the obvious answer was to approximate it with a discrete model (but the usual approximations are not acceptable once the problem is formulated rigorously).

Fortunately, the problem of statistical inference in continuous-time dynamic models has now been dealt with satisfactorily (in addition to the works by Bergstrom ed., 1976, Bergstrom, 1984, Gandolfo, 1981, Wymer, 1972, see also the papers included in Gandolfo ed., 1993). This allows a rigorous estimation of the parameters of systems of stochastic differential equations on the basis of samples of discrete observations such as are available in reality. This methodology marks a great advance, because it allows the economist who has chosen a continuous-time model at the theoretical level, to estimate its parameters rigorously and independently of the observation interval. This means that the bias in favour of discrete-time models (imposed by current econometric methods and not by economic theory) disappears. Thus (although continuous-time econometrics has not yet made its way into standard graduate teaching) there is now no practical reason for choosing a discrete-time rather than a continuous-time model as the representation of a theory to be analysed empirically.

Thus we are back to our initial problem or, how to make a choice. I will now list the main arguments in favour of the use of continuous models, which, for simplicity, I have grouped into eight categories. For an in-depth treatment of what I am summarizing, see the works listed above.

1) Although individual economic decisions are generally made at discrete time intervals, it is difficult to believe that they are coordinated in such a way as to be perfectly synchronized (that is, made at the same moment and with reference to the same time interval as postulated by period analysis). On the contrary, it is plausible to think that they overlap in time in some stochastic manner. As the variables that are usually considered and observed by the economist are the outcome of a great number of decisions taken by different operators at different points of time, it seems natural to treat economic phenomena as if they were continuous.

2) A specification in continuous time is particularly useful for the formulation of dynamic adjustment processes based on excess demand, a plausible discrete equivalent of which is often difficult to find when both stock and flow variables are involved.

3) What has been said at Point 2) is connected with another advantage of continuous-time models: the estimator of these models is independent of the observation interval. What's more, it explicitly takes into account the fact that a flow variable cannot be measured instantaneously, so that what we actually observe is the integral of such a variable over the observation period (this allows a correct treatment of stock-flow models). These properties do

not hold for the discrete-time models usually employed, which therefore must be formulated explicitly in relation to the data which are available or which one wishes to use. So, for example, a model built to be estimated with quarterly data will be different from one built to be estimated with annual data.

4) A further difficulty of discrete analysis is that usually there is no obvious time interval that can serve as a 'natural' unit. Lacking this, the assumption of a certain fixed period length may unwittingly be the source of misleading conclusions. Thus, it is necessary to check that no essential result of a discrete model depends on the actual time length of the period (the model should give the same results when such a period is, say, doubled or halved). But, if the results are unvarying with respect to the period length, they should remain valid when this length tends to zero (that is when one goes from discrete to continuous analysis). Some authors attach a special importance to this property and argue that the test of the invariance of results with respect to the length of the period is fundamental in order to ascertain whether a discrete model is well defined and consistent. Moreover, the lack of clarification, in such models, of whether equilibrium obtains at the beginning or at the end of the period, may give rise to confusion between stock equilibria and flow equilibria.

5) The partial adjustment functions discussed under 2), above (and, more generally, all continuously distributed lag adjustment functions) may have very high adjustment speeds and therefore very short mean time-lags with respect to the observation period. Because of this, it may happen that, when the variables are measured in discrete time, the desired value practically coincides with the observed value over the period, so that it is not possible to obtain an estimate of the adjustment speed. On the contrary, with the continuous formulation, it is always possible to obtain asymptotically unbiased estimates of the adjustment speed even for relatively long observation periods.

This possibility has important implications, especially when adjustment speeds play a crucial role, for example to determine which markets clear more rapidly.

6) The use of a continuous model may allow a more satisfactory treatment of distributed lag processes. In a discrete model the disturbances in successive observations are usually assumed to be independent, but this assumption can be maintained only if the size of the time unit inherent in the model is not too small relative to the observation period. The assumption of independence thus entails a lower limit on the permissible size of the inherent time unit, and this precludes the correct treatment of a number of economic problems. The lags in the system are not always integral multiples of one time unit whose size is compatible with the assumption of independence. As it may happen that distributed time-lags with a lower time limit of almost zero have to be considered, a continuous time specification is more correct.

7) From the analytical point of view, differential equation system are usually more easily handled than difference systems.

8) The availability of a model formulated as a system of differential equations enables its user—once the parameter estimates have been obtained—to get forecasts and simulations for any time interval, and not only for the time unit inherent in the data. In fact, the solution of the system of differential equations yields continuous paths for the endogenous variables, given the initial conditions concerning these variables and the time paths of the exogenous variables. This is very important for policy purposes. Suppose that, for example, a policy simulation tells us that the money supply should be brought from 100 now to 102 within three months. A discrete quarterly model cannot say more than this. But a continuous model based on the same data can trace the continuous path of the money supply and hence offer the policy maker precious suggestions on the day-to-day management of this variable.

To conclude the case for continuous time models, it is interesting to note that, quite independently of each other, the first writers to explicitly advocate the use of continuous models were an economic theorist (Goodwin, 1948, pp. 113-4) and an econometrician (Koopmans, 1950). It is also interesting to note that Koopmans put forward the idea of formulating econometric models in continuous time in a short essay contained in the volume that laid the foundation for modern econometric methodology. In this essay Koopmans also illustrated some of the advantages of these models over those formulated in discrete time. These advantages led Marschak to state, in the general introduction to the volume mentioned, that 'if proper mathematical treatment of stochastic models can be developed, such models [that is, those formulated in continuous time] promise to be a more accurate and more flexible tool for inference in economics than the discrete models used heretofore' (Marschak, 1950, p. 39; square brackets are ours). However—except in occasional works—these suggestions were not followed up and the topic was taken up again and given an operational content only in the 1970s, as shown in the references listed above.

However, our case for continuous-time formulations does not exclude the possibility that discrete-time lags and other similar phenomena are present. This is where mixed differential-difference equations come in. Hence, in our opinion, these equations are much more suitable than differential equations alone or difference equations alone for an adequate treatment of dynamic economic phenomena. We think so because we believe that real-life dynamic economic phenomena are approximately continuous phenomena, in which, however, discontinuous lags, influences on the current value of a variable of previous values of the same variable (placed at a finite, and not infinitesimal, distance), and so on, are present in an essential way.

This is why we have devoted a chapter to mixed differential-difference equations, in spite of their yet scarce use in economic dynamics (see below, Sect. 27.7).

27.3 Linear mixed equations

The treatment here will be limited to linear equations with constant coefficients and of the first order both in derivatives and in differences, which have the form

$$a_0 y'(t) + a_1 y'(t-\omega) + b_0 y(t) + b_1 y(t-\omega) = g(t) \ , \qquad (27.7)$$

where $a_0, a_1, b_0, b_1, \omega$ are given constants and $g(t)$ is a known functions. Eq. (27.7) is non-homogeneous; the corresponding homogeneous equation is

$$a_0 y'(t) + a_1 y'(t-\omega) + b_0 y(t) + b_1 y(t-\omega) = 0. \qquad (27.8)$$

Linear equations may of course also arise from the linear approximation of a non-linear equation. We shall now give a few more definitions. An equation having form (27.7) is called '*of retarded type*' if $a_0 \neq 0, a_1 = 0$; '*of neutral type*' if $a_0 \neq 0, a_1 \neq 0$; '*of advanced type*' if $a_0 = 0, a_1 \neq 0$. Equations of retarded type are also called *differential equations with time delay* and *hystero-differential equations.*

For Eqs. (27.7) and (27.8) the same general principles hold as for difference equations (see Part I, Chap. 2, Sect. 2.2) and differential equations (see Part II, Chap. 11, Sect. 11.2), that is:

(1) if $y_1(t)$ is a solution of (27.8), $Ay_1(t)$ also is a solution, where A is an arbitrary constant;

(2) if $y_1(t)$ and $y_2(t)$ are two linearly independent solutions of (27.8), $A_1y_1(t)+A_2y_2(t)$ is also a solution, where A_1 and A_2 are arbitrary constants; this property (the principle of superposition) can be extended to any number of linearly independent solutions;

(3) if $w(t)$ is a solution of (27.8) and $\bar{y}(t)$ is a particular solution of (27.7), then $w(t) + \bar{y}(t)$ is a solution of (27.7).

The reader can check these principles by direct substitution in Eqs. (27.7) and (27.8). Given (3), the main problem is to find the general solution of (27.8), then adding to this solution, if the equation is non-homogeneous, a particular solution $\bar{y}(t)$ of (27.7).

27.4 The method of solution

Property (2) of the previous section suggests the possibility of finding the general solution of (27.8) as a linear combination of elementary solutions, according to the procedure already applied with success to both differential and difference equations. Moreover, analogy suggests that such elementary solutions might be of exponential type $e^{\lambda t}$, where λ is a constant to be

determined[2]. Substituting in (27.8) and collecting terms, we have

$$e^{\lambda t}\left(a_0\lambda + a_1\lambda e^{-\omega\lambda} + b_0 + b_1 e^{-\omega\lambda}\right) = 0 \ . \qquad (27.9)$$

Therefore $e^{\lambda t}$ is a solution of (27.8) if, and only if, Eq. (27.9) is identically satisfied, which in turn is true if, and only if,

$$a_0\lambda + a_1\lambda e^{-\omega\lambda} + b_0 + b_1 e^{-\omega\lambda} = 0 \ . \qquad (27.10)$$

Eq. (27.10) is called the *characteristic equation* of (27.8) and its roots are called the *characteristic roots.*

Eq. (27.10) is a transcendental equation, which has an infinite number of roots. Under suitable conditions (Bellman and Cooke, 1963, Chap. 5), the general solution of (27.14) can then be written in the form

$$y(t) = \sum_{r=1}^{\infty} c_r \exp(\lambda_r t) \ , \qquad (27.11)$$

if all the characteristic roots λ_r are distinct (the coefficients c_r are arbitrary constants), and in the form

$$y(t) = \sum_{r=1}^{\infty} p_r(t) \exp(\lambda_r t) \ , \qquad (27.12)$$

if there are multiple roots, where $p_r(t)$ is a polynomial in t (of degree less than the multiplicity of the corresponding root). If there are complex roots (that always occur in conjugate pairs, see below, footnote 3), these will be transformed into a trigonometric oscillation, as we know from the treatment of linear differential equations (see, for example, Chap. 14, Sect. 14.1.3).

The study of the transcendental equation (27.10) is rather complicated, hence we shall concentrate on equations of retarded type, both because their study is sufficient to illustrate the general method of solution based on the characteristic equation and because it is this type which has been more frequently used in economic applications.

We shall then consider equations of the type

$$a_0 y'(t) + b_0 y(t) + b_1 y(t-\omega) = g(t) \ , \qquad (27.13)$$

and of the type

$$a_0 y'(t) + b_0 y(t) + b_1 y(t-\omega) = 0 \ . \qquad (27.14)$$

[2] The reader may wonder why $e^{\lambda t}$ and not μ^t (the former is suggested by differential, and the latter by difference, equations). An intuitive reason is that, if the unknown function is differentiable, the 'differential' nature is stronger than the 'difference' nature. Another (practical) reason is that, when differentiability obtains, the function $e^{\lambda t}$ is easier to treat than the function μ^t.

The characteristic equation (27.10) becomes, when the equation is of retarded type,

$$a_0\lambda + b_0 + b_1 e^{-\omega\lambda} = 0 \ . \tag{27.15}$$

As far as multiple roots are concerned, it can be proved that Eq. (27.15) has at most one multiple root, and that, if

$$a_0 \neq b_1\omega \exp\left(1 + \frac{b_0\omega}{a_0}\right) , \tag{27.16}$$

then all the roots are simple (See Bellman and Cooke 1963, p. 109; see also below).

If we define a new variable $s \equiv \omega\lambda$, Eq. (27.15) can be rewritten as

$$\frac{a_0}{\omega}s + b_0 + b_1 e^{-s} = 0 \ . \tag{27.17}$$

Let us now multiply through by $(-\omega/a_0)\,e^s$. Defining $a \equiv -b_0\omega/a_0, b \equiv -b_1\omega/a_0$ and rearranging terms we have

$$ae^s + b - se^s = 0 \ . \tag{27.18}$$

If, instead, we multiply by $-\omega/a_0$ only, we obtain (a and b are defined as before)

$$s = a + be^{-s} \ . \tag{27.19}$$

Of course, the reason for these transformations is that the alternative forms (27.18) and (27.19) will be useful in the following treatment.

We can now tackle the problem of locating the characteristic roots. For this purpose we shall use (27.19), which is in a more convenient form. Consider the function

$$z = f(s) = s - be^{-s}, \tag{27.20}$$

so that the real roots of (27.19) are given by the intersection of $f(s)$ with the function $z = a$, a straight line parallel to the s axis. We have

$$\frac{\mathrm{d}f}{\mathrm{d}s} = 1 + be^{-s} , \qquad \frac{\mathrm{d}^2 f}{\mathrm{d}s^2} = -be^{-s} \ . \tag{27.21}$$

Two cases must be distinguished:

(1) $b > 0$. Then $\mathrm{d}f/\mathrm{d}s$ is always positive, and from (27.20), $\lim_{s\to\pm\infty} f(s) = \pm\infty$. Therefore the function $f(s)$ will intersect once, and only once, the function $z = a$; furthermore, the intersection point cannot be an inflection point, so that the real root will be a simple root. Note that, given the definition of b, if $b > 0$ then $b_1\omega$ and a_0 have opposite signs, so that (27.16) is satisfied, and conversely.

(2) $b < 0$. Then $\mathrm{d}f/\mathrm{d}s$ is zero for

$$-be^{-s} = 1 , \tag{27.22}$$

that is

$$s = \log(-b) \ . \tag{27.23}$$

Furthermore, $\mathrm{d}^2 f/\mathrm{d}s^2$ is always positive, so that (27.23) yields a unique global minimum to $f(s)$. Substituting (27.23) in (27.20) we have

$$\min f(s) = \log(-b) + 1 \ . \tag{27.24}$$

Now, from obvious geometric considerations, if $\min f(s) \gtreqless a$, the function $f(s)$ will lie entirely above, be tangent to, intersect twice, the straight line $z = a$. Therefore, Eq. (27.19) will have no real roots, two equal real roots, two distinct real roots according to whether

$$\log(-b) + 1 \gtreqless a , \tag{27.25}$$

that is, rearranged,

$$a - \log(-b) \lesseqgtr 1 \ . \tag{27.26}$$

It follows that, if $\log(-b) + 1 \neq a$, no multiple real roots may occur. This condition can be written as $\log(-b) \neq -1 + a$, that is $-b \neq e^{-1+a}$, i.e., given the definitions of a and b, as $b_1\omega/a_0 \neq \exp[-1 - (b_0\omega/a_0)]$, from which, taking reciprocals and multiplying through by $b_1\omega$, we obtain Eq. (27.16).

Thus we are able to conclude that the characteristic equation under examination has at most two real roots. Therefore all the other roots (which are infinite in number) are complex.

Regarding complex roots, let

$$s = \alpha \pm \mathrm{i}\theta \tag{27.27}$$

be the representative pair of complex roots[3]. Substituting in (27.19) we have

$$\alpha \pm \mathrm{i}\theta = a + be^{-\alpha}\exp(\mp\mathrm{i}\theta) ,$$

and, using the well-known transformation

$$\exp(\mp\mathrm{i}\theta) = \cos\theta \mp \mathrm{i}\sin\theta ,$$

we obtain

$$\alpha \pm \mathrm{i}\theta = \left(a + be^{-\alpha}\cos\theta\right) \mp ibe^{-\alpha}\sin\theta ,$$

so that, equating the corresponding real and imaginary parts,

$$\begin{aligned} \alpha &= a + be^{-\alpha}\cos\theta , \\ \theta &= -be^{-\alpha}\sin\theta \ . \end{aligned} \tag{27.28}$$

[3]We have implicitly assumed that complex roots always occur in conjugate pairs: It can be proved that it is indeed so. See Bellman and Cooke (1963, p. 108). Another way to prove this is to observe that if $\alpha + i\theta$ satisfies (27.28) below, also $\alpha - i\theta$ satisfies them.

We are not in a much better position than before, since Eqs. (27.28) are transcendental equations too. However, some conclusions concerning the intervals in which θ lies can be reached. Since, in (27.27), θ can be taken as a positive number, from the second equation of (27.28) we obtain the information that $\sin\theta$ and b must have opposite sign. This restricts the values of θ within the following intervals:

$$\begin{array}{ll} 2k\pi < 0 < (2k+1)\,\pi & \text{if } b < 0\,, \\ & k = 0, 1, 2, \ldots, n, \ldots\,, \\ (2k+1)\,\pi < \theta < (2k+2)\,\pi & \text{if } b > 0\,. \end{array} \qquad (27.29)$$

Further information can be obtained by means of a graphical analysis, an illustration of which will be given in the next section (Fig. 27.1).

The general conclusion of the foregoing analysis is that the solution of Eq. (27.10) shows in any case an oscillatory path, to which a monotonic path may or may not be superimposed.

Because of property (3) of the previous section, to complete our analysis it is sufficient to show how a particular solution $\bar{y}(t)$ of (27.13) can be found. The simplest case occurs when $g(t)$ is a constant, let us call it B. As a particular solution, we can try $\bar{y}(t) = D$, where D is an undetermined constant. Substituting in (27.13) we have

$$b_0 D + b_1 D = B\,,$$

which gives

$$D = \frac{B}{b_0 + b_1}\,.$$

If $b_0 + b_1 = 0$, try $\bar{y}(t) = Dt$. Substitution in (27.13) yields, after manipulation, $D = B/(a_0 - b_1\omega)$. If also $a_0 - b_1\omega$, try $\bar{y}(t) = Dt^2$; substitution in (27.13) and manipulation yields $D = B/b_1\omega^2$, where it must be true that $b_1\omega^2 \neq 0$, since in the contrary case Eq. (27.13) would no more be a mixed differential-difference equation, but would reduce to a simple differential equation. The above is nothing else but an application of the general principle of *undetermined coefficients*, which has already served us well in difference and in differential equations. We recall that this method consists in trying as a particular solution a function having the same form of the given function $g(t)$ but with undetermined coefficients; substituting in (27.13) one tries to determine such coefficients so that the equation is identically satisfied. If the first attempt does not succeed, one tries next the same type of function but multiplied by a polynomial in t (starting with the first degree) and proceeds in like manner.

We examine as a further example $g(t) = Be^{\gamma t}$, where B and γ are given constants. As a particular solution let us try $\bar{y}(t) = De^{\gamma t}$, where D is an undetermined constant. Substituting in (27.13) and collecting terms, we have

$$e^{\gamma t}\left[D\left(a_0\gamma + b_0 + b_1 e^{-\gamma\omega}\right) - B\right] = 0\,,$$

which will be identically satisfied if, and only if, the expression in square brackets is zero, so that

$$D = \frac{B}{a_0\gamma + b_0 + b_1 e^{-\gamma\omega}}\,.$$

In the case in which γ happens to coincide with one of the real roots of (27.10) the denominator of the fraction is zero. In such a case we try $\bar{y}(t) = Dte^{\gamma t}$. Substituting in (27.13) and proceeding as above we obtain

$$D = \frac{B}{a_0 - b_1\omega e^{-\gamma\omega}}\,.$$

The reader may examine as an exercise the case in which $a_0 - b_1\omega e^{-\gamma\omega} = 0$.

27.5 Stability conditions

It would be desirable, at this point, to have at hand *stability conditions* to check whether the real parts of the roots of the characteristic equation—be they real or complex—are all negative without having to solve the characteristic equation itself. Fortunately, such conditions exist, and can be given in two forms. Before proceeding further, note that, since ω is a positive magnitude, the negativity of the real part of s is necessary and sufficient for the negativity of the real part of λ. We can now state (see Bellman and Cooke, 1963, pp. 444-6) the following:

Hayes' Theorem. All the roots of (27.18) have negative real parts if, and only if,

(1) $a < 1$,

and

(2) $a < -b < (a_1^2 + a^2)^{1/2}$,

where a_1 is the root of the equation $x = a\tan x$ such that $0 < x < \pi$. If $a = 0$, we take $a_1 = 1/2\pi$.

The stability conditions given in this theorem are not easy to apply, since we must first solve the transcendental equation $x = a\tan x$. Therefore we think that the conditions given by Burger are preferable, since their application requires only the solution of the simpler equation $\cos x = \alpha$ (where α is a given constant). Therefore we state[4]:

[4]For the proof of this theorem and of its equivalence with Hayes' theorem, see Burger (1956). Given its superiority over Hayes' theorem, it is surprising that this theorem has generally passed unnoticed both in the mathematical and in the mathematical economics literature (e.g., in Bellman's and Cooke's treatise on differential-difference equations Burger's theorem is not even mentioned; and in the mathematical economics literature, as far as we know, only Hayes' theorem has been used).

Burger's Theorem. For all roots of (27.19) to possess negative real parts it is necessary and sufficient

(1) in the case $b \geq -1$ that $a < -b$;

(2) in the case $b < -1$ that $a < b$ or that $b \leq a < -b$ and $\arccos(-a/b) > +(b^2 - a^2)^{1/2}$, where the value of the function arccos is restricted by $0 < \arccos(-a/b) \leq \pi$.

As we said above, Burger's conditions are, in general, simpler to apply than Hayes' since to solve the equation $\cos x = -a/b$ is simpler than to solve the equation $a \tan x = x$. The only case in which Hayes' theorem is easier to apply is when $a = 0$. In that case, in fact, the stability condition is $0 < -b < 1/2\pi$, which we immediately obtain by using Hayes' theorem, whereas to obtain it by applying Burger's theorem is a slightly longer business.

27.6 Delay differential equations and chaos

Chaos has been treated at length in the previous chapter, where we examined both difference equations and differential equations. Since time is treated as a continuous variable in delay differential equations, their behaviour might seem more similar to that of differential equations, including the fact that one needs at least a 3×3 system for chaos to arise (see Chap. 26, Sect. 26.3). This, however, would be a wrong impression. As made clear in the procedure of solution (see above, Sect. 27.4), a first-order delay differential equation is essentially equivalent to an infinite dimensional problem. Hence even a *first-order delay differential equation may in principle give rise to chaos.*

Hopf bifurcation theory has been extended by Hale (1979; see also Busenberg and Martelli eds., 1991) to delay differential equations by considering the delay ω as a parameter, but general analytical tools for studying the onset of chaos in delay differential equations are still lacking. Numerical simulations have shown that chaotic behaviour is quite possible even in relatively simple first-order equations (Wen, Chen and Zhang, 1992; Jarsulic, 1993). We must wait for further mathematical research into this topic.

27.7 Some economic applications

Whereas differential equations and difference equations both have a lot of economic applications, the economic applications of mixed differential–difference equations are so far very few. In addition to the applications examined in this section, we recall—no claim to exhaustiveness being made—Tinbergen's (1959) shipbuilding cycle model, Frisch's (1933) essay (see also Zambelli, 1991), an application by Leontief (1953) to his own dynamic model, some

applications cited by Burger (1956), an application by Furuno (1965) to two-sector growth models, an application by Burmeister and Turnovsky (1976) to adaptive expectations in continuous time models, an application by Padoan (in Padoan and Petit, 1978, Sect. 3) to a model of the wage-price spiral, a few papers by Chukwu (see, e.g., 1992) in a controllability context, a study by Wen, Chen and Zhang (1992) on the soft-bouncing oscillator in business cycle models, a paper by Jarsulic (1993) on the Keynesian system, a study by Wen (1995) on stock market dynamics.

Perhaps it is the greater formal difficulty of mixed equations with respect to 'pure' differential or difference equations that has prevented a larger use of mixed equations in economic dynamics. But the formal difficulty should be no hindrance if the superiority of mixed equations with respect to pure equations were definitely proved.

Let us now pass to the applications. The first is Kalecki's model; the second is a formalization of ours concerning the classical price-specie flow mechanism of balance of payments adjustment.

27.7.1 Kalecki's business cycle model

27.7.1.1 The model

The model concerns a closed economic system without trend (we are then in the short run). Indicating with B the total real income of capitalist, we have the relation

$$B = C + A\,, \tag{27.30}$$

where C is capitalists' consumption, while A—if we neglect workers' saving and capital incomes—coincides with gross capital accumulation[5]. Capitalists' consumption can be related to gross real profits (capitalists' income) by means of the linear function

$$C = C_1 + \lambda B\,, \tag{27.31}$$

where C_1 is a positive constant and λ is a positive constant too, but smaller than 1. From (27.30) and (27.31) we obtain

$$B = \frac{C_1 + A}{1-\lambda}\,. \tag{27.32}$$

Although Kalecki's model is pre-Keynesian, we have here the ingredients made famous by Keynes: the consumption function and the multiplier.

Regarding investment, Kalecki assumes that there is an average *gestation lag* of investment equal to a positive constant θ. The gestation period of

[5]This amounts to the assumption that all wages are consumed, an extreme case—made to simplify the analysis—of the plausible assumption that workers' propensity to consume is greater than capitalists' propensity to consume.

investment is the time interval between the decision to invest and the delivery of the finished capital goods. More precisely, in each investment three stages can be distinguished:

(1) investment orders, i.e. all the orders for capital goods, both for replacement purposes and for net additions to the capital stock, they are called I;

(2) production of capital goods, that is gross capital accumulation, call it A;

(3) deliveries of finished capital goods, call them L.

Given the assumption made above, we can write

$$L(t) = I(t-\theta) \ . \tag{27.33}$$

To find the relation between A and I we proceed as follow. Let $W(t)$ be the total amount in time t of unfilled investment orders. Since each order requires a period of time θ to be filled, and assuming that the construction of the ordered capital goods proceeds at an even pace (that is, $1/\theta$ of each order is executed per unit of time)[6], it follows that the production of capital goods is equal to $(1/\theta)W$; therefore

$$A = \frac{W}{\theta} \ . \tag{27.34}$$

As regard $W(t)$, it equals the sum of all orders made in the interval $(t-\theta, t)$; in fact, since θ is the gestation period of investment, all orders placed during the said interval are still unfilled, whereas all orders placed previously are already filled. Therefore

$$W(t) = \int_{t-\theta}^{t} I(\tau)\, d\tau, \tag{27.35}$$

and so, given (27.34),

$$A(t) = \frac{1}{\theta}\int_{t-\theta}^{t} I(\tau)\, d\tau \ . \tag{27.36}$$

The meaning of Eq. (27.36) is that the output of capital goods at time t is equal to an average—expressed in continuous terms—of the orders placed in the interval $(t-\theta, t)$.

If we call K the capital stock, its first derivative with respect to time is its net increment, so that

$$K'(t) = L(t) - U \ , \tag{27.37}$$

[6]Kalecki (1935, p. 328) seems to consider this fact as a consequences of the existence of a gestation lag. We think that it is more correct to regard it as a separate assumption.

where U indicates physical depreciation. Kalecki assumes that in the period under consideration U is a constant.

To 'close' the model we need an investment function. According to Kalecki, there are two main determinants of investment: the gross profit rate B/K and the money rate of interest which he calls p. However, such variables, in Kalecki's opinion, do not influence the absolute level of investment but rather its level relative to the capital stock, that is the ratio I/K; in fact, when B and K increase in the same proportion, so that the ratio B/K remains unchanged, I probably rises. Thus we have the equation

$$\frac{I}{K} = f\left(\frac{B}{K}, p\right) \ . \tag{27.38}$$

In the absence of external actions and except for situations of 'financial panic', the money rate of interest usually varies according to the general business conditions, which are represented by B/K. Thus the money rate p can be assumed to be an increasing function of B/K, and consequently f is a function of B/K only. Since B is proportional to $C_1 + A$ by Eq. (27.32), we can write

$$\frac{I}{K} = \phi\left(\frac{C_1 + A}{K}\right) , \tag{27.39}$$

where ϕ is an increasing function. Taking a linear approximation we have

$$\frac{I}{K} = m\frac{C_1 + A}{K} - n \ , \tag{27.40}$$

where m and n are positive constants.

The positivity of m is obvious, given the assumption that ϕ is an increasing function. The assumption $n > 0$ is not made by Kalecki in the 1935 article. However, he had shown the positivity of n in an earlier work (Kalecki, 1933, pp. 7-8 of the 1971 reprint) by means of the following considerations. From (27.39) we get $n = [m(C_1 + A) - I]/K$. Now, suppose that $I < mC_1$: consequently, $[m(C_1 + A) - I]/K > mA/K$, and so $mA/K < n$. As $m > 0$ and being A and K positive, it follows that $n > 0$.

A simpler way of arriving at the same result is to observe that I is *gross* investment, so that the investment function must be such as to admit zero or negative values of I, which requires the positivity of n.

From Eq. (27.40) we have

$$I = m(C_1 + A) - nK \ . \tag{27.41}$$

27.7.1.2 The dynamics

Let us list here for the sake of convenience the basic equations of the model:

$$L(t) \quad = \quad I(t-\theta) \ , \tag{27.42}$$

$$A(t) = \frac{1}{\theta}\int_{t-\theta}^{t} I(\tau)\,d\tau\ , \tag{27.43}$$

$$K'(t) = L(t) - U\ , \tag{27.44}$$

$$I = m(C_1 + A) - nK\ . \tag{27.45}$$

Differentiating (27.41) we have

$$I'(t) = mA'(t) - nK'(t)\ , \tag{27.46}$$

and differentiating (27.36) we obtain

$$A'(t) = \frac{I(t) - I(t-\theta)}{\theta}\ . \tag{27.47}$$

From (27.33) and (27.37) we obtain

$$K'(t) = I(t-\theta) - U\ . \tag{27.48}$$

Substituting in (27.46) from (27.47) and from (27.48) we have

$$I'(t) = \frac{m}{\theta}[I(t) - I(t-\theta)] - n[I(t-\theta) - U]\ . \tag{27.49}$$

Denoting by $J(t)$ the difference $I(t) - U$, that is net investment, we have, taking account of the fact that U is a constant, so that $J'(t) = I'(t)$, the equation

$$J'(t) = \frac{m}{\theta}[J(t) - J(t-\theta)] - nJ(t-\theta)\ ,$$

that is

$$\theta J'(t) - mJ(t) + (m+\theta n)J(t-\theta) = 0\ . \tag{27.50}$$

Eq. (27.50) is a mixed differential-difference equation of the type analysed in the previous section. Its characteristic equation is

$$\theta\lambda - m + (m+\theta n)e^{-\theta\lambda} = 0\ , \tag{27.51}$$

that is, letting $\theta\lambda \equiv s$ and rearranging terms, we have

$$s = m - (m+\theta n)e^{-s}\ , \tag{27.52}$$

which has the form of (27.19) in Sect. 27.4. For the real roots, applying Eq. (27.26) of the previous section we obtain the result that Eq. (27.52) will have no real roots, two equal real roots, two distinct real roots according to whether

$$m - \log(m+\theta n) \lesseqqgtr 1\ . \tag{27.53}$$

On *a priori* grounds it is not possible to determine which is the sign which obtains in inequality (27.53).The reader who feels inclined to try with numerical values of the parameters may find the following transformations useful. Consider Eq. (27.41) and write it as

$$I = n\left[\frac{m}{n}(C_1 + A) - K\right]\ .$$

Now, from (27.32), $C_1 + A = (1-\lambda)B$. Let $\rho = B/Y$ be the capitalists' share in national income. Then

$$I = n\left\{\left[\frac{m}{n}(1-\lambda)\rho\right]Y - K\right\}\ . \tag{27.54}$$

Eq. (27.54) is the usual capital stock adjustment equation, where n = reaction coefficient, and $[(m/n)(1-\lambda)\rho]$ can be interpreted as the (desired) capital/output ratio, so that $[(m/n)(1-\lambda)\rho]Y$ is the desired capital stock. Thus, given the values of the capital/output ratio and of n, λ, ρ, the value of m can be determined.

For the complex roots, applying Eqs. (27.28), we have

$$\begin{aligned} \alpha &= m - (m+\theta n)e^{-\alpha}\cos\beta\ , \\ \beta &= (m+\theta n)e^{-\alpha}\sin\beta\ , \end{aligned} \tag{27.55}$$

where $\alpha \pm i\beta$ is the typical pair of complex conjugate roots. From the second equation in (27.55) we have[7]

$$e^{\alpha} = (m+\theta n)\frac{\sin\beta}{\beta}\ ,$$

which gives

$$\alpha = \log(m+\theta n) + \log\frac{\sin\beta}{\beta}\ .$$

Substituting in the first equation of (27.55) we have

$$\log(m+\theta n) + \log\frac{\sin\beta}{\beta} = m - (m+\theta n)\left(\frac{1}{m+\theta n}\frac{\beta}{\sin\beta}\right)\cos\beta\ ,$$

so that

$$\log\frac{\sin\beta}{\beta} + \frac{\beta}{\tan\beta} = m - \log(m+\theta n)\ . \tag{27.56}$$

Letting

$$\begin{aligned} f(\beta) &= \log\frac{\sin\beta}{\beta} + \frac{\beta}{\tan\beta}\ , \\ C &= m - \log(m+\theta n)\ , \end{aligned}$$

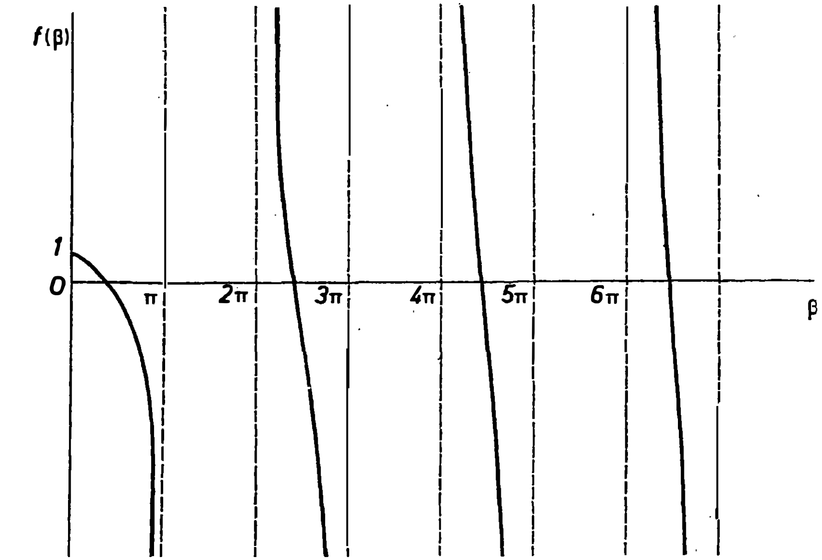


Figure 27.1: Roots of Kalecki's model

we obtain the values of β as the intersections of $f(\beta)$ with C. The function $f(\beta)$ does not depend on the 'structural' parameters of the model and so can be plotted once and for all. Figure 27.1 is due to Frisch and Holme.

Imagine a straight line parallel to the β-axis. It will, in any case, have one intersection with $f(\beta)$ in each of the intervals

$$2k\pi < \beta < (2k+1)\pi \ , \qquad k = 1, 2, ..., n \ , \ ... \ , \tag{27.57}$$

and, if $C < 1$, it will also have one intersection in the interval

$$0 < \beta < \pi \ . \tag{27.58}$$

If $C > 1$, the latter intersection does not exist, and, if $C = 1$, it exists for $\beta = 0$, which is irrelevant.

The cycle corresponding to the intersection in interval (27.58) is called by Frisch and Holme, and by Kalecki a *major* cycle, since its period is greater than 2θ (twice the gestation lag). In fact, let us recall that $s \equiv \theta\lambda$ and so $\lambda = \alpha/\theta \pm i\beta/\theta$, so that the period of the oscillation is $2\pi/(\beta/\theta)$. If β falls in range (27.58), then clearly $2\pi\theta/\beta > 2\theta$.

[7]The following treatment is due to Frisch and Holme (1935).

The other cycles, given by intersections in intervals (27.57), all have periods smaller than θ (*minor* cycles).

Thus, taking account of the definition of C and of (27.53), we can conclude that the necessary and sufficient condition for the existence of a major cycle is that the characteristic equation has no real roots.

Let us now examine the stability of the model. Applying Burger's theorem we obtain the following necessary and sufficient stability conditions:

(1) if $-(m+\theta n) \geq -1$, that is if $m + \theta n \leq 1$, the condition is

$$m < m + \theta n \ ,$$

which is obviously satisfied;

(2) if $-(m+\theta n) < -1$, that is, if $m + \theta n > 1$, the conditions are

$$\text{(a)} \quad -(m+\theta n) < m < m + \theta n \ ,$$

$$\text{(b)} \quad \arccos\left(\frac{m}{m+\theta n}\right) > \left(\theta^2 n^2 + 2m\theta n\right)^{1/2} \ .$$

Condition (b) is the crucial one, since condition (a) is obviously satisfied. Now, since $m/(m+\theta n) < 1, \arccos(...) < 1/2\pi \simeq 1.57$. Thus, is the expression under square root is not smaller than $(1.57)^2 = 2.4649$, inequality (b) is certainly not satisfied (of course, the converse is not true).

27.7.2 A formalization of the classical price-specie-flow mechanism of balance of payments adjustment

27.7.2.1 The model

The classical price-specie-flow mechanism can thus be described: a balance of payments[8] surplus causes an inflow of gold in the surplus country, that is an increase in the supply of money and consequently, according to the quantity theory, an increase in the price level. This increase, on the one hand, tends to curb exports, since the goods of the surplus country are now relatively more expensive on the international market, and to stimulate imports, since foreign goods are now relatively cheaper. Therefore, a gradual disappearance of the surplus takes place[9]. Similar reasoning explains the elimination of a deficit: gold flows out, that is, the supply of money decreases, the price level decreases and this stimulates exports and curbs imports, thus leading to a gradual disappearance of the deficit[10].

We shall now formalize this theory. First of all, let us state some assumptions which are present, explicitly or implicitly, in the classical theory:

[8]'Balance of payments' here is used in the sense of 'balance of trade' only.

[9]As we shall see, this requires that certain stability conditions are satisfied.

[10]Let us note that, in the case in which the circulating medium is paper money, it must be of the type with 100% gold reserve or with a fixed fractional gold reserve.

(1) There is free trade, with no interference by government or by monopolistic agents.

(2) The level of national output is given, normally at the full employment level, and the price level is determined by the quantity theory of money.

(3) During the adjustment period changes in the supply of money are occurring only because of the surpluses (or deficit) in the balance of payments. We could introduce other causes of variation, but this would complicate the analysis without any great advantage.

The following are simplifying assumptions:

(4) Since the rate of exchange is fixed, we assume that it is equal to unity (this does not involve any loss of generality, since it is only a matter of choice of units). Moreover, we assume that international payments take place in gold.

(5) Transport costs, insurance, etc., are neglected both for goods and for gold.

(6) Interactions with the 'rest of the world' are neglected, that is, we are considering a 'small' country.

In what follows we shall use the following symbols:

Q = supply of money,
V = velocity of circulation of money (assumed constant),
Y = level of natinal output (assumed given),
P = level of domestic prices,
P_M = level of foreign prices (assumed given),
M = quantity of imports,
X = quantity of exports.

The model can be expressed by the following equations:

$$QV = PY\ , \tag{27.59}$$

which expresses the quantity theory;

$$X = X(P)\ , \qquad dX/dP < 0\ , \tag{27.60}$$

i.e. exports are a decreasing function of the home price level, given the foreign price level;

$$M = M(P)\ , \qquad dM/dP > 0\ , \tag{27.61}$$

i.e. imports are an increasing function of the home price level;

$$PX(P) - P_M M(P) = 0\ , \tag{27.62}$$

the above being the balance of payments equilibrium equation.

We have four equations to determine the four unknowns Q, X, M, P. Note that Q is also an unknown. In fact, there can be no equilibrium if the supply

of money is not such as to give rise to a price level which is the one appropriate for the balance of payments equilibrium (this is an aspect of the 'optimum distribution of specie'). We shall assume that the equilibrium point exists and is economically meaningful.

If the system is not in equilibrium, flows of gold will take place. Assuming that the supply of money and the quantity of gold coincide[11], we have that the variation in the supply of money is in each instant equal to the surplus or deficit in the balance of payments, so that

$$dQ/dt = PX(P) - P_M M(P)\ . \tag{27.63}$$

In turn, the variation in the supply of money causes a change in the price level, given (27.59). It seems plausible to believe that this change does not occur immediately, but after some lag, say ω. In other words, the change in the price level occurring at time t depends, through (27.59), on the change in the money supply which occurred at time $t-\omega$, where ω is fixed. Thus we have the equation

$$\frac{dP}{dt} = \frac{V}{Y}\frac{dQ}{d(t-\omega)}\ . \tag{27.64}$$

From (27.63) and (27.64) we have

$$\frac{dP}{dt} = \frac{V}{Y}\{P(t-\omega)X[P(t-\omega)] - P_M M[P(t-\omega)]\}\ . \tag{27.65}$$

27.7.2.2 Stability

We shall study local stability, so that we can use the linear approximation

$$\frac{d\overline{P}}{dt} = \frac{V}{Y}\alpha\overline{P}(t-\omega), \tag{27.66}$$

where $\overline{P} = P - P^e$ indicates the deviations from equilibrium, and

$$\alpha \equiv X^e + P^e\left(\frac{dX}{dP}\right)^e - P_M\left(\frac{dM}{dP}\right)^e .$$

The characteristic equation of (27.66) is

$$\lambda - \frac{V}{Y}\alpha e^{-\lambda\omega} = 0\ . \tag{27.67}$$

Setting $s \equiv \lambda\omega$ and multiplying both members of (27.67) by $-\omega e^s$, we obtain

$$\frac{V\omega}{Y}\alpha - se^s = 0\ , \tag{27.68}$$

[11] If it were not so, the only difference would be a multiplicative constant in the right-hand side of (27.63).

which has the form (27.68) of Sect. 27.4. Applying Hayes' theorem we obtain the following necessary and sufficient stability conditions

$$0 < -\frac{V\omega}{Y}\alpha < \frac{1}{2}\pi \ . \tag{27.69}$$

The left-hand part of the double inequality (27.69) implies that $\alpha < 0$, the right-hand part that α be in absolute value smaller than $1/2\pi Y/V\omega$. It is interesting to note that such conditions are *more restrictive* than those holding in the case in which no lag is assumed. In fact, in the latter case we would obtain the differential equation

$$\frac{\mathrm{d}\overline{P}}{\mathrm{d}t} = \frac{V}{Y}\alpha\overline{P}(t) \ ,$$

whose characteristic equation is $\lambda - (V/Y)\,\alpha = 0$, so that the stability condition would be $\alpha < 0$.[12]

We want now to give an economic interpretation to the stability conditions. For this purpose some manipulations on α are required. We have

$$X^e + P^e\left(\frac{\mathrm{d}X}{\mathrm{d}P}\right)^e - P_M\left(\frac{\mathrm{d}M}{\mathrm{d}P}\right)^e = X^e\left[1+\frac{P^e}{X^e}\left(\frac{\mathrm{d}X}{\mathrm{d}P}\right)^e - \frac{P_M}{X^e}\left(\frac{\mathrm{d}M}{\mathrm{d}P}\right)^e\right] \ .$$

Now, in the equilibrium point we have $P^eX^e = P_MM^e$, which gives $P_M/X^e = P^e/M^e$, and so

$$X^e\left[1+\frac{P^e}{X^e}\left(\frac{\mathrm{d}X}{\mathrm{d}P}\right)^e - \frac{P_M}{X^e}\left(\frac{\mathrm{d}M}{\mathrm{d}P}\right)^e\right] = X^e\left[1+\frac{P^e}{X^e}\left(\frac{\mathrm{d}X}{\mathrm{d}P}\right)^e - \frac{P^e}{M^e}\left(\frac{\mathrm{d}M}{\mathrm{d}P}\right)^e\right] \ ,$$

from which

$$\alpha \equiv X^e\left[1-\eta_X^e-\eta_M^e\right] \ ,$$

where

$$\eta_X^e \equiv -\frac{P^e}{X^e}\left(\frac{\mathrm{d}X}{\mathrm{d}P}\right)^e \ , \qquad \eta_M^e \equiv \frac{P^e}{M^e}\left(\frac{\mathrm{d}M}{\mathrm{d}P}\right)^e$$

[12] The same condition would hold also if, instead of a lag, we postulated a continuous adjustment process in the price level of the type

$$\frac{\mathrm{d}P}{\mathrm{d}t} = c(QV - PY) \ , \qquad c > 0 \ .$$

Differentiating and using (27.63) we would obtain, after linearization, the second-order differential equation

$$\frac{\mathrm{d}^2\overline{P}}{\mathrm{d}t^2} + cY\frac{\mathrm{d}\overline{P}}{\mathrm{d}t} - cV\alpha\overline{P} = 0 \ ,$$

whose characteristic equation is

$$\lambda^2 + cY\lambda - cV\alpha = 0 \ ,$$

and the stability condition turns out to be $\alpha < 0$ again.

are, respectively, the elasticity of exports and of imports with respect to P. Since $X^e > 0$, the condition $\alpha < 0$ is equivalent to

$$1 - \eta_X^e - \eta_M^e < 0 \ ,$$

that is,

$$\eta_X^e + \eta_M^e > 1 \ ,$$

which coincides with the so-called 'Marshall-Lerner' condition.

The condition $-\alpha < 1/2\pi Y/V\omega$ can be written, after simple manipulation, as

$$\eta_X^e + \eta_M^e < 1 + \frac{Y}{X^eV\omega}\frac{\pi}{2} \ .$$

Thus, finally, we can rewrite (27.69) as

$$1 < \eta_X^e + \eta_M^e < 1 + \frac{Y\pi}{2X^eV\omega} \ . \tag{27.70}$$

The economic interpretation is straightforward: the sum of the elasticities must not only be greater than 1 (and this is the usual condition, well-known in monetary international economics), but also *smaller* than another critical value (of course, the latter is greater than 1), as given by (27.70). Therefore, instability might occur not only because the sum of the elasticities is too small, but also because such a sum is *too big*.

27.8 Exercises

1. Consider the neoclassical aggregate model (Chap. 13, Sect. 13.2) and suppose that it takes some time for the homogeneous output, when saved and invested, to give rise to an increase in the capital stock. This means that the fundamental dynamic equation becomes

$$r'(t) = sf(r(t-\omega)) - nr(t),$$

where without loss of generality we can set $\omega = 1$. Examine the local stability of the steady-state equilibrium point (i.e., where $r'(t) = 0$) and show that with a well-behaved production function the equilibrium is always (locally) stable, but the approach to equilibrium will be oscillatory. Exemplify with a Cobb-Douglas production function (Hint: see Fig. 13.2; at the equilibrium point, $s\mathrm{d}f/\mathrm{d}r < n$. Then use Burger's theorem).

2. Suppose that, in an imperfectly competitive market, firms adjust prices toward a partial equilibrium level determined by a markup equation, namely

$$\begin{aligned} p'(t) &= g[\hat{p}(t) - p(t)], \quad g > 0, \\ \hat{p}(t) &= \beta w(t), \quad \beta = mA > 0, \end{aligned}$$

where g is the speed of adjustment (on partial adjustment equations see Chap. 12, Sect. 12.4), $(m-1)$ is the markup over direct labour costs per unit of output, A a coefficient related to the average productivity of labour, and w the money wage rate. In the short rum we may take m and A as constant. Further suppose that the money wage rate is fully indexed to the price p (for example through an escalator clause with a quarterly lag) according to the equation

$$w(t) = \mu p(t-1),$$

where μ is the agreed-upon real wage rate and is also given in the short run.

(2.a) Show that the condition for avoiding the possibility of prices converging to zero is $\beta\mu = 1$.

(2.b) Show that, when $\beta\mu = 1$, there is only one real root which is a simple root of the characteristic equation.

(2.c) Show that, when $\beta\mu = 1$, the oscillations of prices will be damped.

3. In the model examined in the previous exercise, show that an increase in the speed of adjustment g will cause (i) a *decrease* in the frequency of each oscillation, and (ii) an *increase* in the amplitude of each oscillation (Hint: both exercises are based on Padoan and Petit, 1978, Sect. 3).

27.9 References

Allen, R.G.D., 1959, *Mathematical Economics,* Chap. 8, Sects. 8.4, 8.5.

Allen, R.G.D., 1967, *Macro-Economic Theory: A Mathematical Treatment,* Chap. 19, Sects. 19.3, 19.4.

Bellman, R. and K.L. Cooke, 1963, *Differential-Difference Equations,* Chaps.3, 4, 13.

Bergstrom, A.R., 1967, *The Construction and Use of Economic Models*, pp. 1-2.

Bergstrom, A.R. (ed.), 1976, *Statistical Inference in Continuous Time Models.*

Bergstrom, A.R., 1984, Continuous Time Stochastic Models and Issues of Aggregation over Time.

Boole, G., 1960, *A Treatise on the Calculus of Finite Differences*, pp. 277-288.

Burger, E., 1956, On the Stability of Certain Economic Systems.

Burmeister, E. and S.J. Turnovsky, 1976, The Specification of Adaptive Expectations in Continuous Time Dynamic Economic Models.

Busenberg, S. and M. Martelli (eds.), 1991, *Delay Differential Equations and Dynamical Systems.*

Chukwu, E.N., 1991, Global Behaviour of Linear Retarded Functional Differential Equations.

Chukwu, E.N., 1992, Mathematical Controllability Theory of Capital Growth of Nations.

Frisch, R.., 1933, Propagation Problems and Impulse Problems in Dynamic Economics.

Frisch, R. and H. Holme, 1935, The Characteristic Solution of a Mixed Difference and Differential Equation Occurring in Economic Dynamics.

Furuno, Y., 1965, The Period of Production in Two-Sector Models of Economic Growth.

Gandolfo, G., 1981, *Qualitative Analysis and Econometric Estimation of Continuous Time Dynamic Models.*

Gandolfo, G., 1993, *Continuous Time Econometrics: Theory and Applications.*

Goodwin, R.M., 1948, Secular and Cyclical Aspects of the Multiplier and the Accelerator.

Hale, J.K., 1979, Nonlinear Oscillations in Equations with Delay.

Jarsulic, M., 1993, Complex Dynamics in a Keynesian Growth Model.

Kalecki, M., 1933, Outline of Theory of the Business Cycle.

Kalecki, M., 1935, A Macrodynamic Theory of Business Cycles.

Kalecki, M., 1954, *Theory of Economic Dynamics.*

Koopmans, T. C., 1950, Models Involving a Continuous Time Variable.

Leontief, W., et. al., 1953, *Studies in the Structure of the American Economy,* Chap. 3, Mathematical Note 2.

Marschak, J., 1950, Statistical Inference in Economics: An Introduction

Padoan, P.C. and Petit, M.L., 1978, Inflazione in un'economia indicizzata: alcune generalizzazioni.

Tinbergen, J., 1959, *Selected Papers*, pp. 1-14.

Wen, K.-H., 1995, Continuous-Time Chaos in Stock-Market Dynamics.

Wen, K.-H., P. Chen and Z. Zhang, 1992, Bifurcation and Chaos in the Soft-Bouncing Oscillator.

Wymer, C.R., 1972, Econometric Estimation of Stochastic Differential Equation Systems.

Zambelli, S., 1992, The Wooden Horse that wouldn't Rock: Reconsidering Frisch.

Bibliography

Allen, R.G.D., 1959, *Mathematical Economics,* 2nd edition, London, Macmillan.

Allen, R.G.D., 1967, *Macro-Economic Theory: A Mathematical Treatment,* London, Macmillan.

American Economic Association and Royal Economic Society, 1966, *Surveys of Economic Theory,* Vols. I, II, III, London, Macmillan.

Anderson, P., K.J. Arrow and D. Pines (eds.), 1988, *The Economy as an Evolving Complex System* (Santa Fe Institute Studies in the Sciences of Complexity, Vol. V), Redwood City, CA, Addison-Wesley.

Andronov, A.A., A.A. Vitt and S.E. Khaikin, 1966, *Theory of Oscillators,* London, Pergamon Press.

Aoki, M., 1994, New Macroeconomic Modeling Approaches: Hierarchical Dynamics and Mean-Field Approximation, *Journal of Economic Dynamics and Control* 18, 865-877.

Arneodo, A., P. Coullet and C. Tresser, 1982, Oscillations with Chaotic Behavior: An Illustration of a Theorem by Shil'nikov, *Journal of Statistical Physics* 27, 171-182.

Arrow, K.J., 1962, The Economic Implications of Learning by Doing, *Review of Economic Studies* 29, 155-173. Reprinted in: R. Becker and E. Burmeister (eds.), 1991, *Growth Theory,* Vol. I, Chap. 17.

Arrow, K.J., 1974, Stability Independent of Adjustment Speed, in: G. Horwich and P.A. Samuelson (eds.), *Trade, Stability, and Macroeconomics: Essays in Honour of L.A. Metzler,* New York, Academic Press, 181-202.

Arrow, K.J., H.D. Block and L. Hurwicz, 1959, On the Stability of Competitive Equilibrium, II, *Econometrica* 27, 82-109.

Arrow, K.J. and F.H. Hahn, 1971, *General Competitive Analysis,* San Francisco, Holden-Day, Inc., and Edinburgh, Oliver and Boyd; Amsterdam, North-Holland.

Arrow, K.J. and L. Hurwicz, 1958, On the Stability of Competitive Equilibrium, I, *Econometrica* 26, 448-454.

Arrow, K.J., S. Karlin and P. Suppes (eds.), 1960, *Mathematical Methods in the Social Sciences, 1959,* Stanford, Stanford University Press.

Arrow, K.J. and M. McManus, 1958, A Note on Dynamic Stability, *Econometrica* 26, 522-552.

Atkinson, A.B., 1969, The Timescale of Economic Models: How Long is the Long Run?, *Review of Economic Studies* 36, 137-152. Reprinted in: F.H. Hahn (ed.), 1971, *Readings in the Theory of Growth,* 248-263.

Barnett, S., 1973, New Reductions of Hurwitz Determinants, *International Journal* *of Control* 18, 977-991.

Barnett, S., 1974, Applications of the Routh Array to the Stability Discrete-Time Linear System, *International Journal of Control* 19, 47-55.

Barnett, S. and C. Storey, 1970, *Matrix Methods in Stability Theory,* London, Nelson.

Barnett, S., 1983, *Polynomials and Linear Control Systems,* New York and Basel, M. Dekker Inc.

Barnett, W.A. and P. Chen, 1988, The Aggregation-Theoretic Monetary Aggregates are Chaotic and Have Strange Attractors: An Econometric Application of Mathematical Chaos, in W.A. Barnett, E.R. Berndt, H. White (eds.), *Dynamic Econometric Modeling,* Cambridge University Press, 199-245.

Barnett, W.A., J. Geweke and K. Shell (eds.), 1989, *Economic Complexity: Chaos, Sunspots, Bubbles, and Nonlinearity,* Cambridge University Press.

Barnett, W.A., G. Gandolfo and C. Hillinger (eds.), 1996, *Dynamic Disequilibrium Modelling,* Cambridge University Press.

Barnett, W.A., M. Salmon and A. Kirman (eds.), 1996, *Nonlinear Dynamics in Economics,* Cambridge University Press.

Barro, R.J. and X. Sala-i-Martin, 1992, Convergence, *Journal of Political Economy* 100, 223-251.

Barro, R.J. and X. Sala-i-Martin, 1995, *Economic Growth,* New York, McGraw-Hill.

Bassett, L., 1968, The Solution of Qualitative Comparative Static Problems: Comment, *Quarterly Journal of Economics* 82, 519-523.

Bassett, L., J. Maybee and J. Quirk, 1968, Qualitative Economics and the Scope of the Correspondence Principle, *Econometrica* 36, 544-563.

Baumol, W.J., 1958, Topology of Second Order Linear Difference Equations with Constant Coefficients, *Econometrica* 26, 258-285.

Baumol, W.J., 1970, *Economic Dynamics,* 3rd edn., New York, Macmillan.

Baumol, W.J., and R.E. Quandt, 1964, Rules of the Thumb and Optimally Imperfect Decisions, *American Economic Review* 54, 23-46.

Baumol, W.J. and J. Benhabib, 1989, Chaos: Significance, Mechanism, and Economic Applications, *Journal of Economic Perspectives* 3, 77-105.

Baumol, W.J. and R.E. Quandt, 1985, Chaos Models and their Implications for Forecasting, *Eastern Economic Journal* 11, 3-15.

Bear, D.V.T., 1966, Distributed Lags and Economic Theory, *Review of Economic Studies* 33, 235-243.

Beavis, B. and I. Dobbs, 1990, *Optimization and Stability Theory for Economic Analysis,* Cambridge (UK), Cambridge University Press.

Becker, R. and E. Burmeister (eds.), 1991, *Growth Theory, Vol. I, Vol. II, Vol. III,* Aldershot, England, E. Elgar.

Beckmann, M.J. and H.E. Ryder, 1969, Simultaneous Price and Quantity Adjustment in a Single Market, *Econometrica* 37, 470- 484.

Beckmann, M.J. and J.P. Wallace, 1967, Marshallian Versus Walrasian Stability, *Kyklos* 20, 935-948.

Bellman, R., 1953, *Stability Theory of Differential Equations,* New York, McGraw-Hill.

Bellman, R. and K.L. Cooke, 1963, *Differential-Difference Equations,* New York, McGraw-Hill

Beltrami, E., 1987, *Mathematics for Dynamic Modeling,* New York, Academic Press.

Benhabib, J. (ed.), 1992, *Cycles and Chaos in Economic Equilibrium,* Princeton, Princeton University Press.

Benhabib, J. and R. Day, 1981, Rational Choice and Erratic Behaviour, *Review of Economic Studies* 48, 459-471.

Benhabib, J. and R. Day, 1982, A Characterization of Erratic Dynamics in the Overlapping Generations Model, *Journal of Economic Dynamics and Control* 4, 37-55.

Benhabib, J. and K. Nishimura, 1979, The Hopf Bifurcation and Existence and Stability of Closed Orbits in Multisector Models of Optimal Economic Growth, *Journal of Economic Theory* 21, 421-444. Reprinted in: R. Becker and E. Burmeister (eds.), 1991, *Growth Theory,* Vol. II, Chap. 7.

Bergstrom, A.R. (ed.), 1976, *Statistical Inference in Continuous Time Models,* Amsterdam, North-Holland.

Bergstrom, A.R., 1967, *The Construction and Use of Economic Models,* London, The English Universities Press.

Bergstrom, A.R., 1984, Continuous Time Stochastic Models and Issues of Aggregation over Time, in Z. Griliches and M. D. Intriligator (eds.), *Handbook of Econometrics,* Vol. 2, Chap. 20, Amsterdam, North-Holland.

Biederman, D.K., 1993, Permanent Income and Long-run Stability in a Generalized Multiplier/Accelerator Model, *Journal of Macroeconomics* 15, 240-72.

Black, J., 1962, The Technical Progress Function and the Production Function, *Economica* (n.s.) 29, 166-170.

Blatt, J.M., 1983, *Dynamic Economic Systems - A Post Keynesian Approach,* Armonk, New York, Sharpe.

Blinder, A.S. and L.J. Maccini, 1991, The Resurgence of Inventory Research: What Have We Learned?, *Journal of Economic Surveys* 5, 291-328.

Boldrin, M. and M. Woodford, 1990, Equilibrium Models Displaying Endogenous Fluctuations and Chaos, *Journal of Monetary Economics* 25, 189-222.

Boole, G., 1960 (1872), *A Treatise on the Calculus of Finite Differences,* New York, Dover Publications (reprint of the 1872 second edition, London, Macmillan).

Bothwell, F.E., 1952, The Method of Equivalent Linearization, *Econometrica* 20, 269-283.

Boulding, K.E. and G.J. Stigler (eds.), 1953, *Readings in Price Theory,* London, Allen &Unwin.

Braun, M., 1986, *Differential Equations and their Applications,* 3rd ed., Berlin and New York, Springer-Verlag.

Brauer, A., 1964, On the Characteristic Roots of Non-Negative Matrices, in: Schneider H., ed., 1964, *Recent Advances in matrix Theory,* 3-38.

Brock, W.A., 1986, A Revised Version of Samuelson's Correspondence Principle: Application of Recent Results on the Asymptotic Stability of Optimal Control to the Problem of Comparing Long-Run Equilibrium, in: H. Sonnenschein (ed.), *Models of Economic Dynamics,* Berlin and New York, Springer-Verlag, pp. 86-116.

Brock, W.A. and J.A. Scheinkman, 1975, Some Results on Global Asymptotic Stability of Difference Equations, *Journal of Economic Theory* 10, 265-268.

Brock, W.A., D.A. Hsieh and B. LeBaron, 1991, *Nolinear Dynamics, Chaos, and Instability: Statistical Theory and Economic Evidence,* Cambridge (MA, USA), MIT Press.

Brock, W.A. and A.G. Malliaris, 1989, *Differential Equations, Stability and Chaos in Dynamic Economics,* Amsterdam, North-Holland.

Bródy, A. and M. Farkas, 1987, Forms of Economic Motion, *Acta Oeconomica* 38, 361-370.

Buchanan, N.S., 1939, A Reconsideration of the Cobweb Theorem, *Journal of Political Economy* 47, 67-81.

Buiter, W.H., 1984, Saddlepoint Problems in Continuous-Time Rational Expectations Models: A General Method and some Macroeconomic Examples, *Econometrica* 62, 665-680. Reprinted in: Buiter, W., 1990, *International Macroeconomics,* Oxford, Oxford University Press, Chap. 1.

Buiter, W.H. and M. Miller, 1981, Monetary Policy and International Adjustment: The Problems of Adjustment, *Oxford Economic Papers* 33 (Supplement), 143-175. Reprinted in: Buiter, W., 1990, *International Macroeconomics,* Oxford, Oxford University Press, Chap. 2.

Burger, E., 1956, On the Stability of Certain Economic Systems, *Econometrica* 24, 488-493.

Burley, P. and J. Foster (eds.), 1994, *Economics and Thermodynamics: New Perspectives on Economic Analysis,* Dordrecht, Kluwer.

Burmeister, E., 1967, The Existence of Golden Ages and Stability in the Two-Sector Models, *Quarterly Journal of Economics* 81, 146-154.

Burmeister, E., 1968, The Role of the Jacobian Determinant in the Two-Sector Model, *International Economic Review* 9, 195-203. Reprinted in: R. Becker and E. Burmeister (eds.), 1991, *Growth Theory,* Vol. I, Chap. 7.

Burmeister, E. and A.R. Dobell, 1970, *Mathematical Theories of Economic Growth,* London, Macmillan.

Burmeister, E. and S.J. Turnovsky, 1976, The Specification of Adaptive Expectations in Continuous Time Models, *Econometrica* 44, 879-905.

Busenberg, S. and M. Martelli (eds.), 1991, *Delay Differential Equations and Dynamical Systems,* Berlin and New York, Springer-Verlag.

Cagan, P., 1956, The Monetary Dynamics of Hyperinflation, in: M. Friedman (ed.), *Studies in the Quantity Theory of Money,* Chicago, University of Chicago Press.

Cartigny, P. and A. Venditti, 1994, Turnpike Theory: Some New Results on the Saddle Point Property of Equilibria and on the Existence of Endogenous Cycles, *Journal of Economic Dynamics and Control* 5, 957-974.

Casti, J., 1989, *Alternate Realities: Mathematical Models of Nature and Man,* New York, Wiley.

Chang, W.W. and D.J. Smyth, 1971, The Existence and Persistence of Cycles in a Non-linear Model: Kaldor's 1940 model Re-examined, *Review of Economic Studies* 38, 37-44.

Chen, P., 1988, Empirical and Theoretical Evidence of Economic Chaos, *System Dynamics Review* 4, 81-108.

Chen, P., 1993, Searching for Economic Chaos: A Challenge to Econometric Practice and Nonlinear Tests, in: R.H. Day and P. Chen (eds.), *Nonlinear Dynamics and Evolutionary Economics.*

Chiarella, C., 1988, The Cobweb Model: Its Instability and the Onset of Chaos, *Economic Modelling* 5, 377-384.

Chiarella, C., 1990, *The Elements of a Nonlinear Theory of Economic Dynamics,* Berlin and New York, Springer-Verlag.

Chipman, J.S., 1950, The Multi-Sector Multiplier, *Econometrica* 18, 355-374.

Chukwu, E.N., 1991, Global Behaviour of Linear Retarded Functional Differential Equations, *Journal of Mathematical Analysis and Applications* 162, 277-293.

Chukwu, E.N., 1992, Mathematical Controllability Theory of Capital Growth of Nations, *Applied Mathematics and Computation,* 52, 317-344.

Clower, R.W. and D.W. Bushaw, 1954, Price Determination in a Stock-Flow Economy, *Econometrica* 22, 238-243.

Coddington, E.A. and N. Levinson, 1955, *Theory of Ordinary Differential Equations,* New York, McGraw-Hill.

Collar, A.R. and A. Simpson, 1987, *Matrices and Engineering Dynamics,* Wiley, New York.

Conlisk, J., 1973, Quick Stability Checks and Matrix Norms, *Economica* (n.s.) 40, 402-409.

Courant, R., 1961(1937), *Differential and Integral Calculus,* Vol. I and Vol. II, London, Blackie & Son.

Cournot, A., 1899 (1838), *Researches into the Mathematical Principles of the Theory of Wealth* (transl. of the 1838 French edition), New York, Macmillan.

Creedy, J. and V. L. Martin (eds.), 1994, *Chaos and Non-linear Models in Economics: Theory and Applications,* Aldershot, U.K., Elgar.

Day, R., 1982, Irregular Growth Cycles, *American Economic Review* 72, 406-414.

Day, R.H. and P. Chen (eds.), 1993, *Nonlinear Dynamics and Evolutionary Economics,* Oxford, Oxford University Press.

Day, R., 1994, *Complex Economic Dynamics, Vol. I: An Introduction to Dynamical Systems and Market Mechanisms,* Cambridge (Mass.), MIT Press.

Deardoff, A.V., 1970, Growth Paths in the Solow Neoclassical Growth model, *Quarterly Journal of Economics* 84, 134-139.

Debreu, G., 1974, Excess Demand Functions, *J. Mathematical Economics* 1, 15-21.

DeGrauwe, P., H. Dewachter and M. Embrechts, 1993, *Exchange Rate Theory: Chaotic Models of Foreign Exchange Markets,* Oxford, Basil Blackwell.

Desai, M., 1973, Growth Cycles and Inflation in a Model of the Class Struggle, *Journal of Economic Theory* 6, 527-545.

Diener, M. and T. Poston, 1984, The Perfect Delay Convention, or The Revolt of the Slaved Variables, in: H. Haken (ed.), *Chaos and Order in Nature,* 2nd edition, Berlin Heidelberg New York, Springer-Verlag, 249-268.

Dockner, E.J. and G. Feichtinger, 1991, On the Optimality of Limit Cycles in Dynamic Economic Systems, *Journal of Economics* 51, 31-50.

Dohtani, A.M., 1993, Global Stability of the Competitive Economy Involving

Complementary Relations among Commodities, *Journal of Mathematical Economics* 22, 73-83.
Domar, E.D., 1946, Capital Expansion, Rate of Growth and Employment, *Econometrica* 14, 137-147. Reprinted in: Domar, E.D., 1957, *Essays in the Theory of Economic Growth,* 70-82; Stiglitz, J.E. and H. Uzawa, (eds.), 1969, *Readings in the Modern Theory of Growth,* 34-44.
Domar, E.D., 1957, *Essays in the Theory of Economic Growth,* New York, Oxford University Press.
Dore, M.H.I., 1993, *The Macrodynamics of Business Cycles,* Cambridge and Oxford, Blackwell.
Dornbusch, R., 1976, Expectations and Exchange Rate Dynamics, *Journal of Political Economy* 84, 1161-1176.
Dornbusch, R., 1980, *Open Economy Macroeconomics*, New York, Basic Books.
Drandakis, E.M., 1963, Factor Substitution in the Two-Sector Growth Model, *Review of Economic Studies* 30, 217-228.
Edgeworth, F.Y., 1881, *Mathematical Psychics,*, London, Kegan Paul (reprinted 1967, New York, Kelley).
Enthoven, A.C. and K.J., Arrow, 1956, A Theorem on Expectations and the Stability of Equilibrium, *Econometrica* 24, 288-293.
Ezekiel, M., 1938, The Cobweb Theorem, *Quarterly Journal of Economic* 52, 255-288. Reprinted in: Haberler, G. (ed.), 1950, *Readings in Business Cycle Theory,* 422-42.
Farebrother, R.W., 1973, Simplified Samuelson Conditions for Cubic and Quartic Equations, *The Manchester School of Economic and Social Studies* 41, 396-400.
Farley, A.M. and K.-P. Lin, 1990, Qualitative Reasoning in Economics, *Journal of Economic Dynamics and Control* 14, 465-490.
Feichtinger, G., 1992, Hopf Bifurcation in an Advertising Diffusion Model, *Journal of Economic Behavior and Organization* 17, 401-411.
Fellner W. and B.F. Haley (eds.), 1950, *Readings in the Theory of Income Distribution,* London, Allen & Unwin.
Fiorito, R. (ed.), 1994, *Inventory, Business Cycles and Monetary Transmission*, Berlin and New York, Springer-Verlag.
Fisher, F.M., 1962, An Alternate Proof and Extension of Solow's Theorem of Non-negative Square Matrices, *Econometrica* 30, 349-350.
Fisher, F.M., 1965, Choice of Units, Column Sums, and Stability in Linear Dynamic Systems with Non-negative Square Matrices, *Econometrica* 33,445-450.
Fisher, F.M., 1983, *Disequilibrium Foundations of Equilibrium Economics,* Cambridge (UK), Cambridge University Press.
Fisher, M.E. and A.T. Fuller, 1958, On the Stabilization of Matrices and Convergence of Linear Iterative Processes, *Proceedings of the Cambridge Philosophical Society* 54, 417-425.
Foley, D., 1992, A Contribution to the Theory of Business Cycles, *Quarterly Journal of Economics* 107, 1071-1088.
Folsom, R.N., D.C. Boger and H.C. Mullikin, 1976, Stability Conditions for Linear

Constant Coefficient Difference Equations in Generalized Difference Form, *Econometrica* 44, 575-591.
Forsyth, A.R., 1890-1906, *Theory of Differential Equations*, Vols. I-IV, Cambridge, Cambridge University Press (reprinted 1956, New York, Dover Publications).
Foster, J., 1994, The Self-Organisation Approach in Economics, in P. Burley and J. Foster (eds.), 183-201.
Franke, R. and T. Lux, 1993, Adaptive Expectations in a Nonlinear Metzlerian Model of the Inventory Cycle, *Scandinavian Journal of Economics* 28, 355-364.
Frankel, J.A. and K.A. Froot, 1988, Chartists, Fundamentalists, and the Demand for Dollars, *Greek Economic Review* 10, 49-102.
Frazer, R.A., W.J. Duncan and A.R. Collar, 1938, *Elementary Matrices and some Applications to Dynamics and Differential Equations,* London, Cambridge University Press.
Frisch, R., 1933, Propagation Problems and Impulse Problems in Dynamic Economics, in: *Economic Essay in Honour of Gustav Cassel,* London, Allen & Unwin, (reprinted in: Gordon, R.A. and L.R. Klein (eds.), 1966, *Readings in Business Cycles,* 155-185).
Frisch, R., 1936, On the Notion of Equilibrium and Disequilibrium, *Review of Economic Studies* 3, 100-105.
Frisch, R., 1963, Parametric Solution and Programming of the Hicksian Model, in Rao, C.R., et al., (eds.), *Essays on Econometrics and Planning* (in Honour of P.C. Mahalanobis), London, Pergamon Press, 45-82.
Frisch, R. and H. Holme, 1935, The Characteristic Solutions of a Mixed Difference and Differential Equation Occurring in Economic Dynamics, *Econometrica* 3, 225-239.
Fujimoto, T. and P. Indelli, 1986, A Note on Quick Stability Checks for Linear Difference Equations, *Economic Studies Quarterly* 37, 81-86.
Fuller, A.T., 1968, Conditions for a Matrix to Have Only Characteristic Roots with Negative Real Part, *Journal of Mathematical Analysis and Applications* 23, 71-98.
Furuno, Y., 1965, The Period of Production in Two-Sector Models of Economic Growth, *International Economic Review* 6, 240-244.
Gabisch, G. and H.-W. Lorenz, 1989, *Business Cycle Theory: A Survey of Methods and Concepts,* 2nd ed., Berlin and New York, Springer-Verlag.
Gale, D., 1973, Pure Exchange Equilibrium of Dynamic Economic Models, *Journal of Economic Theory* 6, 12-36.
Gale, D. and H. Nikaidô, 1965, The Jacobian Matrix and Global Univalence of Mappings, *Mathematische Annalen* 159, 81-93. Reprinted in: Newman, P.K. (ed.), 1968, *Readings in Mathematical Economics,* Vol. I, 68-80.
Galeotti M. and F. Gori, 1989, Uniqueness of Periodic Orbits in Liénard-type Business Cycle Models, *Metroeconomica* 40, 135-146.
Gandolfo, G., 1971, *Mathematical Methods and Models in Economic Dynamics,* Amsterdam, North-Holland.
Gandolfo, G., 1981, *Qualitative Analysis and Econometric Estimation of Continuous Time Dynamic Models,* Amsterdam, North-Holland.

Gandolfo, G., 1992, Sensitivity Analysis in Continuous Time Econometric Models, *Computers and Mathematics with Applications* 24, 43-55.

Gandolfo, G. (ed.), 1993, *Continuous Time Econometrics: Theory and Applications,* London, Chapman & Hall.

Gandolfo, G., 1994, *International Economics I: The Pure Theory of International Trade,* Berlin and New York, Springer-Verlag.

Gandolfo, G., 1995, *International Economics II: International Monetary Economics and Open-Economy Macroeconomics*, Berlin and New York, Springer-Verlag.

Gandolfo, G. and P. C. Padoan, 1990, The Italian Continuous Time Model: Theory and Empirical Results, *Economic Modelling* 7, 91-132.

Gandolfo, G., P.C. Padoan, G. De Arcangelis and C.R. Wymer, 1996, Nonlinear Estimation of a Nonlinear Continuous Time Model, in: W.A. Barnett, G. Gandolfo and C. Hillinger (eds.), *Dynamic Disequilibrium Modelling,* Cambridge, Cambridge University Press.

Gantmacher, F.R., 1959, *Applications of the Theory of Matrices,* New York, Interscience.

Georgescu-Roegen, N., 1951, Relaxation Phenomena in Linear Dynamic Models, in: T.C. Koopmans (ed.), *Activity Analysis of Production and Allocation,* New York, Wiley, 116-131.

Gillespie, R.P., 1959, *Integration,* Edinburgh, Oliver and Boyd.

Goldberg, S., 1958, *Introduction to Difference Equations,* New York, John Wiley.

Goodwin, R.M., 1947, Dynamical Coupling with Especial Reference to Markets Having Production Lags, *Econometrica* 15, 181-204.

Goodwin, R.M., 1948, Secular and Cyclical Aspects of the Multiplier and the Accelerator, in: *Income, Employment and Public Policy: Essays in Honor of A.H. Hansen* (various authors), New York, Norton, 108-132.

Goodwin, R.M., 1951, The Nonlinear Accelerator and the Persistence of Business Cycles, *Econometrica* 19, 1-17.

Goodwin, R.M., 1955, A Model of Cyclical Growth. In: Lundberg, E. (ed.), *The Business Cycle in the Post-War World,* London Macmillan, 203-21. Reprinted in: Gordon, R.A. and L.R. Klein, (eds.), 1966, *Readings in Business Cycles,* 6-22.

Goodwin, R.M., 1965, 1967, A Growth Cycle, paper presented at the First World Congress of the Econometric Society, Rome, 1965. Published in: Feinstein, C.H. (ed.), 1967, *Socialism, Capitalism and Economic Growth* (Essays presented to Maurice Dobb), London, Cambridge University Press, 54-58.

Goodwin, R.M., 1982, *Essays in Economic Dynamics,* London, Macmillan.

Goodwin, R.M., 1990, *Chaotic Economic Dynamics,* Oxford, Oxford University Press.

Goodwin, R.M., 1991, Non-linear Dynamics and Economic Evolution, in Thygesen et al. eds., 424-442.

Goodwin, R.M., M. Krüger and A. Vercelli (eds.), 1984, *Nonlinear Models of Fluctuating Growth,* Berlin and New York, Springer-Verlag.

Gordon, R.A. and L.R., Klein (eds.), 1966, *Readings in Business Cycles*, London, Allen & Unwin.

Grandmont, J.-M., 1985, On Endogenous Competitive Business Cycles, *Econometrica* 22, 995-1037.

Grandmont, J.-M., 1986, Periodic and Aperiodic Behaviour in Discrete One-Dimensional Dynamical Systems, in: W. Hildenbrand and A. Mas-Colell (eds.), *Contributions to Mathematical Economics,* Amsterdam, North-Holland, pp. 227-65.

Grandmont, J.-M. (ed.), 1988, *Nonlinear Economic Dynamics,* New York, Academic Press.

Granger, C.W.J., 1994, Is Chaotic Economic Theory Relevant for Economics?, *Journal of International and Comparative Economics* 3, 139-145.

Grasman, J. and J.J. Wentzel, 1994, Co-existence of a Limit Cycle and an Equilibrium in Kaldor's Business Cycle Model and Its Consequences, *Journal of Economic Behavior and Organization* 24, 369-377.

Guckenheimer, J. and P. Holmes, 1986, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields,* Berlin and New York, Springer-Verlag.

Hababigahi, H. and Quirk, J., 1973, Hicksian Stability and Walras' Law, *Review of Economic Studies* 40, 249-258.

Haberler, G. (ed.), 1950, *Readings in Business Cycle Theory,* London, Allen & Unwin.

Hadar, J., 1965, A Note on Dominant Diagonals in Stability Analysis, *Econometrica* 33, 442-444.

Hahn, F.H., 1965, On Two-Sector Growth Models, *Review of Economic Studies* 32, 339-346. Reprinted in: A. Sen (ed.), 1970, *Growth Economics: Selected Readings,* 281-292.

Hahn, F.H., 1966, Equilibrium Dynamics with Heterogeneous Capital Goods, *Quarterly Journal of Economics* 80, 633-646. Reprinted in: R. Becker and E. Burmeister (eds.), 1991, *Growth Theory,* Vol. I, Chap. 8.

Hahn, F.H., 1991, Hicksian Themes on Stability, in L.W. McKenzie and S. Zamagni (eds.), *Value and Capital Fifty Years Later,* New York, New York University Press, pp. 64-83.

Hahn, F.H. (ed.), 1971, *Readings in the Theory of Growth,* London, Macmillan.

Hahn, W., 1967, *Stability of Motion,* Berlin and New York, Springer-Verlag.

Haken, H., 1983a, *Synergetics: An Introduction,* 3rd edition, Berlin and New York, Springer-Verlag.

Haken, H., 1983b, *Advanced Synergetics,* Berlin and New York, Springer-Verlag.

Hale, J.K., 1979, Nonlinear Oscillations in Equations with Delay, in: F.C. Hoppestead (ed.), *Nonlinear Oscillations in Biology,* American Mathematical Society.

Hall, S.G., 1984, On the Solution of High Order, Symmetric, Difference Equations, *Oxford Bulletin of Economics and Statistics* 47, 85-88.

Hamburger, L., 1931, Analogie des fluctuations économiques et des oscillations de relaxation, *Supplément aux Indices du Mouvement des Affaires* 9, 3-36.

Hamburger, L., 1934, Note on Economic Cycles and Relaxation Oscillations, *Econometrica* 2, 112.

Hansen, A.H. and R.V. Clemence (eds.), 1953, *Readings in Business Cycles and National Income,* New York, Norton.

Harrod, R.F., 1939, An Essay in Dynamic Theory, *Economic Journal* 49, 14-33. Reprinted in: Hansen, A.H. and R.V. Clemence (eds.), 1953, *Readings in Business Cycles and National Income,* 200-219; Stiglitz, J.E. and H. Uzawa (eds.), 1969, *Readings in the Modern Theory of Economic Growth,* 14-33.

Harrod, R.F., 1948, *Towards a Dynamic Economics,* London, Macmillan.

Heiner, R.A., 1989, The Origin of Predictable Dynamic Behavior, *Journal of Economic Behavior and Organization* 12, 233-257.

Hicks, J.R., 1937, Mr. Keynes and the 'Classics': A Suggested Interpretation, *Econometrica* 5, 147-159. Reprinted in: Fellner W. and B.F. Haley (eds.), 1950, *Readings in the Theory of Income Distribution,* 461-476; Hicks, J.R., 1967, *Critical Essays in Monetary Theory,* London, Oxford University Press, 126-142.

Hicks, J.R., 1939, *Value and Capital,* 1st edition, London, Oxford University Press.

Hicks, J.R., 1946, *Value and Capital,* 2nd edition, London, Oxford University Press.

Hicks, J.R., 1949, Mr. Harrod's Dynamic Theory, *Economica* (n.s.) 16, 106-121. Reprinted in: Gordon R.A. and L.R. Klein (eds.), 1966, *Readings in Business Cycles,* 23-38.

Hicks, J.R., 1950, *A Contribution to the Theory of Trade Cycle,* London, Oxford University Press.

Hicks, J.R., 1965, *Capital and Growth,* London, Oxford University Press.

Hillinger, C. (ed.), 1992, *Cyclical Growth in Market and Planned Economies,* Oxford, Oxford University Press.

Hirsch, M.W. and S. Smale, 1974, *Differential Equations, Dynamical Systems, and Linear Algebra,* New York, Academic Press.

Hobson, E.W., 1957, *The Theory of Functions of a Real Variable & a Theory of Fourier's Series,* Vols. I, II, New York, Dover Publications Reprint.

Horwich, G. and P.A. Samuelson (eds.), 1974, *Trade, Stability, and Macroeconomics: Essays in Honour of L.A. Metzler,* New York, Academic Press.

Hsu, S.-B. and T.-W. Huang, 1995, Global Stability for a Class of Predator-Prey Systems, *SIAM Journal on Applied Mathematics* 55, 763-783.

Hyslop, J.M, 1959, *Infinite Series,* Edinburgh, Oliver and Boyd.

Ichimura, S., 1955, Toward a General Nonlinear Macrodynamic Theory of Economic Fluctuations. In: Kurihara, K.K. (ed.), *Post-Keynesian Economics,* London, Allen & Unwin, 192-226.

Inada, K.-I., 1963, On a Two-Sector Model of Economic Growth: Comments and a Generalization, *Review of Economic Studies* 30, 105-118.

Inada, K.-I., 1964, On the Stability of Growth Equilibria in Two-Sector Models, *Review of Economic Studies* 31, 127-142.

Ince, E.L., 1956 (1926), *Ordinary Differential Equations,* New York, Dover Publications (reprint of the 1926 edition, Longmans Green).

Intriligator, M.D., 1971, *Mathematical Optimization and Economic Theory,* Amsterdam, North-Holland.

Invernizzi, S. and A. Medio, 1991, On Lags and Chaos in Economic Dynamic Models, *Journal of Mathematical Economics* 20, 521-550.

Jaffé, W., 1967, Walras' Theory of *Tâtonnement*: A Critique of Recent Interpretations, *Journal of Political Economy* 75, 1-19.

Jarsulic, M., 1993, Complex Dynamics in a Keynesian Growth Model, *Metroeconomica* 44, 43-64.

Johnson, C.R., 1974, Sufficient Conditions for D-stability, *Journal of Economic Theory* 9, 53-62.

Jones, L.E. and R.E. Manuelli, 1990, A Convex Model of Equilibrium Growth: Theory and Policy Implications, *Journal of Political Economy* 98, 1008-1038.

Jorgenson, D.W., 1960, A Dual Stability Theorem, *Econometrica* 28, 892-899.

Jorgenson, D.W., 1961a, Stability of a Dynamic Input-Output System, *Review of Economic Studies* 28, 105-116.

Jorgenson, D.W., 1961b, The Structure of Multi-Sectoral Dynamic Models, *International Economic Review* 2, 276-293.

Kaldor, N., 1934, A Classificatory Note on the Determinateness of Equilibrium, *Review of Economic Studies* 1, 122-136. Reprinted as: Kaldor, N., 1960, The Determinateness of Static Equilibrium, *Essays on Value and Distribution,* London, G. Duckworth, 13-33.

Kaldor, N., 1940, A Model of the Trade Cycle, *Economic Journal* 50, 78-92.

Kaldor, N., 1957, A Model of Economic Growth, *Economic Journal* 67, 591-624. Reprinted in: Kaldor, N., 1960, *Essays on Economic Stability and Growth,* London, G. Duckworth, 259-300.

Kaldor, N., 1971, A Comment, *Review of Economic Studies* 38, 45-46.

Kaldor, N. and J.A. Mirrlees, 1961, A New Model of Economic Growth, *Review of Economic Studies* 29, 174-190.

Kalecki, M., 1933, Outline of a Theory of a Business Cycle. In: Kalecki, M., 1969, *Studies in the Theory of Business Cycles, 1933-39,* London, Basil Blackwell, 3-15; and in: Kalecki, M., 1970, *Selected Essays on the Dynamics of Capitalist Economy, 1933-70,* Cambridge, Cambridge University Press, 1-14.

Kalecki, M., 1935, A Macrodynamic Theory of Business Cycles, *Econometrica* 3, 327-344.

Kalecki, M., 1954, *Theory of Economic Dynamics,* London, Allen & Unwin.

Kalman, R.E. and J.E. Bertram, 1960, Control System Analysis and Design Via the 'Second Method' of Lyapunov, I (Continuous-Time Systems) and II (Discrete-Time Systems), *Journal of Basic Engineering* (Trans. ASME), 371-400.

Kamien, M.I. and N.L. Schwartz, 1991, *Dynamic Optimization,* 2nd ed., Amsterdam, North-Holland.

Karlin, S., 1959, *Mathematical Methods and Theory in Games, Programming and Economics,* Vols. I, II, Reading, Massachusetts, Addison-Wesley.

Kelsey, D., 1988, The Economics of Chaos or the Chaos of Economics, *Oxford Economic Papers* 40, 663-693.

Kemp, M.C., Y. Kumura and M. Tawada, 1990, The Global Correspondence Principle, *Economic Letters* 34, 1-4.

Kenkel, J.L., 1974, *Dynamic Linear Economic Models,* New York, Gordon and Breach.

Khalil, H.K., 1980, A New Test for D-stability, *Journal of Economic Theory* 23, 120-122.
Koblo, R., 1991, *The Visible Hand: Synergetic Microfoundations of Macroeconomic Systems,* Berlin and New York, Springer-Verlag.
Koopmans, T. C., 1950, Models Involving a Continuous Time Variable, in: T.C. Koopmans (ed.), *Statistical Inference in Dynamic Economic Models,* New York, Wiley, 384-389.
Kosobud, R.F. and W.D. O'Neil, 1972, Stochastic Implications of Orbital Asymptotic Stability of a Nonlinear Trade Cycle Model, *Econometrica* 40, 69-86.
Krasovskii, N.N., 1963, *Stability of Motion: Applications of Liapunov's Second Method to Differential Systems and Equations with Delay,* Stanford, Stanford University Press.
Kuenne, R.E., 1963, *The Theory of General Economic Equilibrium,* Princeton, Princeton University Press.
Lancaster, K., 1966, The Solution of Qualitative Comparative Static Problems, *Quarterly Journal of Economics* 80, 278-295.
Lancaster, K., 1968, *Mathematical Economics*, New York, Macmillan.
Lange, O., 1952, *Price Flexibility and Employment,* Bloomington, Indiana, Principia Press.
LaSalle, J.P. and S. Lefschetz, 1961, *Stability of Liapunov's Direct Method with Applications,* New York, Academic Press.
Laughton, M.A., 1964, Sensitivity in Dynamical System Analysis, *Journal of Electronics and Control* 17, 577-591.
Le Corbeiller, P., 1933, Les systèmes autoentretenus et les oscillations de relaxation, *Econometrica* 1, 328-332.
Leontief, W., et al., 1953, *Studies in the Structure of the American Economy,* New York, Oxford University Press.
Levinson, N. and O.K. Smith, 1942, A General Equation for Relaxation Oscillations, *Duke Mathematic Journal* 9, 382-403.
Leydesdorff, L. and P. Van den Besselaar (eds.), 1994, *Evolutionary Economics and Chaos Theory: New Directions in Technology Studies*, New York, St. Martin's Press.
Liapounoff, A., 1907, Problème général de la stabilité du mouvement, *Annales de la Faculté des Sciences de Toulouse* (2), 9. Reprinted in: 1949, *Annals of Mathematical Studies,* 17, Princeton University Press. English translation in: Liapunov A.M., 1966, *Stability of Motion,* New York, Academic Press.
Liu, T., C.W.J. Granger and W.P. Heller, 1992, Using the Correlation Exponent to Decide whether an Economic Series is Chaotic, in: M.H. Pesaran and S.M. Potter (eds.), S25-S40.
Lloyd, P.J., 1969, Qualitative Calculus and Comparative Static Analysis, *Economic Record* 45, 343-353.
Lorenz, H.-W., 1987a, Goodwin's Nonlinear Accelerator and Chaotic Motion, *Journal of Economics* 47, 413-418.
Lorenz, H.-W., 1987b, International Trade and the Possible Occurrence of Chaos, *Economic Letters* 23, 135-138.
Lorenz, H.-W., 1993, *Nonlinear Dynamical Economics and Chaotic Motion,* 2nd edition, Berlin and New York, Springer-Verlag (1st ed. 1989).

Lotka, A.J., 1956 (1926), *Elements of Mathematical Biology* (formerly published under the title *Elements of Physical Biology,* London, Constable), (New York, Dover Publications reprint).
Luenberger, D.G., 1979, *Introduction to Dynamic Systems: Theory, Models, & Applications,* New York, Wiley.
Machlup, F., 1943, *International Trade and the National Income Multiplier,* Philadelphia, Blakiston.
Machlup, F., 1959, Statics and Dynamics: Kaleidoscopic Words, *Southern Economic Journal* 26, 91-110. Reprinted in: Machlup, F., 1963, *Essays in Economic Semantics*, Englewood Cliffs, N.J., Prentice-Hall, 4-42.
Malliaris, A.G. and W.A. Brock, 1982, *Stochastic Methods in Economics and Finance,* Amsterdam, North-Holland.
Mandelbrot, B., 1983, *The Fractal Geometry of Nature,* New York, Freeman.
Mankiw, N.G., D. Romer and D.N. Weil, 1992, A Contribution to the Empirics of Economic Growth, *Quarterly Journal of Economics* 108, 407-437.
Mantel, R.R., 1974, On the Characterization of Aggregate Excess Demand, *Journal of Economic Theory* 7, 348-353.
Marotto, F.R., 1978, Snap-Back Repellers Imply Chaos in R^n, *Journal of Mathematical Analysis and Applications* 72, 199-223.
Marrama, V., 1946, Short Notes on a Model of the Trade Cycle, *Review of Economic Studies* 14, 34-40.
Marschak, J., Statistical Inference in Economics: An Introduction, in T.C. Koopmans (ed.), *Statistical Inference in Dynamic Economic Models,* New York, Wiley, 1-50.
Marshall, A., 1879a, *The Pure Theory of Foreign Trade;* 1879b, *The Pure of Domestic Values,* privately printed and circulated in 1879. Reprinted together by: London School of Economics, 1930 (and subsequently), *Reprints of Scarce Tracts on Political Economy,* No. 1, London, Lund Humphries.
Marshall, A., 1920 (1890), *Principles of Economics,* 8th edn., London, Macmillan.
Marshall, A., 1923, *Money, Credit and Commerce,* London, Macmillan.
Mas-Colell, A., 1986, Notes on Price and Quantity Tâtonnement Dynamics, in: H. Sonnenschein (ed.), *Models of Economic Dynamics,* Berlin and New York, Springer-Verlag, 49-68.
McFadden, D., 1968, On Hicksian Stability. In: J.N. Wolfe, (ed.), 1968, *Value, Capital and Growth: Papers in Honour of Sir John Hicks,* Edinburgh University Press, 329-351.
McFadden, D., 1969, On the Controllability of Decentralized Macroeconomic Systems: The Assignment Problem. In: H.W. Kuhn and G.P. Szegö (eds.), 1969, *Mathematical Systems Theory and Economics I,* Berlin and New York, Springer-Verlag, 221-239.
McKenzie, L., 1969, Matrices with Dominant Diagonals and Economic Theory. In: K.J. Arrow, S. Karlin and P. Suppes, (eds.), 1960, *Mathematical Methods in the Social Sciences,* 47-62.
McManus, M., 1962, Dynamic Cournot-type Oligopoly Models–A Correction, *Review of Economic Studies* 29, 337-339.

McManus, M., 1963, Notes on Jorgenson's Model, *Review of Economic Studies* 30, 141-147.

Medio, A., 1984, Synergetics and Dynamic Economic Models, in: R.M. Goodwin, M. Krüger and A. Vercelli (eds.), *Nonlinear Models of Fluctuating Growth,* 166-191.

Medio, A., 1987, Oscillations in Optimal Growth Models, *Journal of Economic Behavior and Organization* 8, 413-427.

Medio, A., 1992, *Chaotic Dynamics: Theory and Applications to Economics,* Cambridge, Cambridge University Press.

Metzler, L.A., 1941, The Nature and Stability of Inventory Cycles, *Review of Economics and Statistics* 23, 113-129. Reprinted in: Gordon R.A. and L.R. Klein (eds.), 1966, *Readings in Business Cycles,* 100-129 (pages refer to this reprint); Metzler, L.A., 1973, *Collected Papers,* 393-427.

Metzler, L.A., 1942, Underemployment Equilibrium in International Trade, *Econometrica* 10, 97-112. Reprinted in: Metzler, L.A., 1973, *Collected Papers,* 258-274.

Metzler, L.A., 1945, Stability of Multiple Markets: The Hicks conditions, *Econometrica* 13, 277-292. Reprinted in: Newman, P.K. (ed.), 1968, *Readings in Mathematical Economics,* Vol. I, 197-212; L.A. Metzler, 1973, *Collected Papers,* 499-515.

Metzler, L.A., 1950, A Multiple Region Theory of Income and Trade, *Econometrica* 18, 329-354. Reprinted in: L.A. Metzler, 1973, *Collected Papers,* 516-544.

Metzler, L.A., 1951, Wealth, Saving and the Rate of Interest, *Journal of Political Economy* 59, 93-116. Reprinted in: L.A. Metzler, 1973, *Collected Papers,* 311-344.

Metzler, L.A., 1973, *Collected Papers,* Harvard University Press.

Milne-Thomson, L.M., 1960 (1933), *The Calculus of Finite Differences,* London, Macmillan.

Minorsky, N., 1962, *Nonlinear Oscillations,* New York, Van Nostrand.

Mirowski, P., 1990, From Mandelbrot to Chaos in Economic Theory, *Southern Economic Journal* 57, 289-307.

Mirowski, P., 1992, Do Economists Suffer from Physics Envy?, *Finnish Economic Papers* 2, 61-68.

Montrucchio, L., 1992, Dynamical Systems that Solve Continuous-Time Concave Optimization Problems: Anything Goes, in: J. Benhabib (ed.), Chap. 12.

Morishima, M., 1964, *Equilibrium Stability, and Growth,* London, Oxford University Press.

Mukherji, A., 1973, On the Sensitivity of Stability Results to the Choice of the Numeraire, *Review of Economic Studies* 40, 427-433.

Mundell, R.A., 1960, The Monetary Dynamics of International Adjustment under Fixed and Flexible Exchange Rates, *Quarterly Journal of Economics* 74, 227-257., Reprinted as Chap. 11 in: Mundell, R.A., 1968, *International Economics,* 152-176.

Mundell, R.A., 1968, *International Economics,* New York, Macmillan.

Murata, Y., 1977, *Mathematics for Stability and Optimization of Economics Systems,* New York, Academic Press.

Muth, J.F., 1960, Optimal Properties of Exponentially Weighted Forecasts, *Journal of the American Statistical Association* 55, 299-306.

Muth, J.F., 1961, Rational Expectations and the Theory of Price Movements, *Econometrica* 29, 315-335.

Negishi, T., 1962, The Stability of a Competitive Economy: A Survey Article, *Econometrica* 30, 635-669.

Negishi, T., 1972, *General Equilibrium Theory and International Trade,* Amsterdam, North-Holland.

Nerlove, M., 1958, Adaptive Expectations and Cobweb Phenomena, *Quarterly Journal of Economics* 73, 227-240.

Newhouse, S., D. Ruelle and F. Takens, 1978, Occurrence of Strange Axiom *A* Attractor near Quasi-Periodic Flows on $T^m, m \geq 3$, *Communications in Mathematical Physics* 64, 35-40.

Newman, P.K., 1959, Some Notes on Stability Conditions, *Review of Economic Studies* 27, 1-9.

Newman, P.K., 1961, Approaches to Stability Analysis, *Economica* (n.s.) 28, 12-29.

Newman, P.K. (ed.), 1968, *Readings in Mathematical Economics,* Vols. I, II, Baltimore, Johns Hopkins Press.

Nikaidô, H., 1968, *Convex Structures and Economic Theory,* New York, Academic Press.

Nusse, H.E. and C.H. Hommes, 1990, Resolution of Chaos with Application to a Modified Samuelson Model, *Journal of Economic Dynamics and Control* 14, 1-19.

Okuguchi, K., 1968, The Labour Participation Ratio and the Speed of Adjustment, *Economica* (n.s.) 35, 445-450.

Okuguchi, K.,1969, On the Stability of Price Adjusting Oligopoly Equilibrium under Product Differentiation, *Southern Economic Journal* 35, 244-246.

Okuguchi, K., 1970, Adaptive Expectations in an Oligopoly Model, *Review of Economic Studies* 37, 233-237.

Okuguchi, K., 1976, *Expectations and Stability in Oligopoly Models,* Berlin and New York, Springer-Verlag.

Okuguchi, K. and K. Irie, 1990, The Schur and Samuelson Conditions for a Cubic Equation, *Manchester School* 58, 414-18.

Okuguchi, K. and F. Szidarovszky, 1988, A Linear Oligopoly Model with Adaptive Expectations: Stability Reconsidered, *Journal of Economics* 48, 79-82.

Okuguchi, K. and F. Szidarovszky, 1990, *The Theory of Oligopoly with Multi-Product Firms,* Berlin, Springer-Verlag.

Olech, C., 1963, On the Global Stability of an Autonomous System on the Plane. In: J.P. LaSalle, J.B. Diaz et al. (eds.), 1963, *Contributions to Differential Equations,* Vol. I, New York, Interscience, 389-400.

Padoan, P.C. and M.L. Petit, 1978, Inflazione in un'economia indicizzata: alcune generalizzazioni, *Moneta e Credito* 31, 193-209.

Patinkin, D., 1952, The Limitations of Samuelson's 'Correspondence Principle', *Metroeconomica* 4, 37-43.

Patinkin, D., 1965, *Money, Interest, and Prices,* 2nd edn., New York, Harper &

Row.
Pesaran, M.H. and S. Potter (eds.), 1992, Nonlinear Dynamics and Econometrics, Special Issue of the *Journal of Applied Econometrics*, Vol. 7, December.
Peters, E.E., 1991, *Chaos and Order in the Capital markets,* New York, Wiley.
Peters, E.E., 1994, *Fractal Market Analysis: Applying Chaos Theory to Investment and Economics,* New York, Wiley.
Petit, M.L., 1990, *Control Theory and Dynamic Games in Economic Policy Analysis,* Cambridge (UK), Cambridge University Press.
Phelps, E., 1966, *Golden Rules of Economic Growth,* New York, Norton.
Phillips, A.W., 1954, Stabilisation Policy in a Closed Economy, *Economic Journal* 64, 290-323.
Phillips, A.W., 1957, Stabilisation Policy and the Time-Form of Lagged Responses, *Economic Journal* 67, 265-277.
Phillips, P.C.B., 1982, A Simple Proof of the Latent Root Sensitivity Formula. *Economic Letters* 9, 57-59.
Poston, T. and I. Stewart, 1978, *Catastrophe Theory and its Applications,* Boston, Pitman.
Poincaré, H., 1908, *Science et méthode,* Paris, Flammarion.
Pol, E., 1993, Theoretical Economics and Catastrophe Theory: An Appraisal, *Australian Economic Papers* 32, 258-271.
Prigogine, I., 1980, *From Being into Becoming: Time and Complexity in the Physical Sciences,* San Francisco, Freeman.
Puu, T., 1987, Complex Dynamics in Continuous Models of the Business Cycle, in D. Batten, J. Casti, B. Johansson (eds.), *Economic Evolution and Structural Change,* Berlin and New York, Springer-Verlag.
Puu, T., 1993, *Nonlinear Economic Dynamics,* 3rd edition, Berlin and New York, Springer-Verlag.
Pyndyck, R.S., 1991, Irreversibility, Uncertainty and Investment, *Journal of Economic Literature* 29, 1110-1148.
Quirk, J.P., 1974, A Class of Generalized Metzlerian Matrices. In: G. Horwich, and P.A. Samuelson, (eds.) 1974, *Trade, Stability, and Macroeconomics: Essays in Honour of L.A. Metzler,* 203-220.
Quirk, J. and R. Saposnick, 1968, *Introduction to General Equilibrium Theory and Welfare Economics,* New York, McGraw-Hill.
Rader, T., 1972, Impossibility of Qualitative Economics: Excessively Strong Correspondence Principles in Production-Exchange Economies, *Zeitschrift für Nationalökonomie* 32, 397-416.
Ramsey, F.P., 1928, A Mathematical Theory of Saving, *Economic Journal* 38, 543-559. Reprinted in: R. Becker and E. Burmeister (eds.), 1991, *Growth Theory,* Vol. II, Chap. 1.
Rand, D., 1978, Exotic Phenomena in Games and Duopoly Models, *Journal of Mathematical Economics* 3, 139-154.
Ritschard, G., 1983, Computable Qualitative Comparative Static Techniques, *Econometrica* 51, 1145-1168.
Romer, P.M., 1994, The Origins of Endogenous Growth, *Journal of Economic Perspectives* 8, 3-22.

Rose, H., 1967, On the Non-Linear Theory of the Employment Cycle, *Review of Economic Studies* 34, 153-173.
Rosser, J.B. Jr., 1991, *From Catastrophe to Chaos: A General Theory of Economic Discontinuities,* Dordrecht, Kluwer.
Ruelle, D., 1979, Strange Attractors, *Mathematical Intelligencer* 2, 126-137.
Samuelson, P.A., 1939, Interactions between the Multiplier Analysis and the Principle of acceleration, *Review of Economics and Statistics* 21, 75-78. Reprinted in: J.E. Stiglitz (ed.), 1966, *Collected Scientific Papers of Paul A. Samuelson,* Vol. II, 1107-1110.
Samuelson, P.A., 1941a, Conditions that the Roots of a Polynomial be less than Unity in Absolute Value, *Annals of Mathematics and Statistics* 12, 360-364. Reprinted in: J.E. Stiglitz (ed.), 1966, *Collected Scientific Papers of Paul A. Samuelson,* Vol. I, 689-693.
Samuelson, P.A., 1941b, The Stability of Equilibrium: Comparative Statics and Dynamics, *Econometrica* 9, 97-120. Reprinted in: J.E. Stiglitz (ed.), 1966, *Collected Scientific Papers of Paul A. Samuelson,* Vol. I, 539-562.
Samuelson, P.A., 1944, The Relation between Hicksian Stability and True Dynamics, *Econometrica* 12, 256-257. Reprinted in: J.E. Stiglitz (ed.), 1966, *Collected Scientific Papers of Paul A. Samuelson,* Vol. I, 563-564.
Samuelson, P.A., 1947, *Foundations of Economic Analysis,* Harvard University Press.
Samuelson, P.A., 1949, Dynamic Process Analysis. In: Ellis, H.S. (ed.), *A Survey of Contemporary Economics,* Philadelphia, Blakiston, Chap. 10. Reprinted in: Stiglitz, J.E., (ed.)1966, *Collected Scientific Papers of Paul A. Samuelson,* Vol. I, 590-625.
Samuelson, P.A., 1972, The General Saddle Point Property of Optimal Control Motions, *Journal of Economic Theory* 5, 102-120.
Sansone, G. and R. Conti, 1964, *Non-Linear Differential Equations,* London, Pergamon Press.
Sargent, T.J., 1993, *Bounded Rationality in Macroeconomics-The Arne Ryde Memorial Lectures,* Oxford, Oxford University Press.
Sasakura, K., 1994, On the Dynamic Behavior of Schinasi's Business Cycle Model, *Journal of Macroeconomics* 16, 423-444.
Sato, R., 1963, Fiscal Policy in a Neo-Classical Growth Model: An Analysis of Time Required for Equilibrating Adjustment, *Review of Economic Studies* 30, 16-23.
Sato, R., 1970, A Further Note on a Difference Equation Recurring in Growth Theory, *Journal of Economic Theory* 2, 95-102.
Sato, R., 1980, Adjustment Time and Economic Growth Revisited, *Journal of Macroeconomics* 2, 239-246.
Saunders, P.T., 1980, *An Introduction to Catastrophe Theory,* Cambridge, Cambridge University Press.
Scarf, H., 1960, Some Examples of Global Instability of the Competitive Equilibrium, *International Economic Review* 1, 157-172.
Scarfe, B.L., 1977, *Cycles, Growth, and Inflation: A Survey of Contemporary Macrodynamics,* New York, McGraw-Hill.

Scheinkman, J.A. and M. Woodford, 1994, Self-Organized Criticality and Economic Fluctuations, *American Economic Review* 84 (Papers and Proceedings), 417-421.
Schinasi, G.J., 1981, A Nonlinear Dynamic Model of Short Run Fluctuations, *Review of Economic Studies* 48, 649-656.
Schinasi, G.J., 1982, Fluctuations in a Dynamic, Intermediate-run IS-LM Model: Applications of Poincaré-Bendixson Theorem, *Journal of Economic Theory* 28, 369-375.
Schneider, H. (ed.), 1964, *Recent Advances in Matrix Theory*, University of Winsconsin Press.
Schoonbeek, L., 1995, The Instantaneous Speed of Adjustment Assumption and Stability of Economic Models, *Economic Theory* 5, 353-359.
Schwartz, J.T., 1961, *Lectures on the Mathematical Method in Analytical Economics*, New York, Gordon and Breach.
Semmler, W., 1987, A Macroeconomic Limit Cycle with Financial Perturbations, *Journal of Economic Behavior and Organization* 8, 469-495.
Sen, A., (ed.) 1970, *Growth Economics: Selected Readings,* Harmondsworth, Middlesex, Penguin Books.
Shell, K. and J.E. Stiglitz, 1967, The Allocation of Investment in a Dynamic Economy, *Quarterly Journal of Economics* 81, 592-609.
Shinkai, Y., 1960, On Equilibrium Growth of Capital and Labour, *International Economic Review* 1, 107-111.
Silverberg, G., 1984, Embodied Technical Progress in a Dynamic Economic Model: The Self-Organization Paradigm, in: R.M. Goodwin, M. Krüger and A. Vercelli (eds.), *Nonlinear Models of Fluctuating Growth,* 192-208.
Silverberg, G. and L. Soete (eds.), 1994, *The Economics of Growth and Technical Change: Technologies, Nations, Agents,* Aldershot, Edwar Elgar.
Sims, C., 1994, Review of Zarnowitz's book *Business Cycles*, in *Journal of Economic Literature* 32, 1885-1888.
Slutsky, E., 1953 (1915), On the Theory of the Budget of the Consumer (English translation of the 1915 Italian-language paper) in: K.E. Boulding and G.J. Stigler (eds.), 1953, *Readings in Price Theory,* 27-56.
Smithies, A., 1942, The Stability of Competitive Equilibrium, *Econometrica* 10, 258-274.
Solow, R.M., 1952, On the Structure of Linear Models, *Econometrica* 20, 29-46.
Solow, R.M., 1956, A Contribution to the Theory of Economic Growth, *Quarterly Journal of Economics* 70, 65-94. Reprinted in: Stiglitz, J.E. and H. Uzawa (eds.), 1969, *Readings in the Modern Theory of Economic Growth,* 58-87; Sen, A. (ed.), 1970, *Growth Economics: Selected Readings,* 161-192.
Solow, R.M., 1959, Competitive Valuation in a Dynamic Input-Output System, *Econometrica* 27, 30-53.
Solow, R.M., 1960, Investment and Technical Progress. In: K.J. Arrow, S. Karlin and P. Suppes (eds.), 1960, *Mathematical Methods in the Social Sciences, 1959,* 89-104; reprinted in: J.E. Stiglitz and H. Uzawa (eds.), 1969, *Readings in the Modern Theory of Economic Growth,* 156-171, and in R. Becker and E. Burmeister (eds.), 1991, *Growth Theory,* Vol. I, Chap. 1.

Solow, R.M., 1992, *Siena Lectures on Endogenous Growth Theory,* University of Siena, Department of Economics Monographs Series, No. 6. A revised and enlarged edition has been published in Italian as R.M. Solow, 1994, *Lezioni sulla teoria della crescita endogena* (with appendixes by Serena Sordi), Roma, La Nuova Italia Scientifica.
Sonnenschein, H., 1972, Market Excess Demand Functions, *Econometrica* 40, 549-563.
Sordi, S., 1988, *Theories of the Business Cycle,* PhD. Thesis, Florence, European University Institute, Department of Economics. A revised and enlarged edition has been published in Italian as S. Sordi, 1990, *Teorie del ciclo economico: uno studio comparato*, Bologna, CLUEB.
Sordi, S., 1993, *Nonlinear Models of the Business Cycle and Chaos Theory: A Critical Analysis*, University of Siena, Quaderni del Dipartimento di Economia Politica No. 153.
Sotomayor, J., 1973, Generic Bifurcations of Dynamical Systems, in M.M. Peixoto (ed.), *Dynamical Systems,* New York, Academic Press, 549-560.
Stiglitz, J.E., 1967, A Two-Sector Two Class Model of Economic Growth, *Review of Economic Studies* 34, 227-238.
Stiglitz, J.E., (ed.) 1966, *Collected Scientific Papers of Paul A. Samuelson,* Vol. I, II, Cambridge, MA, MIT Press.
Stiglitz, J.E. and H., Uzawa (eds.), 1969, *Readings in the Modern Theory of Economic Growth,* Cambridge, MIT Press.
Swan, T., 1956, Economic Growth and Capital Accumulation, *Economic Record* 32, 334-361.
Szidarovszky, F., 1990, Global Asymptotic Stability of Non-Linear Difference Equations with Non-Differentiable Functions, *Economic Letters* 33, 333-336.
Szidarovszky, F. and K. Okuguchi, 1988, A Note on Global Asymptotic Stability of Non-Linear Difference Equations, *Economic Letters* 26, 349-352.
Szpiro, G.G., 1994, Exchange Rate Speculation and Chaos-Inducing Intervention, *Journal of Economic Behavior and Organization* 24, 363-368.
Takayama, A., 1963, On a Two-Sector Model of Economic Growth: A Comparative Statics Analysis, *Review of Economic Studies* 30, 95-104.
Takayama, A., 1985, *Mathematical Economics,* 2nd edn., Cambridge, Cambridge University Press.
Tarr, D.G., 1976, Distributed Lags, Morishima Matrices, and the Stability of Economic Models, *Econometrica* 44, 597-600.
Tarr, D.G., 1978, Expectations and Stability with Gross Complements, *Review of Economic Studies* 45, 617-20.
Taussky, O., 1964, On the Variation of the Characteristic Roots of a Finite Matrix under Various Changes of its Elements. In: H. Schneider (ed.), 1964, *Recent Advances in Matrix Theory,* 124-138.
Thalberg, B., 1966, *A Trade Cycle Analysis,* Lund, Lund Economic Studies.
Theocharis, R.D., 1960, On the Stability of the Cournot Solution on the Oligopoly Problem, *Review of Economic Studies* 27, 133-134.
Thom, R. 1975 (1972), *Structural Stability and Morphogenesis: An Outline of a General Theory of Models,* Reading, Benjamin (English translation of the

1972 French edition).

Thygesen, N., K. Velupillai and S. Zambelli, 1991, *Business Cycles: Theories, Evidence and Analysis,* London, Macmillan.

Tinbergen, J., 1959 (1931), A Shipbuilding Cycle? In: *Selected Papers,* eds. Claassen, L.H., L.M., Koyck and H.J. Witteveen, Amsterdam, North-Holland (English translation of the paper originally published in German in 1931).

Torre, V., 1977, Existence of Limit Cycles and Control in Complete Keynesian System by Theory of Bifurcations, *Econometrica* 45, 1457-1466.

Tu, P.N.V., 1994, *Dynamical Systems: An Introduction with Applications in Economics and Biology,* 2nd edition, Berlin and New York, Springer-Verlag.

Turnbull. H.W., 1957, *Theory of Equations,* Edinburgh, Oliver and Boyd.

Uzawa, H., 1961, On a Two-Sector Model of Economic Growth, *Review of Economic Studies* 29, 40-47. Reprinted in Hahn, F.H. (ed.), 1971, *Readings in Theory of Growth,* 19-26.

Uzawa, H., 1963, On a Two-Sector Model of Economic Growth, II, *Review of Economic Studies* 30, 105-118. Reprinted in: Stiglitz, J.E. and H. Uzawa (eds.), 1969, *Readings in the Modern Theory of Economic Growth,* 411-424.

Varian, H.R., 1979, Catastrophe Theory and the Business Cycle, *Economic Inquiry* 17, 14-28.

Varian, H. R., 1992, *Microeconomic Analysis,* 3rd ed., New York, Norton.

Vercelli, A., 1984, Fluctuations and Growth: Keynes, Schumpeter, Marx, and the Structural Instability of Capitalism, in: R.M. Goodwin, M. Krüger and A. Vercelli (eds.), *Non-linear Models of Fluctuating Growth,* Berlin and New York, Springer-Verlag, 209-231.

Volterra, V., 1927, Variazioni e fluttuazioni del numero d'individui in specie animali conviventi. In: *Memorie del R. Comitato talassografico italiano,* memoria CXXXI. Reprinted in: Volterra, V., 1962, *Opere Matematiche: memorie e note,* a cura dell'Accademia Nazionale dei Lincei, Roma, Cremonese, Vol. V, 1-106.

Volterra, V., 1931, *Leçons sur la théorie mathématique de la lutte pour la vie,* Paris, Gauthier.

Volterra, V., 1959 (1929), *Theory of Functionals and of Integral and Integro-Differential Equations,* New York, Dover Publications reprint of the 1929 English edition, London, Blackie & Son.

von Neumann, J., 1945 (1937), A Model of General Equilibrium, *Review of Economic Studies* 13, 1-9 (English translation of the original 1937 paper in German). Reprinted in: R. Becker and E. Burmeister (eds.), 1991, *Growth Theory,* Vol. II, Chap. 19.

von Stackelberg, H., 1933, Sulla teoria del duopolio e del polipolio, *Rivista Italiana di Statistica, Economia e Finanza* 5, 275-279. Reprinted in N. Kloten and H. Möller (eds.), 1992, *Heinrich Freiherr von Stackelberg: Gesammelte wirtschaftswissenschaftliche Abhandlungen,* Regensburg, Transfer Verlag, Band 1, 169-183.

Walras, L., 1954 (1874, 1926), *Elements of Pure Economics (*English translation by W. Jaffé of the 1926 definitive French edition), London, Allen & Unwin.

Wan, H.Y., Jr., 1971, *Economic Growth,* New York, Harcourt Brace Jovanovich.

Waugh, F.V., 1964, Cobweb Models, *Journal of Farm Economics* 46, 732-50.

Weintraub, E.R., 1987, Stability Theory via Liapunov's Method: A Note on the Contribution of Takuma Yasui, *History of Political Economy* 19, 615-620.

Weintraub, E.R., 1991, Allais, Stability, and Liapunov Theory, *History of Political Economy* 23, 383-396.

Wen, K.-H., 1995, Continuous-Time Chaos in Stock-Market Dynamics, in W.A. Barnett, M. Salmon and A. Kirman (eds.), *Nonlinear Dynamics in Economics.*

Wen, K.-H., P. Chen and Z. Zhang, 1992, Bifurcation and Chaos in the Soft-Bouncing Oscillator, IC^2 Institute, The University of Texas at Austin, Working Paper #92-10-05.

Woodcock, A. and M. Davis, 1978, *Catastrophe Theory,* New York, E.P. Dutton.

Wu, J.-W. and D.P. Brown, 1989, Global Asymptotic Stability in Discrete Systems, *Journal of Mathematical Analysis and Applications* 140, 224-227.

Wymer, C.R., 1968, The Computation of the Eigensystem of an Econometric Model, unpublished manuscript.

Wymer, C.R., 1972, Econometric Estimation of Stochastic Differential Equation Systems, *Econometrica* 40, 565-577.

Wymer, C.R., 1987, Sensitivity Analysis of Economic Policy, in: G. Gandolfo and F. Marzano (eds.), *Essays in Memory of Vittorio Marrama, Vol. II,* Milano, Giuffré.

Wymer, C.R., 1996, Advances in the Estimation and Analysis of Non-Linear Differential Equation Models in Economics, forthcoming.

Zambelli, S., 1992, The Wooden Horse that wouldn't Rock: Reconsidering Frisch, in: K. Velupillai (ed.), *Nonlinearities, Disequilibria and Simulation* (Proceedings of the Arne Ryde Symposium on Quantitative Methods in the Stabilization of Macrodynamic Systems - Essays in Honour of B. Thalberg), London, Macmillan, 27-56.

Zarnowitz, V., 1992, *Business Cycles: Theory, History, Indicators, and Forecasting,* Chicago, University of Chicago Press.

Zeeman, E.C., 1974, On the Unstable Behavior of Stock Exchanges, *Journal of Mathematical Economics* 1, 39-44.

Zhang, W.-B., 1991, *Synergetic Economics: Time and Change in Nonlinear Economics,* Berlin and New York, Springer-Verlag.

Name Index

A

Allais, M., 413
Allen, R.G.D., 25
Anderson, P., 515, 529
Andronov, A.A., 338–339, 347, 350, 355, 438–439, 449, 453–454, 456, 469, 471
Aoki, M., 536
Arneodo, A., 522
Arrow, K.J., 183, 317, 413–415, 418, 420

B

Barnett, S., 89, 221, 252
Barnett, W.A., 515, 529, 531
Barro, R.J., 175, 187, 189, 286, 293, 379, 386–387, 393–394, 404
Basset, L., 310
Baumol, W.J., 1, 120, 365–366, 420–421, 503, 530
Bear, D.V.T., 132
Beckmann, M.J., 299, 466
Bellman, R., 361, 551–553, 555
Benhabib, J., 481, 483, 485, 503, 515, 529, 540
Bergstrom, A.R., 547
Bernoulli, James, 436
Bernoulli, John, 435
Bertram, J.E., 336, 410
Blatt, J.M., 340
Blinder, A.S., 95
Block, H.D., 413–415
Boldrin, M., 503
Boole, G., 25, 430
Bothwell, F.E., 341
Bródy, A., 458
Braun, M., 251, 362
Brock, W.A., 317, 412, 504, 528–529, 532
Brown, D.P., 412
Buiter, W.H., 402–403, 405
Burger, E., 555, 557
Burmeister, E., 175, 481–482, 557
Busenberg, S., 556
Bushaw, D.W., 413

C

Cagan, P., 49
Cartigny, P., 481, 485
Casti, J., 513
Chang, W.W., 441, 443, 445, 447, 449
Chen, P., 529, 531–532, 556–557
Chiarella, C., 529–530
Chipman, J.S., 89
Chukwu, E.N., 557
Clower, R.W., 413
Coddington, E.A., 361
Collar, A.R., 130, 250–251, 268, 271
Conlisk, J., 121, 132
Conti, R., 359, 361, 410, 438, 440
Cooke, K.L., 551–553, 555
Courant, R., 163–164, 275
Cournot, A., 137, 139, 421
Creedy, J., 515, 529

D

Davis, M., 536–537
Day, R.H., 1, 503, 515–516, 529, 532, 540
Debreu, G., 317
DeGrauwe, P., 515–517, 532
Diener, M., 536
Dobell, A.R., 175, 481–482
Dockner, E.J., 481
Dohtani, A., 420
Domar, E.D., 190–191, 293
Dore, M.H.I., 441
Dornbusch, R., 397

Duncan, W.J., 130, 271

E

Edgeworth, F.Y., 337

F

Farebrother, R.W., 90–91
Farkas, M., 458
Farley, A.M., 310
Feichtinger, G., 481, 500–501
Fiorito, R., 95
Fisher, F.M., 120, 337
Fisher, M.E., 255
Foley, D., 494–495, 497
Forsyth, A.R., 332, 430
Frankel, J.A., 516
Frazer, R.A., 130, 271
Frisch, R., 1, 556, 561–562
Froot, K.A., 516
Fujimoto, T., 132
Fuller, A.T., 251–252, 255
Furuno, Y., 557

G

Gabisch, G., 340, 441, 446, 458, 538
Gale, D., 306, 413, 540
Gandolfo, G., 143, 208, 210, 256–257, 291, 293, 340–341, 397, 481–482, 499, 522, 531, 547
Gantmacher, F.R., 120, 122, 221, 250, 263, 295
Georgescu-Roegen, N., 439–440, 445
Gillespie, R.P., 163
Goodwin, R.M., 79, 208, 341, 445, 458–459, 461–464, 521–522, 527–529, 549
Grandmont, J.-M., 503, 529–530
Granger, C.W.J., 532
Grasman, J., 441
Guckenheimer, J., 339–340, 362, 469, 471, 475, 477, 491–492, 505, 512, 518, 520–522

H

Hahn, F.H., 286, 317, 420, 481
Hahn, W., 410
Haken, H., 533, 535–536
Hale, J.K., 556
Hamburger, L., 445
Hansen, A.H., 71–72
Harrod, R.F., 48, 176, 191, 293, 528
Heiner, R.A., 530
Hicks, J.R., 1, 77–78, 98–101, 280–281, 286, 308, 312, 326, 332
Hillinger, C., 208–209
Hirsch, M.W., 331, 455
Hobson, E.W., 306
Holme, H., 561–562
Holmes, P., 339–340, 362, 469, 471, 475, 477, 491–492, 505, 512, 518, 520–522
Hommes, C.H., 529
Hsu, S.-B., 456
Huang, T.-W., 456
Hurwicz, L., 413–415, 418
Hyslop, J.M., 86

I

Ichimura, S., 441, 443, 445–446
Inada, K.-I., 177
Ince, E.L., 332, 430, 433–434
Indelli, P., 132
Intriligator, M.D., 374
Invernizzi, S., 477
Irie, K., 90

J

Jaffé, W., 280
Jarsulic, M., 556–557
Johnson, C.R., 254
Jones, L.E., 386
Jorgenson, D.W., 298

K

Kaldor, N., 37, 183, 337, 441–443, 445, 447, 449, 480, 494, 499–500, 526, 538
Kalecki, M., 449, 557–559, 562
Kalman, R.E., 336, 410
Kamien, M.I., 374, 377
Kelsey, D., 503, 530
Kemp, M.C., 317
Keynes, J.M., 403
Khalil, H.K., 254

Koblo, R., 536
Koopmans, T.C., 549
Kosobud, R.F., 446
Krasovskiĭ, N.N., 360, 410, 413

L

Lancaster, K., 310
LaSalle, J.P., 410
Laughton, M.A., 256
Le Corbeiller, P., 445
Lefschetz, S., 410
Leibniz, G.W., 436
Leontief, W., 293–294, 299, 556
Levinson, N., 361, 440
Leydesdorff, L., 529
Liapounoff, A., 407
Liapunov, A.M., 263, 315, 332–334, 337, 339–341, 344, 361, 366, 407, 413, 421, 430, 466
Lin, K.-P., 310
Liu, T., 532
Lorenz, E.N., 505, 518
Lorenz, H.-W., 340, 430, 441, 446, 458, 480, 503, 521–523, 526, 529–530, 538–539
Lotka, A.J., 449–450
Luenberger, D.G., 85

M

Maccini, L.J., 95
Machlup, F., 1
Malliaris, A.G., 528
Mandelbrot, B., 505
Mankiw, N.G., 286–287
Mantel, R.R., 317
Manuelli, R.E., 386
Marotto, F.R., 512
Marrama, V., 441
Marschak, J., 549
Marshall, A., 291, 293, 331, 337
Martelli, M., 556
Martin, V.L., 515, 529
Mas-Colell, A., 299
McFadden, D., 255
McKenzie, L., 120, 263, 317
McManus, M., 143
Medio, A., 477, 481, 485, 503, 507, 512–513, 522, 529–530, 536
Metzler, L.A., 95–96, 101, 234, 285, 291, 298
Miller, M., 405
Milne- Thomson, L.M., 25, 430
Minorsky, N., 339, 341, 350, 438–440, 449, 456, 469, 471
Mirowski, P., 504–505, 529, 532
Mirrlees, J.A., 183
Montrucchio, L., 485–486
Mundell, R.A., 291
Murata, Y., 91, 115, 120, 251–252
Muth, J.F., 75, 80, 394

N

Negishi, T., 280, 420
Nerlove, M., 40
Neumann, P.K., 293
Newhouse, S., 525
Newman, P.K., 254
Nikaidô, H., 306, 413
Nishimura, K., 481, 483, 485
Nusse, H.E., 529

O

O'Neil, W.D., 446
Okuguchi, K., 90, 140, 412, 421, 426
Olech, C., 354

P

Padoan, P.C., 531, 557, 568
Patinkin, D., 316–317
Perron, O., 361
Pesaran, M.H., 532
Peters, E.E., 529
Petit, M.L., 227–228, 300, 311, 403, 557, 568
Phelps, E., 183–184
Phillips, A.W., 228–231, 234
Phillips, P.C.B., 256
Poincaré, H., 361, 504–505, 529
Pol, E., 538
Poston, T., 536, 538
Potter, S., 532
Prigogine, I., 536
Puu, T., 526

Q
Quandt, R.E., 420–421, 530

R
Ramsey, F.P., 377–378, 380, 403
Ritschard, G., 310
Romer, P.M., 189
Rosser, J.B., Jr., 505, 529, 538
Ruelle, D., 520, 525
Ryder, H.E., 299, 466

S
Sala-i-Martin, X., 175, 187, 189, 286, 293, 379, 386–387, 393–394, 404
Samuelson, P.A., 1, 71, 73, 76–77, 95, 161, 207, 284, 305, 308, 310–312, 314, 319, 324, 331, 335, 337, 377
Sansone, G., 359, 361, 410, 438, 440
Sargent, T.J., 530
Sasakura, K., 486
Sato, R., 91, 184–185
Saunders, P.T., 538
Scarf, H., 418
Scheinkman, J.A., 412, 536
Schinasi, G.J., 486–487
Schoonbeek, L., 533
Schwartz, J.T., 117, 296
Schwartz, N.L., 374, 377
Shinkai, Y., 300
Silverberg, G., 175, 536
Simpson, A., 130, 250–251, 268, 271
Sims, C., 531
Slutsky, E., 308
Smale, S., 331, 455
Smith, D.J., 441
Smith, O.K., 440
Smithies, A., 91
Smyth, D.J., 443, 445, 447, 449
Soete, L., 175
Solow, R.M., 120, 143, 175–177, 183, 189, 293, 297, 378–379, 464
Sonnenschein, H., 317
Sordi, S., 441, 445, 529
Sotomayor, J., 471, 475
Stewart, I., 538
Storey, C., 252
Swan, T., 175–176, 293, 378–379
Szidarovszky, F., 140, 412
Szpiro, G.G., 518

T
Takens, F., 525
Taussky, O., 121
Theocharis, R.D., 138
Thom, R., 536, 538
Thygesen, N., 441, 458
Tinbergen, J., 311, 556
Torre, V., 500
Tu, P.N.V., 538
Turnbull, H.W., 62, 86, 220, 274, 478
Turnovsky, S.J., 557

U
Uzawa, H., 177

V
Van den Besselaar, P., 529
Varian, H.R., 137, 425, 538
Venditti, A., 481, 485
Vercelli, A., 340
Volterra, V., 449–450, 454
von Stackelberg, H., 425

W
Walras, L., 280, 285, 331, 414
Wan, H.Y., Jr., 175
Waugh, F.V., 37
Weintraub, E.R., 413
Wen, K.-H., 556–557
Wentzel, J.J., 441
Wittaker, E.T., 450
Woodcock, A., 536–537
Woodford, M., 503, 536
Wu, J.-W., 412
Wymer, C.R., 257, 531, 547

Y
Yasui, T., 413

Z
Zambelli, S., 556
Zarnowitz, V., 441

Zeeman, E.C., 536
Zhang, W.-B., 536
Zhang, Z., 556–557

Subject Index

A

Absolute value of a complex number, 56
Accelerator, 72, 77, 208, 234, 294
  distributed lag, 98
  flexible, 208, 464
  inventory, 96
  multi-sector, 294
  non-linear, 101, 464
  second-order, 208
Adaptive expectations, 40, 190, 530
Addition formulae for the sine and cosine, 24
Adiabatic approximation, 535
Adjustment speeds, 28, 30, 163, 254, 298, 533
  and D-stability, 255, 285
  and stabilization of matrices, 255, 285
  fast and slow variables, and synergetics, 533
  infinite, 165, 534
Adjustment time in the neoclassical aggregate growth model, 184
Alternation, 18, 366, 491
Arbitrary constants, 2, 9, 11, 56–57, 108, 114, 148–149, 195–197, 214, 218, 225, 242, 249–250, 347, 432, 450–451, 551
  and conditional stability, 118, 261, 386, 394
  and variation of parameters method, 161, 200, 215, 264, 436
  determination of, 13, 18, 27, 63, 76, 87, 111, 113, 125, 127, 129, 152, 156, 162, 203, 260, 265, 268, 270, 296, 299
Arrow diagrams, 290, 353, 355, 385, 393, 449
Assignment problem, 300
Attractor, 335
  strange, 520
Auctioneer, 280
Autonomous
  equations, 342, 363
  systems, 333, 338, 341, 346, 354, 361, 368, 408, 410–411
Auxiliary equation, 54, 156, 194, 213

B

Backward solution, 26, 62, 76, 86, 123–124
Balance of payments adjustment
  and capital flows, 301
  and full employment, 301
  and target-instrument pairing, 300
  critical sum of elasticities, 567
  multiplier, 46
  price-specie-flow mechanism, 563
Balanced growth, 175, 293, 295, 300, 404
  optimal, 384, 390
Basin of attraction, 335
Bernoulli equation, 179, 320, 436
β-convergence in neoclassical growth model, 187
Bifurcation, 350, 469
  and sensitivity analysis, 479
  and structural stability, 471
  in continuous time systems, 469ff
  in discrete time systems, 488ff
Boundary conditions, 152, 384, 388
Budget constraint, 312, 495
Burger's stability conditions for mixed difference-differential equations, 556, 563

Business cycle models
Goodwin, 464
Hicks, 77, 98
IS-LM, 486, 500
Kaldor, 441, 445, 447, 480, 492, 499
Kalecki, 557
Metzler (inventory cycles), 95, 234
policy-induced cycles, 231
Samuelson, 71

C

Canonical equations, 376
Capital flows and internal and external equilibrium, 301
Capital intensity condition, 300
Capital stock adjustment principle, 207, 464, 561
Capital stock, desired, 207, 464, 561
Cash-in-advance, 494
Casorati determinant, 11
Catastrophe theory, 536
and inflation, 537
Centre, 349
Centre manifold, 476
Chaos
and synergetics, 536
correlation dimension, 529
correlation integral, 532
definitions of, 503–505
high- and low-dimensional, 529
topological, 514
transient, 521
Chaos in continuous time, 518ff
coupled oscillator, 522
forced oscillator, 522
horseshoe map, 521
Lorenz equations, 518
Rössler attractor, 521
Ruelle-Takens scenario, 525
strange attractor, 520
transient, 521
Chaos in delay differential equations, 556
Chaos in discrete time, 505ff
in n-dimensional systems, 514
intermittency, 512
Li-Yorke theorem, 513
logistic map, 505
Marotto theorem, 514
period-three implies chaos, 514
Sarkovskii theorem, 512
Singer theorem, 514
snap-back repeller, 514
unimodal maps, 511
Chaos in economics
and econometric forecasting, 530
and irreversibility of time, 530
and rational expectations, 530
and stochastic dynamics, 529, 531
in consumption, 539
in cyclical growth, 527
in exchange rates, 516
in growth theory, 515
in international trade, 526
in overlapping generations, 539
significance and detection, 528
Characteristic curve, 346
Characteristic equation, 54, 83, 111–112, 126, 128, 130, 156, 194, 213, 241, 247–248, 267, 269–270, 273, 551
complex roots, 56, 84, 109, 196, 214, 244, 551
conditions for the roots to be less than unity in absolute value, 58–59, 88–91, 115–118, 120–121, 132
conditions for the roots to have negative real parts, 198, 220–223, 251–254, 262–264, 555
multiple roots, 55, 84, 105, 108, 195, 214, 239, 243, 551
relation between the equation coefficients and the minors of the matrix, 113, 248, 315
transcendental, 551
Characteristic exponent, 355, 438
Characteristic roots of a matrix, 113, 248, 551
Characteristic vector, 113, 248
associated with dominant root, 119
Cobweb theorem, 35, 190

and expectations, 38, 79, 190
Codimension one bifurcations
in continuous time, 471ff
in discrete time, 489ff
Col, 348
Comparative dynamics, 324
Comparative statics
and dynamics, 314, 309
and Hopf bifurcation, 318
and qualitative calculus, 310
and second-order conditions for an extremum, 312
basic theorems, 306
definition of, 305
inverse problem, 310
Complementary function, 12
Complex dynamics, 503
Complex numbers
Cartesian form, 56
De Moivre's theorem, 56
exponential form, 197
imaginary part, 56
modulus, 56
pure imaginary, 197
real part, 56
trigonometric (polar) form, 56
Conditional stability, 118, 259, 337
Conservative systems, 456
Contraction mapping, 412, 424
Control
feedback, 227
optimal, 374ff
Control trajectories, 374
Control variables, 374
Correspondence principle, 45, 314
and bifurcations, 318
criticism and qualifications, 316
Costate variables, 375
Coupled oscillator, 522
Coupling and de-coupling, 522
Cusp catastrophe, 538
Cyclical growth, 100, 458
chaotic, 527

D

D-stability, 254, 285, 323
De Moivre's theorem, 56
Decentralized policy making, 300
Delay differential equations, 545
Dependence, linear, 10, 149
Descartes' theorem, 54
Determinantal equation, 112, 248, 315
Diagonal
negative and dominant, 253–254, 285, 290, 415, 417
positive and dominant, 116–117
Diagonalization of a matrix
and solution of difference systems, 114ff
and solution of differential systems, 250ff
Difference equation, definition of, 8
Difference equations (non-linear), qualitative analysis
2x2 systems, 368
bifurcations, 488
chaos, 505
Hopf bifurcation, 491
limit cycles, 366
monotonic movements, 365
oscillations, 366
phase diagram, 363
via Liapunov's second method, 411
Difference equations, linear with constant coefficients
arbitrary constants, 9, 13
backward and forward form, 10
definition and general principles, 9
first order, 17ff
fundamental set of solutions, 11
general solution, 11–12
higher order, 83ff
homogeneous and non homogeneous, 10
order of, 8
particular solution, 12
second order, 53ff
simultaneous systems, 103ff
Differences, definition of, 7
Differential equation, definition of, 147
Differential equations (non-linear), qualitative analysis

2x2 systems, 346
bifurcations, 469
chaos, 518
conservative and dissipative systems, 456
Hopf bifurcation, 475
limit cycles, 355, 437
Olech's theorem, 354
phase diagram, 342
phase plane, 346
relaxation oscillations, 439
via Liapunov's second method, 410
Differential equations with time delay, 550
Differential equations, linear with constant coefficients
arbitrary constants, 148, 152
definition and general principles, 148
first order, 155ff
fundamental set of solutions, 150
general solution, 149, 151
higher order, 213ff
homogeneous and non homogeneous, 148
order of, 147
particular solution, 151
second order, 193ff
simultaneous systems, 237ff
Differential-difference equations, 545
Disinvestment, 101, 465
Dissipative systems, 456
Distance, 409
Distributed lags, 98
continuously, 163
exponentially, 163
Koyck scheme, 29
Divergence, 456
Domar growth model, 190
Dominant root, 55, 91, 119–120, 263, 295
Dual stability theorem, 298
Duopoly, 138
Dynamic optimization, 374
Dynamical economics, definitions of, 1
Dynamics
comparative, 324
definitions of, 1

E
Economic dynamics, definitions of, 1
Efficiency units, 182
Eigenvalues, 113, 248
Eigenvector, 113, 248
Equilibrium
and particular solution, 13, 152
deviations from, 13, 152, 360
displacement of, 305
moving, 13, 152
multiple, 337, 366
stability, 331
stationary, 13, 152
Equivalent linearization, 341
Erratic dynamics, 503
Exact differential equations, 432
Excess demand price, 170
Exchange rate dynamics
and overshooting, 400
chartists and fundamentalists, 516
under rational expectations, 397
Exchange rate speculation, 211
and chaos, 516
and official intervention, 211
Existence and uniqueness theorem, 430
Exotic dynamics, 503
Expanding fixed point, 514
Expectations
adaptive, 40, 190, 530
and capital stock adjustment principle, 208, 210
and cobweb theorem, 38, 79
and demand, 298
and foreign exchange speculation, 211, 516
and inventory cycles, 96, 234
conservative, 298
extrapolative, 79, 96, 190, 210, 298
in general equilibrium, 298
perfect foresight, 49
rational, 75, 80, 394
regressive, 79, 190, 210

static, 79, 298
Export-led models, 143
External and internal balance, 301
Extremum points
and comparative statics, 309, 311
and dynamic stability, 318

F
Fast and slow variables, 533
Feedback control, 227
Feigenbaum constant, 507
First integral, 456
Fisher-Fuller theorem on stabilization of matrices, 255, 285, 300
Flip bifurcation, 490
and period doubling, 491
subcritical and supercritical, 491
Focal point, 349
Focus, 349
Fold (or saddle-node) bifurcation, 472, 489
Fold bifurcation
in the foreign exchange market, 499
Forced oscillator, 522
Forcing functions, 408, 522
Foreign exchange speculation, 211
and chaos, 516
and official intervention, 211
Forward solution, 26, 62, 76, 86, 123–124
Forward-looking variables, 395
Fractal attractor, 505
Fractals, 505
Frequency of an oscillation, 58
Frobenius theorem on nonnegative indecomposable matrices, 120, 296
Fuller's modified Routh-Hurwitz conditions, 251
Functional equation, 2
Fundamental set of solutions, 11, 14, 87, 113, 125, 150, 153, 249, 266

G
General equilibrium stability
global, 413
local, 280
Gestation lag in investment, 557
Global
instability, 411
stability, 333, 335, 337, 355, 407, 410, 412–413
stability of an adjustment process, 337
univalence of mappings, 306
Golden rule, 183
modified, 383
Gross substitutability, 285, 414
Growth and cycles, 100, 458
chaotic, 527
Growth models
adjustment time, 184
and chaos, 515ff
and optimal oscillations, 481ff
convergence hypothesis, 187
Domar, 191
endogenous, 189
export-led, 143
Harrod, 48
multisector, 293ff, 481ff
neoclassical aggregate, 175–189
optimal endogenous, 386ff
optimal neoclassical, 378ff
two sector, 299
with human capital, 286ff

H
Habit formation in consumption, 51
Hamiltonian, 375
Harrod-Domar growth model, 191
Hartman-Grobman theorem, 362
Hayes' stability conditions for mixed difference-differential equations, 555, 566
Henon map, 512
Hessian, 312–313, 322
Heteroclinic orbit, 350
Hicksian matrix, 286
Historical variables in rational expectations, 395
Homoclinic orbit, 349, 521
transversal, 350

Homogeneous and non homogeneous equations, definition and properties
- difference, 10
- differential, 148
- mixed differential-difference, 550

Hopf bifurcation in continuous time, 475
- in advertising model, 500
- in IS-LM model, 488, 500
- in Kaldor's model, 480
- in Kaldor's model with depreciation, 499
- in optimal multisector growth, 481, 485
- secondary Hopf bifurcation, 479, 525
- subcritical, 477
- supercritical, 477
- theorem, 477

Hopf bifurcation in discrete time, 491
- and liquidity costs in the firm, 499
- in Kaldor's model, 492
- theorem, 492

Horseshoe map, 521
Human capital and growth, 286
Hyperbolic fixed point, 362, 470
Hystero-differential equations, 550

I

Implicit
- function differentiation, 351
- function theorem, 306, 413, 422
- solution of a differential equation, 342

Inada-Uzawa conditions, 177, 378, 386, 515
Income effect in consumer theory, 284, 314
Independence, linear, 11, 149
Index of Poincaré, 439
Inflation
- and expectations, 49
- markup, 567

Initial conditions, 14, 18, 63, 87, 111, 125, 152, 157, 163, 203, 219, 245, 249, 260, 265, 359, 373, 454, 463, 465
- and conditional stability, 119, 261, 397, 403
- in optimal control problems, 375
- sensitive dependence on, 504, 511, 520, 523, 530

Input-output dynamic model, 293
- and balanced growth, 295
- dual stability theorem, 298
- relative stability, 297

Integrability condition, 432
Integral curves, 347
Integrating factor, 433
Intermittency, 512
Internal and external balance, 301
Inventory accelerator, 96
Inventory cycles, 95, 234, 522
- in continuous time, 234

Inverse comparative statics problem, 310
Inverse optimal problem, 486
Irreversibility of time, 457, 530
IS-LM model, 325
Isokines, 350

J

Jacobian, 282, 306, 309, 311–312, 317–318, 326, 412–413
- and bifurcations, 470, 475–477, 479–480, 484–486, 488, 493, 497

Jump variables in rational expectations, 395
- and right-hand derivative, 395

K

Keynesian IS-LM model, 325
- and cycles via Hopf's bifurcation, 486
- open-economy extension, 300

Koyck distributed lags, 29

L

Landau scenario, 525
Latent roots, 113, 248
Latent vector, 113, 248
Leading principal minors, 115, 252

Li-Yorke
- theorem, 513
- chaos, 514

Liénard-Chipart stability conditions for differential equations, 221
Liapunov function, 410
Liapunov's second method, 263, 344, 366, 407ff, 466
- and notion of distance, 409
- and oligopoly, 421
- and rules of thumb in the firm, 420
- and stability of general equilibrium, 413
- contraction mapping, 412, 424
- for difference systems, 411
- for differential systems, 410

Liapunov-Perron-Poincaré theorem on local stability, 361
Lie derivative, 456
Limit cycles, 366, 437
- and Hopf bifurcation, 475

Linear
- dependence, 10, 149
- independence, 11, 149

Linearisation and local stability, 360
- Hartman-Grobman theorem, 362
- local stability theorem, 362
- Poincaré-Liapunov-Perron theorem, 361

Linearisation, equivalent, 341
Lipschitzian condition, 430
Liquidity costs in the firm, 494
Local stability theorem, 362
Logistic map, 505ff
- and analytical proof of chaos, 508
- and onset of chaos, 508
- and period doubling, 507
- and windows, 508
- flip bifurcation, 507
- transcritical bifurcation, 506

Lorenz equations, 518
Lotka-Volterra equations, 449–451, 453–455

M

Majorization of a matrix, 117
Manifold
- stable, 260, 337, 350
- unstable, 350

Marotto theorem, 514
Marshall-Lerner condition, 567
Marshallian vs Walrasian stability conditions, 170
Matrix
- and difference systems, 112, 126, 130
- and differential systems, 245, 247, 267, 271
- and sensitivity analysis, 255
- characteristic roots and vectors, 113, 248
- D-stable, 254
- diagonalization, 114, 123, 250
- dominant negative diagonal, 253
- dominant positive diagonal, 116
- dominant root, 119–120
- exponential, 249
- Hicksian, 286
- indecomposable, 116
- majorization, 117
- Metzlerian, 253
- negative definite, 252
- non-negative, 115
- positive, 116
- quasi-dominant negative diagonal, 253
- quasi-negative definite, 252
- stabilization, 255
- stable, 251
- symmetric, 252
- trace, 254

Maximum principle, 375
Mean time-lag
- in continuous time, 164–165
- in discrete time, 29–30, 32

Metzlerian matrix, 253, 285, 298
Mixed differential-difference equations, 545ff
- and continuous vs discrete time in economics, 549
- characteristic equation, 551
- general solution, 550

homogeneous and non-homogeneous, 550
order of, 545
particular solution, 554
stability conditions, 555
superposition principle, 550
Modal matrix, 114, 249
Modes, 533
Modulus of a complex number, 56
Monetary equilibrium, 301, 326
Monopoly, 323
Multi-sector growth, 293ff, 483ff
Multiplier
delay-i, 29
distributed lag, 29
dynamic, 43
foreign trade, 46
impact, 29
income, 43
Lagrange, 312, 376
Lagrange, dynamic, 375
with taxation, 47
Multiplier-accelerator models, 71, 77, 98, 464

N

Negative definiteness, 252, 262
Neimark bifurcation, 479
Neoclassical growth models
adjustment time, 184
aggregate, 175–189
convergence hypothesis, 187
depreciation, 181
golden rule, 183
multisector, 481
optimal, 378, 481
optimal oscillations, 485
technical progress, 181, 286
vintage, 183
with human capital, 286
Node, 348
Non-linear vs linear models in economics, 429–431
Non-tâtonnement, 280
Nondegenerate fixed point, 362
Norm
of matrix, 121
of vector, 409
Normal form
for difference systems, 103, 112, 412
for differential systems, 237, 245, 412
in Hopf bifurcation, 476
Null solution, 360
Numéraire, 419, 482

O

Objective functional, 374
Olech's stability conditions for differential systems, 354
Oligopoly, 137, 421
Operational calculus method, 25, 61, 85, 122
Operator
difference, 7
differential, 270
forward, 61, 123
lag, 25, 122
polynomial, 129, 270
shift, 123, 129
Optimal control, 374ff
and oscillations, 485
and saddle points, 377, 385, 393
canonical equations, 376
control trajectories, 374
control variables, 374
costate variables, 375
equations of motion, 375
Hamiltonian, 375
infinite horizon, 374
initial conditions, 375
maximum principle, 375
objective functional, 374
open loop, 374
state variables, 374
terminal surface, 375
terminal time, 375
transversality condition, 376
Optimal growth
and oscillations, 485
endogenous, 386ff
multisector, 483ff
neoclassical aggregate, 378ff

Ramsey, 403
transitional dynamics in endogenous model, 390ff
transitional dynamics in neoclassical model, 384ff
turnpike property, 404
with finite terminal time, 404
Optimizing behaviour and comparative statics, 311
Order
of a difference equation, 8
of a differential equation, 147
of a mixed differential-difference equation, 545
of a simultaneous difference system, 103
of a simultaneous differential system, 237
Ordinary point, 347
Oscillations
frequency and period, 58
from Hopf bifurcation, 475, 491
in linear systems, 18, 58, 84, 105, 198, 214, 239, 551
limit cycles, 366, 437
linear and non-linear, 465
nonperiodic, 523
optimal, 481
orbitally stable, 437
periodic and aperiodic, 58, 503, 523
Poincaré-Bendixson theory, 438
proper and improper, 18
quasi-periodic, 523
relaxation, 439
Oscillator
coupled, 522
forced, 522
hard-bouncing, 100
relaxation, 439
soft-bouncing, 557
Overlapping generations, 539
Overshooting, 396

P

Partial adjustment equations
and synergetics, 533
in continuous time, 163
in discrete time, 28
in inventory investment, 102
Particular solution, defined, 12, 151, 550
Pearl-Verhulst map, 505
Period analysis, 20
Period doubling, 507
Period of an oscillation, 58
Period-three implies chaos, 514
Periodic, nonperiodic, and quasi-periodic oscillations, 523
Perron's theorem on positive matrices, 120
Perturbed motion, 361
Phase diagrams
for difference equations, 363
for difference systems, 368
for differential equation systems, 346
for differential equations, 341
for linear approximating differential systems, 356
Phase path, 346–347
Phase plane, 346
Pitchfork bifurcation, 474, 490
subcritical and supercritical, 475
Plane of the states, 346
Poincaré
-Bendixson theorem on limit cycles, 438
-Liapunov-Perron theorem on local stability, 361
index, 438
Policy making
decentralized, 300
dynamic, 300, 528
Predetermined variables in rational expectations, 395
Prey-predator equations, 449
Price adjustment
and rules of thumb, 420
in general competitive equilibrium, 280, 413
in oligopoly, 423
in partial competitive equilibrium,

170
Production possibility frontier, 481

Q

Quadratic forms, 262
Quantity adjustment
by duopolist, 138
by monopolist, 323
by oligopolist, 138
in Marshallian demand and supply partial equilibrium, 170
Quantity theory of money, 564
Quasi-negative definiteness, 252
Quasi-stability, 337

R

Rössler attractor, 521
Rational expectations, 75, 80
and exchange rate overshooting, 400
and jump variables, 395, 403
and overshooting, 396
and predetermined variables, 395
and saddle points, 396, 402
definition of, 394
Real equilibrium, 326
Reduced equation, 13
Reduction
of a difference equation to a system, 111
of a difference system to a single equation, 104, 111
of a differential equation to a system, 246
of a differential system to a single equation, 238, 246
of a first-order difference system to normal form, 127
of a first-order differential system to normal form, 268
of a general difference system to normal form, 130, 412
of a general differential system to normal form, 271, 412
Regular point, 347.
Relations between the roots and the coefficients of an algebraic equation, 220, 257–258, 478
Relativity of stability conditions, 170, 175, 318
Relaxation oscillations, 439
and forced oscillator, 522
and Kaldor's cyclical model, 445
and transient chaos, 522
Repeller, 335
snap-back, 514
Reversibility of time, 457
Right-hand-side
continuity, 395
derivative, 396
Roots of unity, 275
Routes to chaos in continuous time
Shil'nikov scenario, 521
coupled oscillator, 522
forced oscillator, 522
Lorenz equations, 518
Rössler attractor, 521
Ruelle-Takens scenario, 525
Routes to chaos in discrete time
intermittency, 512
period doubling, 508
Routh-Hurwitz stability conditions for differential equations, 221
modified, 251
Rules of thumb in the firm, 420

S

Saddle loops, 349
Saddle point, 348, 373
and overshooting of exchange rate, 401
and rational expectations, 396
generalized, 402
in neoclassical optimal growth, 385
in optimal endogenous growth, 393
stable arm, 349
Saddle-node (or fold) bifurcation, 472, 489
Samuelson's stability conditions for difference equations, 88
simplified, 90
Sarkovskii theorem, 512
Sawtooth diagram, 20

Schur's stability conditions for difference equations, 89
Schwarzian derivative, 512
Scrambled set, 513
Semi-stability, 337
Sensitive Dependence on Initial Conditions, 504
Sensitivity analysis in dynamical systems, 256
and bifurcations, 257, 479
and Hopf bifurcation, 258, 487
and structural stability, 257, 340
Separated and separable variables, 434
Separatrix, 349
Shadow price, 377
Shil'nikov scenario, 521
σ-convergence in neoclassical growth model, 187
Sign-preserving function, 172, 283, 420
Simplified stability conditions for a third-order difference equation, 90
Singular points, 347ff
centre, 349
col, 348
focal point, 349
focus, 349
node, 348
saddle point, 348
separatrices, 349
source, 349
spiral sink, 349
star or stellar node, 348
vortex, 348
Sink, 336
Snap-back repeller, 514
Source, 336, 349
Speculation on foreign exchange, 211
and official intervention, 211
Spiral sink, 349
Stabilisation policies, 228
and chaos, 528
derivative, 229
derivative-integral, 234
integral, 229, 233
proportional, 229, 231
proportional-derivative, 232
simultaneous derivative, integral, proportional, 234
Stability
and rapidity of convergence, 338
asymptotic, 334
concept of, in economics, 169, 331
conditional, 118, 259, 337
global, 335
in the large, 335
in the sense of Liapunov, 333
local, 334
of an adjustment process, 337
of the first kind in the small, 335
of the first kind, perfect, 335
one sided stability-instability, 337, 343
orbital, 355
path dependent, 337
perfect and imperfect, 281
qualitative approach, 332
quantitative approach, 332
quasi-stability, 337
relative, 297
semi-stability, 337
stable arm, 260, 337, 349
stable manifold, 260, 337
static and dynamic, 169, 280
structural, 339, 413
uniform, 336
Stability conditions
and comparative statics, 316
for difference equations, 19, 58, 88–91, 365
for difference systems, 115–117
for differential equations, 157, 198, 220–223, 343
for differential systems, 251–255, 273, 353
for mixed differential-difference equations, 555
global, 354, 410–411
relativity of, 170, 175, 318
Stability dogma, 339
Stabilization of matrices, 255, 285, 300
Stable arm, 260, 337, 349

Stable manifold, 260, 337
Star or stellar node, 348
Statics
  and dynamics, 1
  comparative, 305
Stellar node or star, 348
Strange attractor, 520
Structural stability, 338, 413
  and bifurcations, 340, 471
  and definition of 'slight' changes, 341
  and sensitivity analysis, 340
  and stability dogma, 339
  in applied economics, 340
  in economic theory, 340
Substitution effect in consumer theory, 314
Superposition theorem, 11, 149, 550
Synergetics, 532ff
  and adiabatic approximation, 535
  and ajustment speeds, 534
  and fast and slow variables, 533
  and insurgence of chaos, 536
  and overall dynamic behaviour, 535
  and Prigogine's order through chaos, 536
  order variables, 533
  slaving principle, 533

T
Tâtonnement, 280, 315, 413
Taxation, 47, 77
Technical coefficients, 293, 299
Technical progress, 181, 189, 286, 378, 387
Tent map, 512
Terminal condition, 18
Terminal surface, 375
Terminal time, 375
Time
  discrete vs continuous, 20, 67, 153, 545–550
  irreversibility, 457, 530
  reversibility, 457
Time separable utility function, 380
Torus, 523
Trajectory, 291, 346, 350
Transcritical bifurcation, 473, 490
Transfers, net, 47
Transient chaos, 521
Transversality condition, 376
Turnpike property, 404
Two-point boundary value problem, 376
Two-sector growth model, 299
  and capital intensity condition, 300

U
Uncovered interest parity, 397
Undetermined coefficients method, 152, 158, 160, 199, 215, 245, 268, 270, 554
Univalence of mappings
  global, 306, 413
  local, 306, 413
Unperturbed motion, 360

V
Variable coefficients, 434
Variation of parameters method, 161, 200, 204, 215, 264, 435
Verhulst map, 505
Vortex, 348

W
Walras' law, 414
Walrasian vs Marshallian stability conditions, 170
Warranted rate of growth, 48
Windows, 508
Wronskian, 150, 153, 201–204, 217–219

Z
z-transform, 85

Answers to Exercises

Answers to all exercises are given here. In the case of purely numerical exercises only the final result is given, while in other cases to help the reader better to understand the technicalities, a full derivation is provided.

Chapter 3

3.1.a)

(i) $y_t = A(0.5)^t + 20; \quad A = -20.$
(ii) $y_t = A(-\frac{1}{2})^t + 15; \quad A = 3.$
(iii) $y_t = A(-1)^t + 45; \quad A = 5.$
(iv) $y_t = A\left(-\frac{3}{2}\right)^t + \frac{2}{5}; \quad A = -\frac{7}{5}.$

3.1.b)

(i) $\bar{y} = 20$, monotonic convergent movement.
(ii) $\bar{y} = 15$, oscillatory convergent movement.
(iii) $\bar{y} = 45$, oscillatory with constant amplitude movement.
(iv) $\bar{y} = \frac{2}{5}$, oscillatory divergent movement.

3.2.

(i) $y_t = A(\frac{1}{2})^t + 2 + 2t; \quad A = 1.$
(ii) $y_t = A(3)^t + 4^{t+1}; \quad A = -4.$
(iii) $y_t = A(-\frac{\sqrt{2}}{2})^t + \frac{1}{2\sqrt{2}}\left(\cos\frac{\pi}{4}t + \sin\frac{\pi}{4}t\right); \quad A = -\frac{1}{2\sqrt{2}}.$

3.3. Starting from Eq. (3.46) of the text, we can see that $\sum_i w_i = w_0 \sum_i k_i$, but $w_0 = b_0/b = b_0$ having assumed $b = 1$.

Performing all the passages we obtain

$$\sum_{i=0}^{\infty} i(1-k)k^i = (1-k)\left(0 + k + 2k^2 + 3k^3 + \ldots..\right).$$

To evaluate this expression, let S be the quantity

$$S = k + 2k^2 + 3k^3 + \ldots$$

then

$$kS = k^2 + 2k^3 + 3k^4 + \ldots$$

and

$$S - kS = k + k^2 + k^3 + ... = k\left(1 + k + k^2 + ...\right) = \frac{k}{(1-k)},$$

so that

$$(1-k)\,S = \frac{k}{(1-k)}$$

or

$$S = \frac{k}{(1-k)^2}.$$

Thus we obtain the *mean lag* for the Koyck distributed lag model

$$E\,(i) = \frac{k}{(1-k)}.$$

We know from the text that the Koyck distributed lag equation is equivalent to a partial adjustment equation with $\alpha = (1-k)$ and that the adjustment speed is inversely related to the mean time lag which is $1 - \frac{1}{\alpha} = \frac{\alpha - 1}{\alpha}$.

Chapter 4

4.2. Rewriting the model with the new assumption we have

$$\begin{aligned} S_t &= sY_t, \\ I_t &= k\,(Y_t - Y_{t-1})\,, \\ S_t &= I_t. \end{aligned}$$

Substitution from the two first equations into the third one gives

$$Y_t\,(s-k) + kY_{t-1} = 0.$$

The solution of this difference equation is

$$Y_t = Y_0\left(\frac{k}{k-s}\right)^t = Y_0\left(1 + \frac{s}{k-s}\right)^t.$$

There is growth if $s/(k-s) > 0$. This holds only for $k > s$. In the affirmative case, it is easy to see that $s/(k-s) > s/k$. The reason is that, when income grows, saving based on current income is greater than saving based on (the same fraction of) lagged income, hence (in equilibrium) investment is correspondingly greater and income grows faster.

4.3. If $|\alpha\gamma| < |1 + \alpha\gamma|$ we apply formula (3.37) which yields the following particular solution

$$p_e = \sum_{i=1}^{\infty}\left(\frac{\alpha\gamma}{1+\alpha\gamma}\right)^i m_{t-i}.$$

In the opposite case, i.e. $|\alpha\gamma| > |1 + \alpha\gamma|$, we apply formula (3.39) and obtain the particular solution

$$p_e = -\sum_{i=1}^{\infty}\left(\frac{1+\alpha\gamma}{\alpha\gamma}\right)^i m_{t+i}.$$

4.4. By solving the difference equation (4.37) we get

$$p_t = p_0\left(\frac{\alpha - 1}{\alpha}\right)^t + p_e.$$

Since $\alpha < 0$, it follows that $\left|\frac{\alpha-1}{\alpha}\right| > 1$. This makes the movement unstable. Solving forward we obtain the particular solution

$$p_e = -\sum_{i=1}^{\infty}\left(\frac{\alpha}{\alpha - 1}\right)^i m_{t+i}.$$

4.6.a) In this case investment is constant. Hence, after the appropriate substitutions, the income equation becomes

$$Y_t = a + bY_t + cY_{t-1} + I_0,$$

therefore

$$(1-b)\,Y_t - cY_{t-1} = a + I_0,$$

whose solution is

$$Y_t = Y_0\left(\frac{c}{1-b}\right)^t + \frac{a + I_0}{1-b-c}.$$

The model is stable if the marginal propensity to save in the current period, i.e. $1-b$, is greater than the marginal propensity to consume referred to the previous period.

4.6.b) In this case the solution of the model is

$$Y_t = Y_0\left(\frac{c}{1-b}\right)^t + \frac{a}{1-b-c} + \frac{I_0\,(1+g)^{t+1}}{(1-b)\,(1+g) - c}.$$

4.7. Since demand and supply are evaluated at time t, we have to take $\Delta p_t = p_{t+1} - p_t$. The solution of the model is

$$p_t = p_0\,[1 + \alpha\,(b - b_1)]^t - \frac{a - a_1}{b - b_1}.$$

The movement is convergent to its equilibrium point if $|1 + \alpha\,(b-b_1)| < 1$ or

$$-1 < 1 + \alpha\,(b - b_1) < 1,$$

whence

$$-2 < \alpha\,(b - b_1) < 0.$$

If $(b - b_1) < 0$ (which is always the case when the demand curve is downward sloping and the supply curve upward sloping, namely $b < 0, b_1 > 0$) this condition requires

$$\alpha < \frac{2}{b_1 - b},$$

i.e., the adjustment speed must be lower than the critical value $2/(b_1 - b)$. If, on the contrary, $(b - b_1) > 0$, the stability condition can never be satisfied and the equilibrium is unstable.

Chapter 5

5.a)

$$\begin{array}{lll}
(i) & y_t & = (1.55)^t (A_1 \cos \omega t + A_2 \sin \omega t) + 38.46; \\
 & & \text{where} \cos \omega = 0.26 \text{ and} \sin \omega = 0.96. \\
(ii) & y_t & = (0.77)^t (A_1 \cos \omega t + A_2 \sin \omega t) + 2.200; \\
 & & \text{where} \cos \omega = 0.71 \text{ and} \sin \omega = 0.71. \\
(iii) & y_t & = (0.5)^t (A_1 \cos \omega t + A_2 \sin \omega t) + 180; \\
 & & \text{where} \cos \omega = 0 \text{ and} \sin \omega = 1. \\
(iv) & y_t & = (2.37)^t (A_1 \cos \omega t + A_2 \sin \omega t); \\
 & & \text{where} \cos \omega = 0.97 \text{ and} \sin \omega = 0.24. \\
(v) & y_t & = (A_1 \cos \omega t + A_2 \sin \omega t) + 500; \\
 & & \text{where} \cos \omega = 0.9 \text{ and} \sin \omega = 0.43.
\end{array}$$

5.b)

$$\begin{array}{llll}
(i) & y_t & = A_1 (2)^t + A_2; & A_1 = 1,\ A_2 = 2. \\
(ii) & y_t & = A_1 (3)^t + A_2 t (3)^t; & A_1 = 3,\ A_2 = -\frac{7}{3}.
\end{array}$$

Chapter 6

6.1.a) After the introduction of taxation in Samuelson's model, we obtain the following difference equation for income

$$Y_t - b[(1+k)(1-\tau)] Y_{t-1} + bk(1-\tau) Y_{t-2} - G = 0.$$

The characteristic equation is

$$\lambda^2 - \lambda b[(1+k)(1-\tau)] + bk(1-\tau) = 0.$$

Applying Descartes' theorem on the sequence of signs we can exclude negative roots and hence improper oscillations. Moreover, we know that the roots will have modulus less than 1, if and only if, the stability conditions (5.19) hold. In this case they are:

$$\begin{aligned}
1 - b[(1+k)(1-\tau)] + bk(1-\tau) &= 1 - b(1-\tau) > 0, \\
1 - bk(1-\tau) &> 0, \\
1 + b[(1+k)(1-\tau)] + bk(1-\tau) &> 0.
\end{aligned}$$

The first inequality, given the assumptions made on the parameters, is certainly satisfied; the third is also satisfied being a sum of all positive quantities. The crucial inequality is then the second one, which can be rewritten as

$$b < \frac{1}{k(1-\tau)}.$$

We can see that the introduction of the taxation parameter makes the stability condition easier to meet. Taxation (as a proportion of income) makes the model more stable because it reduces disposable income and hence consumption (on whose changes induced investment is based).

6.1.b) The difference equation in this case is

$$Y_t (1-\tau) = b(1+k)(1-\tau) Y_{t-1} - bk(1-\tau) Y_{t-2} + G.$$

Therefore

$$Y_t = b(1+k) Y_{t-1} - bkY_{t-2} + G/(1-\tau),$$

whose homogeneous part—and hence the stability condition—is the same as in equation (6.6) in the text.

In the case in which there is a lag between tax receipts and expenditure, the stability conditions become:

$$\begin{aligned}
1 - [b(1+k)(1-\tau) + \tau] + bk(1-\tau) &= (1-b)(1-\tau) > 0, \\
1 - bk(1-\tau) &> 0, \\
1 + [b(1+k)(1-\tau) + \tau] + bk(1-\tau) &> 0.
\end{aligned}$$

The crucial condition is once again the second one.

6.2.a) The known function of time is an exponential function, so that as a particular solution we may try

$$Y_0 (1+g)^t,$$

where Y_0 is an undetermined constant. After the usual procedure we obtain

$$Y_0 = \frac{A_0 (1+g)^2}{(1+g)^2 - (b+k)(1+g) + k}.$$

Hence the particular solution is

$$\overline{Y} = \frac{A_0 (1+g)^2}{(1+g)^2 - (b+k)(1+g) + k} (1+g)^t.$$

For our particular solution to be economically meaningful, Y_0 must be positive, i.e., the denominator must be positive, the numerator being positive. From the hint we know that the denominator coincides with the characteristic equation when $\lambda = (1+g)$. Consider then the function $f(\lambda) = \lambda^2 - \lambda(b+k) + k$. This is a parabola that never intersects the horizontal axis if the equation $\lambda^2 - \lambda(b+k) + k = 0$ has no real roots. In such a case $f(\lambda) > 0$ because it entirely lies above the horizontal axis, hence $(1+g)^2 - (b+k)(1+g) + k > 0$.

6.2.b) From the previous exercise we know that the characteristic equation of the corresponding homogeneous equation coincides with the denominator of the particular solution when $\lambda = (1+g)$. If the roots are real, the parabola $f(\lambda)$ intersects the horizontal axis at $\lambda = \lambda_1$ and $\lambda = \lambda_2$, lying below it for $\lambda_1 < \lambda < \lambda_2$ and above it elsewhere. Hence $(1+g)^2 - (b+k)(1+g) + k > 0$ for those values of g such that $(1+g) < \lambda_1$ or $(1+g) > \lambda_2$.

6.2.c) In the characteristic equation $\lambda^2 - \lambda(b+k) + k = 0$, we note that the succession of the coefficient signs is $+-+$. This—by Descartes' theorem—excludes the presence of negative real roots (no 'improper' oscillation may occur).

6.2.d) The stability conditions are

$$\begin{aligned} 1-(b+k)+k &= 1-b > 0, \\ 1-k &> 0, \\ 1+(b+k)+k &> 0. \end{aligned}$$

The first inequality is satisfied since we assume a propensity to consume smaller than unity; the third is satisfied since the left-hand side is the sum of positive magnitudes. The crucial inequality is then the second one, and so the stability condition in this model is

$$k < 1.$$

In Samuelson's model the stability condition is, as we know, $bk < 1$. Now, since $b < 1$, the stability condition $k < 1$ is more stringent. Hicks' model is then intrinsically less stable than Samuelson's model. This is not surprising, since in the former, induced investment depends on the variations in total demand, whereas in the latter it depends on the variations in consumption demand, which are evidently smaller.

6.2.e-f) Let us consider the parabola $f(k) = k^2 - (4-2b)k + b^2$, and let k_1 and k_2 be the roots of the equation $k^2 - (4-2b)k + b^2 = 0$, which are

$$k_1, k_2 = \frac{1}{2}\{4 - 2b \pm [(4-2b)^2 - 4b^2]^{\frac{1}{2}}\}.$$

Simple manipulation yields

$$k_1, k_2 = 1 + s \pm 2\sqrt{s} = (1 \pm \sqrt{s})^2,$$

where $s = (1-b)$ is the (marginal and average) propensity to save. Then we shall have $f(k) < 0$ for all values of k in the interval between the roots, and $f(k) > 0$ outside that interval. Therefore we obtain the cases that are listed in Table 6.1. We can combine these results with the stability conditions of the previous exercise (2.d). In order to do that, note that, since $s < 1$, $\sqrt{s}$ is smaller than unity and so

$0 < 1 - \sqrt{s} < 1$, so that $(1-\sqrt{s})^2 < 1$. The following table shows the complete results.

If	Roots	Regions of Fig. 6.2	Behaviour of the deviations from the trend, as $t \to \infty$
$k < (1-\sqrt{s})^2$	R, Di, $\mu < 1$	A	Monotonic, convergent
$k = (1-\sqrt{s})^2$	R, E, $\mu < 1$		Monotonic, convergent
$(1+\sqrt{s})^2 < k < 1$	Co, $\mu < 1$	B	Oscillatory, damped
$k = 1$	Co, $\mu = 1$		Oscillatory, constant amplitude
$1 < k < (1+\sqrt{s})^2$	Co, $\mu > 1$	C	Oscillatory, divergent
$k = (1+\sqrt{s})^2$	R, E, $\mu > 1$		Monotonic, divergent
$k > (1+\sqrt{s})^2$	R, Di, $\mu > 1$	D	Monotonic, divergent

Key: R=real (and positive); Di=distinct; E=equal; Co=complex conjugate; μ=modulus.

6.2.g) Starting from the homogeneous difference equation

$$\frac{1}{Y_0(1+g)^t}[Y_t - (b+k)Y_{t-1} + kY_{t-2}] = 0,$$

we can rewrite it in the following way

$$\frac{Y_t}{Y_0(1+g)^t} - \frac{b+k}{(1+g)}\frac{Y_{t-1}}{Y_0(1+g)^{t-1}} + \frac{k}{(1+g)^2}\frac{Y_{t-2}}{Y_0(1+g)^{t-2}} = 0,$$

or

$$D_t - \frac{b+k}{(1+g)}D_{t-1} + \frac{k}{(1+g)^2}D_{t-2} = 0,$$

where $D_t \equiv Y_t/Y_0(1+g)^t$ represents *relative* deviations. Applying the stability conditions to the last equation, we find that the crucial condition is $k/(1+g)^2 < 1$. It follows that in the opposite case, i.e. when $k/(1+g)^2 > 1$, D_t will be divergent anyhow. Note that with empirically plausible values of k and g, the instability condition is certainly satisfied, and so Hicks' assumption (that in his model relative deviations are explosive) is not unwarranted.

6.3.a) After the introduction of foreign trade into the model we obtain the following second order difference equation

$$Y_t - (b+k-m)Y_{t-1} + kY_{t-2} = A_0(1+g)^t + X_0(1+g_x)^t.$$

The known function of time is a sum of exponential functions, so as a particular solution we may try

$$\overline{Y} = H_0(1+g)^t + B_0(1+g_x)^t.$$

After substitution we obtain

$$H_0 = \frac{A_0 (1+g)^2}{(1+g)^2 - (b+k-m)(1+g)+k},$$
$$B_0 = \frac{X_0 (1+g_x)^2}{(1+g_x)^2 - (b+k-m)(1+g_x)+k}.$$

For $Y_0 = H_0 + B_0$ to be positive, it is sufficient that both denominators are positive. These can be examined along the usual lines (see exercises 6.2.a,b).

To check whether there is balance-of-payments equilibrium along this solution we have to compare the rate of growth of imports and exports. We know that exports grow according to the constant rate g_x. Imports are proportional to lagged income, therefore their growth rate is g plus g_x. This will generally cause balance-of-payments disequilibria. Only in the particular case in which the closed economy is static, i.e. $g = 0$, imports and exports will have the same rate of growth, namely g_x.

6.3.b) The stability conditions are

$$\begin{aligned} 1-(b+k-m)+k &= 1-b+m>0, \\ 1-k &> 0, \\ 1+(b+k-m)+k &> 0. \end{aligned}$$

The first and third inequalities are always satisfied given the assumptions we made on b and m (respectively $b < 1$ and $m < 1$). The crucial inequality is then the second one, i.e. $k < 1$, which is the same as in the case of closed economy, because the introduction of imports and exports does not influence the accelerator coefficient.

6.4.a) After introducing the new assumptions we obtain the following difference equation

$$Y_t - (b+k-m_1 b - m_2 k) Y_{t-1} + (k - m_2 k) Y_{t-2} = (1-m_2) A_0 (1+g)^t + X_0 (1+g_x)^t$$

As in the previous case the known functions of time is a sum of exponential functions, hence we can try as particular solution

$$\overline{Y} = H_0 (1+g)^t + B_0 (1+g_x)^t.$$

After substitution in the equation we get

$$H_0 = \frac{A_0 (1-m_2)(1+g)^2}{(1+g)^2 - [b(1-m_1)+k(1-m_2)](1+g) + k(1-m_2)},$$
$$B_0 = \frac{X_0 (1+g_x)^2}{(1+g_x)^2 - [b(1-m_1)+k(1-m_2)](1+g_x) + k(1-m_2)}.$$

6.4.b) The stability conditions are

$$\begin{aligned} 1-[b(1-m_1)+k(1-m_2)](1+g)+k(1-m_2) &= 1-b(1-m_1)>0, \\ 1-k(1-m_2) &> 0, \\ 1+[b(1-m_1)+k(1-m_2)](1+g)+k(1-m_2) &> 0. \end{aligned}$$

As in the previous case the crucial stability condition is the second one, which is now $1-k(1-m_2)>0$. This is less stringent than $1-k>0$ as $(1-m_2)<1$.

6.5.a) Starting from the non homogeneous difference equation given in the text, we see that the known function is a constant, i.e. $a_1 - a$. We can then try as a particular solution

$$\overline{p} = c.$$

After substitution in the non homogenous difference equation we get

$$bc - b_1(1+\rho)c + b_1\rho c = a_1 - a.$$

Simple manipulation yields

$$c = p_e = \frac{a_1 - a}{b - b_1}.$$

We can now consider the homogenous part of the difference equation which is

$$bp_t - b_1(1+\rho)p_{t-1} + b_1\rho p_{t-2} = 0.$$

Dividing through by b we obtain

$$p_t - \frac{b_1(1+\rho)}{b} p_{t-1} + \frac{b_1\rho}{b} p_{t-2} = 0,$$

which has the associated characteristic equation

$$\lambda^2 - \frac{b_1(1+\rho)}{b}\lambda + \frac{b_1\rho}{b} = 0.$$

6.5.b) When $\rho > 0$, the discriminant

$$\Delta = \left[\frac{b_1(1+\rho)}{b}\right]^2 - 4\frac{b_1\rho}{b}$$

is always positive given that $b_1 > 0$ and $b < 0$. Moreover, since the succession of the coefficient signs is $++-$, by Descartes' theorem we know that the roots are real and distinct, specifically one positive and one negative.

If $\rho < 0$, the discriminant will be negative when

$$\frac{b_1^2(1+\rho)^2}{b^2} < 4\frac{b_1\rho}{b},$$

or, dividing through by $\frac{b_1}{-b} = \frac{b_1}{|b|} > 0$,

$$\frac{b_1}{|b|} < \frac{-4\rho}{(1+\rho)^2}.$$

6.5.c) The stability conditions are

$$\begin{aligned} 1-\tfrac{b_1(1+\rho)}{b}+\tfrac{b_1\rho}{b} &= 1-\tfrac{b_1}{b} && > 0,\\ 1-\tfrac{b_1\rho}{b} & && > 0,\\ 1+\tfrac{b_1(1+\rho)}{b}+\tfrac{b_1\rho}{b} &= 1+\tfrac{b_1(1+2\rho)}{b} && > 0. \end{aligned}$$

For $\rho > 0$ the first and second inequalities are always satisfied and so the crucial inequality is the third, which may be written as

$$\frac{1}{(1+2\rho)} > \frac{b_1}{-b},$$

i.e.

$$\frac{b_1}{-b} < \frac{1}{(1+2\rho)}.$$

Let us now compare the last inequality with the stability condition holding in the original cobweb theorem, which is, in the case of normal demand and supply functions,

$$\frac{b_1}{-b} < 1.$$

Since $1/(1+2\rho)$ is smaller than 1, the new stability condition is more restrictive. This last result confirms that extrapolative expectations are an element of instability.

6.5.d) For $\rho < 0$, the first inequality of the stability conditions is always satisfied, and so the remaining two are the relevant ones. The second may be written as

$$\frac{b_1}{-b} < \frac{1}{-\rho}.$$

As for the third, two sub-cases must be distinguished. If $\rho \leq -1/2$, then such inequality is always satisfied, and so the only relevant inequality is $b_1/(-b) < 1/(-\rho)$. If $\rho > -1/2$, then $1+2\rho > 0$, and the inequality under consideration may be written as

$$\frac{b_1}{-b} < \frac{1}{1+2\rho}.$$

Now, we must check which inequality is more stringent, which is equivalent to checking whether

$$\frac{1}{-\rho} \lesseqqgtr \frac{1}{1+2\rho}$$

is in the range $-1/2 < \rho < 0$. From the inequality $1/(-\rho) \lesseqqgtr 1/(1+2\rho)$ we have

$$1+2\rho \lesseqqgtr -\rho,$$

therefore

$$3\rho \lesseqqgtr -1,$$

and

$$\rho \lesseqqgtr -\frac{1}{3},$$

so

$$\frac{1}{-\rho} \lesseqqgtr \frac{1}{1+2\rho} \qquad \text{for} \qquad \rho \lesseqqgtr -\frac{1}{3}.$$

The results for $\rho < 0$ are shown in table (6.2). Let us now compare, as before, the crucial stability condition with that holding in the original cobweb theorem. The result is that, if $-1 < \rho < 0$, then the new stability condition is less restrictive than $b_1/(-b) < 1$; if $\rho = -1$, they are the same; if $\rho < -1$, then the new stability condition is more restrictive than the original one. In fact, for $-1/3 < \rho < 0$, the stability condition is $b_1/(-b) < 1/(1+2\rho)$, which is less restrictive than the original one, since $1/(1+2\rho) > 1$. For $-1 < \rho \leq -1/3$ the stability condition is $b_1/(-b) < 1/(-\rho)$ and, since $1/(-\rho) > 1$, it is less restrictive than the original one. For $\rho \leq -1$ the stability condition is again $b_1/(-b) < 1/(-\rho)$ and, if $\rho = -1$, then $1/-\rho < 1$ and so $b_1/(-b) < 1/(-\rho)$ is more restrictive than the original one. The economic meaning of these results is as follows. The fact that producers expect price to reverse its movement is an element of stability, provided that the expected inversion is not too great, since in the opposite case it would have the contrary effect, with the result that the stability conditions would have to be more stringent than in the original cobweb theorem (which is the case when $\rho < -1$).

Chapter 7

7.a)

$$\begin{aligned} (i)\quad y_t &= A_1(2)^t + A_2(1.45)^t + A_3(0.55)^t + 500;\\ &\quad \text{where } A_1 = 37.5,\ A_2 = -88.19,\ A_3 = -49.3.\\ (ii)\quad y_t &= (A_1 + A_2 t)(1.5)^t + A_3(0.6)^t + 1000;\\ &\quad \text{where } A_1 = -511,\ A_2 = 184,\ A_3 = 667. \end{aligned}$$

7.b)

$$\begin{aligned} (i)\quad y_t &= A_1(1.41)^t + A_2(-1.41)^t + A_3;\\ &\quad \text{where } A_1 = 4.13,\ A_2 = -0.13,\ A_3 = -4.\\ (ii)\quad y_t &= A_1 + A_2(-1)^t + A_3(2)^t + A_4(-2)^t;\\ &\quad \text{where } A_1 = \tfrac{11}{3},\ A_2 = -\tfrac{8}{3},\ A_3 = -\tfrac{5}{6},\ A_4 = -\tfrac{1}{6}.\\ (iii)\quad y_t &= (0.70)^t\left[(A_1 + A_2 t)\cos\omega t + (A_3 + A_4 t)\sin\omega t\right];\\ &\quad \text{where } \cos\omega = 0 \text{ and} \sin\omega = 1. \end{aligned}$$

Chapter 8

8.1. The proof is the following. Assume that Y_{t-1}, Y_{t-2} and Y_{t-3} have been equal to the corresponding values of the ceiling, i.e. $Y_{t-1} = B_0(1+g)^{t-1}$, $Y_{t-2} = B_0(1+g)^{t-2}$ and $Y_{t-3} = B_0(1+g)^{t-3}$. Substituting these values into Eq. (8.17) we have

$$\begin{aligned} & Y_t - B_0(1+g)^{t-1}(b_1+k_1) - B_0(1+g)^{t-2}(k_2+b_2-k_1) + k_2 B_0(1+g)^{t-3} \\ = \; & A_0(1+g)^t, \end{aligned}$$

which can be written

$$\begin{aligned} Y_t = \; & B_0(1+g)^{t-3}\left[(1+g)^2(b_1+k_1) - (1+g)(k_2+b_2-k_1) + k_2\right] \\ & + A_0(1+g)^t. \end{aligned}$$

The proposition to prove is that such a value of Y_t is smaller than the value of the ceiling at time t, i.e. we have to show that the inequality

$$\begin{aligned} B_0(1+g)^t \; > \; & B_0(1+g)^{t-3}\left[(1+g)^2(b_1+k_1) - (1+g)(k_2+b_2-k_1) + k_2\right] \\ & + A_0(1+g)^t \end{aligned}$$

is true. Simple manipulations show that the last inequality is actually satisfied if, and only if, $Y_0 < B_0$. In fact,

$$\begin{aligned} B_0\left\{(1+g)^t - (1+g)^{t-3}\left[(1+g)^2(b_1+k_1) - (1+g)(k_2+b_2-k_1) + k_2\right]\right\} \\ > A_0(1+g)^t, \end{aligned}$$

hence, on the assumption that $Y_0 > 0$,

$$B_0 > \frac{A_0(1+g)^3}{(1+g)^2(b_1+k_1) - (1+g)(k_2+b_2-k_1) + k_2} = Y_0,$$

so

$$B_0 > Y_0.$$

If, on the contrary, $Y_0 > B_0$ (i.e. the trend lies above the ceiling), then Y_t would be greater than B_t and so income would continue indefinitely along the ceiling.

8.2. Following Hicks' assumption, $a_t = a$, we obtain the following second order difference equation

$$Y_t - b_1 Y_{t-1} - b_2 Y_{t-2} = A_0(1+g)^t - a.$$

As a particular solution, try the function

$$H_1(1+g)^t + H_2,$$

where H_1 and H_2 are undetermined constants. Substituting into the difference equation we obtain

$$H_1(1+g)^t + H_2 - b_1 H_1(1+g)^{t-1} - b_1 H_2 - b_2 H_1(1+g)^{t-2} - b_2 H_2 = A_0(1+g)^t - a,$$

hence

$$H_1(1+g)^{t-2}\left[(1+g)^2 - b_1(1+g) - b_2\right] + H_2(1-b_1-b_2) = A_0(1+g)^t - a,$$

therefore

$$\left\{H_1\left[(1+g)^2 - b_1(1+g) - b_2\right] - A_0(1+g)^2\right\}(1+g)^{t-2} + \{a + H_2(1-b_1-b_2)\} = 0.$$

This equation holds for any value of t if and only if both expressions in curly brackets are zero, i.e.

$$H_1\left[(1+g)^2 - b_1(1+g) - b_2\right] - A_0(1+g)^2 = 0, \quad a + H_2(1-b_1-b_2) = 0.$$

Thus we get

$$H_1 = \frac{A_0(1+g)^2}{(1+g)^2 - b_1(1+g) - b_2}, \qquad H_2 = \frac{a}{1-b_1-b_2},$$

and therefore

$$\bar{Y}_t = \frac{A_0(1+g)^2}{(1+g)^2 - b_1(1+g) - b_2}(1+g)^t - \frac{a}{1-b_1-b_2}$$

is a particular solution of the basic difference equation of the model when $I_t' = -a$.

8.3. To show that in the descending phase the movement converges to the lower limit (or 'floor'), given by the particular solution, we have to solve the homogeneous equation, i.e.

$$Y_t - b_1 Y_{t-1} - b_2 Y_{t-2} = 0.$$

The corresponding characteristic equation is

$$\lambda^2 - b_1\lambda - b_2 = 0.$$

Since the succession of the coefficient signs is $+--$, both roots will be real, one positive and the other negative. Given that $b_1 + b_2 = b < 1$, the stability conditions (5.19) are certainly satisfied, hence the movement will be to the floor.

Putting this together with the results of the previous exercise, we note that the lower limit is seen to be equal to the output corresponding to the multiplier $1/[(1+g)^2 - b_1(1+g) - b_2]$ applied to autonomous investment, less the output corresponding to the multiplier $1/(1-b_1-b_2)$ applied to (the absolute value of) depreciation. This has an economic meaning. In the descending phase, total investment equals autonomous investment less induced disinvestment, the latter being at worst equal to the amount of depreciation; therefore income cannot fall below the minimum level obtained applying the multiplier to that difference.

We note from the particular solution that the floor increases at a proportional rate approximately equal to g. The 'approximately' is due to the fact that, owing to the presence of the term $-a/(1-b_1-b_2)$, the actual rate of growth is slightly different from g.

In the descending phase, income moves toward the floor. But since the floor is increasing, in the course of the approach to it income must sooner or later begin to increase. At this point the accelerator comes back into gear and brings about a positive induced investment. Income starts increasing explosively towards the ceiling and then the story repeats itself.

8.4. Considering the general form of Hicks' model in which investment is evenly distributed over n successive periods, i.e. $k_1 = k_2 = ... = k_n = (1/n)k$, the income equation becomes

$$Y_t - \left(b_1 + \frac{k}{n}\right) Y_{t-1} - b_2 Y_{t-2} - b_3 Y_{t-3} - ... - b_n Y_{t-n} + \frac{k}{n} Y_{t-n-1} = A_0 (1+g)^t.$$

The corresponding characteristic equation of the homogeneous part is

$$\lambda^m - \left(b_1 + \frac{k}{n}\right) \lambda^{m-1} - b_2 \lambda^{m-2} - b_n \lambda + \frac{k}{n} = 0,$$

where $m \equiv n+1$. We can apply the stability condition (7.22), which is a sufficient condition as the coefficients are arbitrary. Since $b_1 + b_2 + ... + b_n = b < 1$, we obtain

$$b + 2\frac{k}{n} < 1,$$

that can always be satisfied for values of n sufficiently large, namely for $n > 2k/(1-b)$.

8.5. Assuming in Metzler's model that producers want to maintain inventories at a fixed level Q_0, Eq. (8.2) of the text has to be replaced with $\hat{Q} = Q_0$. After all the substitutions in the basic equation of the model, i.e. Eq. (8.1), we obtain the following difference equation for income

$$Y_t - b(\rho+2) Y_{t-1} - b(1-2\rho) Y_{t-2} - b\rho Y_{t-3} = I_0.$$

We can now apply the stability conditions (7.19) on the coefficients of the equation

$$\begin{aligned}
1 - b(\rho+2) - b(1-2\rho) - b\rho &> 0, \\
3 - b(\rho+2) + b(1-2\rho) + 3b\rho &> 0, \\
1 + b(\rho+2) - b(1-2\rho) + b\rho &> 0, \\
8\left[-(b\rho)^2 + b(\rho+2)b\rho + b(1-2\rho) + 1\right] &> 0.
\end{aligned}$$

After simple manipulations the inequalities can be written, for $\rho > 0$, as

$$\begin{aligned}
1 - 3b &> 0, \\
1 + 4b\rho + b &> 0, \\
3 - b &> 0, \\
1 + b + 2b^2\rho - 2b\rho &> 0.
\end{aligned}$$

Given the assumption we made on the parameters, the last three inequalities are always satisfied. The crucial inequality then is the first one, $1 - 3b > 0$. Therefore the system will be stable if, and only if

$$b < \frac{1}{3}.$$

If $\rho = 0$, the stability conditions become

$$\begin{aligned}
1 - 3b &> 0, \\
1 + b &> 0, \\
3 - b &> 0, \\
1 + b &> 0.
\end{aligned}$$

Thus the crucial inequality is still the first one.

8.6. In Metzler's model, when $\rho = 1$, the stability condition becomes more restrictive. In fact, the first inequality in (8.9) is unaffected. The second one becomes

$$(1+k)(2+k)b^2 - 3(1+k)b + 1 > 0.$$

It is satisfied for $b < b_1$ and for $b > b_2$, where b_1, b_2 are the roots of the equation $(1+k)(2+k)b^2 - 3(1+k)b + 1 = 0$. It can be checked that such roots are real, distinct and both positive. The inequalities (8.9) are then satisfied in the following intervals

$$\begin{aligned}
b &< \tfrac{3}{2k+3}, \\
b &< b_1, \\
b &> b_2.
\end{aligned}$$

Consider now the parabola $f(b) = (1+k)(2+k)b^2 - 3(1+k)b + 1$. From what we have said above it follows that $f(b) > 0$ for $b < b_1$ and for $b > b_2$ and $f(b) < 0$ for $b_1 < b < b_2$. Substituting in $f(b)$ the value $b^* = 3/(2k+3)$, we see that $f(b^*)$ takes on a negative value, and so b^* lies between b_1 and b_2. The interval $b > b_2$ must then be discarded, and, since the inequality $b < 3/(2k+3)$ is absorbed by the more stringent inequality $b < b_1$, the latter is the crucial stability condition. Substituting $b = 1/(1+k)$ in $f(b)$ we see that $f(b) < 0$, and so $b_1 < 1/(1+k) < b_2$.. This proves that the inequality $b < b_1$ is more stringent then the inequality $b < 1/(1+k)$.

8.7. In the case in which inventories follow a partial adjustment equation, the income equation is

$$Y_t = U_t + \alpha\left(\hat{Q} - Q_{t-1}\right) + I_0.$$

After the substitutions, we obtain the following third-order difference equation

$$\begin{aligned}
&Y_t - b\{1+\rho+\alpha[k(1+\rho)+1]\} Y_{t-1} + b\{\rho+\alpha[k\rho+(1+k)(1+\rho)]\} Y_{t-2} \\
&\quad - b\alpha\rho(1+k) Y_{t-3} = I_0.
\end{aligned}$$

Applying the stability conditions we have

$$\begin{aligned}
&1 - b\{1+\rho+\alpha[k(1+\rho)+1]\} + b\{\rho+\alpha[k\rho+(1+k)(1+\rho)]\} \\
&\quad - b\alpha\rho(1+k) > 0, \\
&1 + b\{1+\rho+\alpha[k(1+\rho)+1]\} + b\{\rho+\alpha[k\rho+(1+k)(1+\rho)]\} \\
&\quad + b\alpha\rho(1+k) > 0, \\
&3 - b\{1+\rho+\alpha[k(1+\rho)+1]\} - b\{\rho+\alpha[k\rho+(1+k)(1+\rho)]\} \\
&\quad + 3b\alpha\rho(1+k) > 0, \\
&-b^2\alpha^2\rho^2(1+k)^2 + b^2\rho\alpha(1+k)\{1+\rho+\alpha[k(1+\rho)+1]\} \\
&\quad - b\{\rho+\alpha[k\rho+(1+k)(1+\rho)]\} + 1 > 0.
\end{aligned}$$

The first and second inequalities are always satisfied since the propensity to consume is smaller than unity and the expectation coefficient is non-negative. The relevant inequalities are then the third and the fourth, that can be written as

$$\begin{aligned} 3-b\{1+2\,[\rho+\alpha\,(1+k-\rho)]\} &> 0, \\ b^2\,\{\alpha\rho\,(1+k)\,[1+\rho+\alpha(1+k-\rho]\} & \\ -b\,\{\rho+\alpha\,[k\rho+(1+k)\,(1+\rho)]\}+1 &> 0. \end{aligned}$$

From these inequalities we can see that if $\rho=0$, the stability conditions become

$$\begin{aligned} b &< \tfrac{3}{1+2\alpha(1+k)}, \\ b &< \tfrac{1}{\alpha(1+k)}, \end{aligned}$$

that are both satisfied for sufficiently low values of α.

When $\rho>0$ it will also normally be possible to find sufficiently low values of α such that the stability conditions will be satisfied. In fact, if we let $\alpha\to 0$ the stability conditions tend to

$$\begin{aligned} 3-b(1+2\rho) &> 0, \\ -b\rho+1 &> 0, \end{aligned}$$

or

$$\begin{aligned} b &< \tfrac{3}{(1+2\rho)}, \\ b &< \tfrac{1}{\rho}. \end{aligned}$$

For example, for $\rho=1$ these both reduce to $b<1$, which is certainly satisfied.

Hence the introduction of a partial adjustment equation in inventories tends to make the model more stable, provided that the adjustment coefficient is sufficiently low.

Chapter 9

9.A) Since the coefficients are non-negative, we can apply the stability conditions I of Sect. 9.2.2, Eq. (9.65), which show that the system is stable.

9.B)

(*i*)

$$\begin{cases} y_{1t} = A_1\alpha_1^{(1)}\,(3)^t+A_2\alpha_1^{(2)}\,(2)^t-3 \\ y_{2t} = A_1\alpha_2^{(1)}\,(3)^t+A_2\alpha_2^{(2)}\,(2)^t \end{cases}$$

where $\alpha_1^{(1)}=3, \alpha_1^{(2)}=2, \alpha_2^{(1)}=1, \alpha_2^{(2)}=1$.

(*ii*)

$$\begin{cases} y_{1t} = A_1\alpha_1^{(1)}\,(1.4)^t+A_2\alpha_1^{(2)}\,(0.95)^t \\ y_{2t} = A_1\alpha_2^{(1)}\,(1.4)^t+A_2\alpha_2^{(2)}\,(0.95)^t \end{cases}$$

where $\alpha_1^{(1)}=1, \alpha_1^{(2)}=1, \alpha_2^{(1)}=0.2, \alpha_2^{(2)}=-1.1$.

(*iii*)

$$\begin{cases} y_{1t} = A_1\alpha_1^{(1)}\left(\frac{2}{3}\right)^t+A_2\alpha_1^{(2)}\left(-\frac{1}{\sqrt{2}}\right)^t+A_3\alpha_1^{(3)}\left(\frac{1}{\sqrt{2}}\right)^t+7 \\ y_{2t} = A_1\alpha_2^{(1)}\left(\frac{2}{3}\right)^t+A_2\alpha_2^{(2)}\left(-\frac{1}{\sqrt{2}}\right)^t+A_3\alpha_2^{(3)}\left(\frac{1}{\sqrt{2}}\right)^t+3 \\ y_{3t} = A_1\alpha_3^{(1)}\,(0.67)^t+A_2\alpha_3^{(2)}\left(-\frac{1}{\sqrt{2}}\right)^t+A_3\alpha_3^{(3)}\left(\frac{1}{\sqrt{2}}\right)^t+1 \end{cases}$$

where $\alpha_1^{(1)}=0.5, \alpha_1^{(2)}=11, \alpha_1^{(3)}=0.66, \alpha_2^{(1)}=0.5, \alpha_2^{(2)}=3.4, \alpha_2^{(3)}=0.59, \alpha_3^{(1)}=1, \alpha_3^{(2)}=1, \alpha_3^{(3)}=1$.

(*iv*)

$$\begin{cases} y_{1t} = (A_1+A_2t)\,(-1)^t-\frac{2a}{(a+1)^2}a^t \\ y_{2t} = \left(A_1'+A_2't\right)(-1)^t+\frac{(a-1)}{(a+1)^2}a^t \end{cases}$$

9.C) Performing the matrix multiplication we see that the first $(n-1)$ rows define the new variables, and the last row expresses the equation in terms of the new variables.

Chapter 10

10.1. The other root of Eq. (10.6) is 1/2 with multiplicity $(n-1)$. In fact, if we put λ=1/2 in the characteristic matrix, we obtain a matrix with all identical elements. It is well known in matrix algebra that the determinant of a matrix which has two lines (rows or columns) equal is null. In our case this will be true for all minors up to the last one of order 1. Hence the matrix under consideration has rank 1 and degeneracy $(n-1)$. This shows that the root 1/2 has multiplicity $(n-1)$.

We know that, in the case of a multiple root, this will appear in the general solution multiplied by a polynomial in t with degree equal to the multiplicity of the root less one, say $(A_i+A_{1i}t+A_{2i}t^2+...+A_{n-2,i}t^{n-2})\left(\frac{1}{2}\right)^t$. However, due to the extreme degeneracy of the matrix, the coefficients of all the non zero powers of t identically vanish, so that the general solution turns out to be

$$X_{it}=A_i\left(\frac{1}{2}\right)^t+B_i\left(\frac{1-n}{2}\right)^t+\frac{a-nc_i+\sum_{j\neq i}c_j}{(n+1)\,b}, \quad i=1,2,...,n,$$

where $\sum_{i=1}^n A_i=0, B_i=B$, with $(n-1)$ of the A_i and B arbitrary.

For further detail see McManus (1962).

10.2. Under the assumption that each country imports directly or indirectly from all other countries and remembering that $b_i+h_i-m_i>0$ because m_i is part of b_i+h_i, the matrix

$$\mathbf{A}=\begin{bmatrix} (b_1+h_1-m_1) & m_{12} & ... & m_{1n} \\ m_{21} & (b_2+h_2-m_2) & ... & m_{2n} \\ ... & ... & ... & ... \\ m_{n1} & m_{n2} & ... & (b_n+h_n-m_n) \end{bmatrix}$$

is a non-negative *indecomposable* matrix. In this case we can apply the stability condition V given in Chap. 9, Sect. 9.2.2. Thus we have

$$(b_i + h_i - m_i) + \sum_{\substack{j=1 \\ j \neq i}}^{n} m_{ji} \leq 1,$$

at least one strict inequality.

Since $m_i = \sum_{\substack{j=1 \\ j\neq i}}^{n} m_{ji}$, we have that

$$b_i + h_i < 1$$

is a set of sufficient stability conditions. Since $1 - b_i - h_i > 0$ is the stability condition for the closed-economy multiplier, we can conclude that the n-country model is stable if each country is stable in isolation.

10.3. After the introduction of the new assumptions on investment, the system representing the multiplier model becomes

$$\begin{cases} Y_{1t} = (b_1 + h_1 - m_1) Y_{1t-1} + m_2 (1 + h_3) Y_{2t-1} + (h_3 M_{02} - M_{01} + M_{02}) \\ Y_{2t} = (b_2 + h_2 - m_2) Y_{2t-1} + m_1 (1 + h_4) Y_{1t-1} + (h_4 M_{01} - M_{02} + M_{01}). \end{cases}$$

The matrix of the system is

$$\begin{bmatrix} (b_1 + h_1 - m_1) & m_2 (1 + h_3) \\ m_1 (1 + h_4) & (b_2 + h_2 - m_2) \end{bmatrix}.$$

Now, since $b + h > m$ (see Chap. 4, Sect. 4.2.), the coefficients of the matrix are all positive. Then we can apply the *in*stability condition IX (Chap. 9, Sect. 9.2.2) and obtain

$$\begin{aligned} (b_1 + h_1 - m_1) + m_2 (1 + h_3) &> 1, \\ (b_2 + h_2 - m_2) + m_1 (1 + h_4) &> 1, \end{aligned}$$

as sufficient conditions for the dominant root (which is real positive, and with a positive associated characteristic vector) to be greater than one, and hence for the model to show growth. After simple manipulations we obtain the relevant values of h_3 and h_4:

$$h_3 > \frac{1 - b_1 - h_1 + m_1 - m_2}{m_2},$$

$$h_4 > \frac{1 - b_2 - h_2 + m_2 - m_1}{m_1}.$$

Chapter 12

12.1

$$\begin{array}{lrcll} (i) & y(t) & = & Ae^{-t}; & A = 1. \\ (ii) & y(t) & = & Ae^{-3t}; & A = 1. \\ (iii) & y(t) & = & Ae^{6t}; & A = 1. \\ (iv) & y(t) & = & Ae^{t}; & A = 1. \\ (v) & y(t) & = & Ae^{-2t}; & A = 1. \end{array}$$

12.2.

$$\begin{array}{lrcll} (i) & y(t) & = & Ae^{-t} + 2t; & A = 1. \\ (ii) & y(t) & = & Ae^{-3t} + \frac{2}{3}; & A = -\frac{1}{3}. \\ (iii) & y(t) & = & Ae^{6t} - \frac{1}{2}\cos 2t - \frac{3}{2}\sin 2t; & A = \frac{3}{2}. \\ (iv) & y(t) & = & Ae^{t} + 1 + t; & A = 2. \\ (v) & y(t) & = & Ae^{-2t} + \frac{1}{3}e^{-2t}; & A = \frac{2}{3}. \end{array}$$

12.3. The primitive of the integrand $\alpha e^{-\alpha t}$ is $-e^{-\alpha t}$, so the definite integral in the interval $(0, \infty)$ is $[-e^{-\infty}] + 1 = 1$.

12.4. By using the hint, we know that the primitive is $\dfrac{(\alpha\tau + 1)}{-e^{\alpha\tau}\alpha^2}$, hence the definite integral in the interval $(0, \infty)$ is $\left[0 - \left(-\frac{1}{\alpha^2}\right)\right] = \frac{1}{\alpha^2}$.

12.5. Following the hint we get

$$\begin{aligned} \Delta y(t) = Y(t) - y(t) = & \int_{-\infty}^{t-\theta} \alpha e^{-\alpha(t-s)} x(s)\, ds + \int_{t-\theta}^{t} \alpha e^{-\alpha(t-s)} [x(s) + \Delta x(s)]\, ds \\ & - \int_{-\infty}^{t} \alpha e^{-\alpha(t-s)} x(s)\, ds \\ = & - \int_{t-\theta}^{t} \alpha e^{-\alpha(t-s)} x(s)\, ds + \int_{t-\theta}^{t} \alpha e^{-\alpha(t-s)} [x(s) + \Delta x(s)]\, ds \\ = & \int_{t-\theta}^{t} \alpha e^{-\alpha(t-s)} \Delta x(s)\, ds \\ = & \Delta x(t-\theta) \int_{t-\theta}^{t} \alpha e^{-\alpha(t-s)} ds. \end{aligned}$$

Chapter 13

13.1. Taking the adjustment mechanism into account, we obtain the following first order differential equation:

$$\tilde{p}' = c\left[(\tilde{a} - \tilde{a}_1) + (b - b_1)\tilde{p}\right],$$

where $\tilde{x} = \ln x$.

The general solution of the system is

$$\tilde{p} = Ae^{c(b-b_1)t} + \frac{(\tilde{a} - \tilde{a}_1)}{(b - b_1)}.$$

The system will be stable if, and only if, $(b - b_1) < 0$, being $c > 0$. This is certainly satisfied in the normal case $(b < 0, b_1 > 0)$.

13.2. After the substitution, we obtain the following differential equation

$$p' = \frac{a_1 - a}{b} + \frac{b_1 - b}{b}p,$$

whose general solution is

$$p(t) = Ae^{\frac{b_1 - b}{b}t} + \frac{a_1 - a}{b_1 - b}.$$

The equilibrium is stable if, and only if, $(b_1 - b)/b < 0$. If $b < 0$, the condition holds for $b_1 - b > 0$, which is the usual one.

13.3. After the introduction of expectations in the previous exercise, the differential equation becomes

$$p' + \frac{b - b_1}{b - b_1 c}p = \frac{a_1 - a}{b - b_1 c}.$$

The general solution is

$$p(t) = Ae^{-\frac{b - b_1}{b - b_1 c}t} + \frac{a_1 - a}{b - b_1}.$$

The equilibrium is stable if, and only if, $(b - b_1)/(b - b_1 c) > 0$. This condition is satisfied only if the numerator and the denominator have the same sign. Then, in the normal case, $b - b_1 < 0$, so that the crucial stability condition is $b - b_1 c < 0$. If $c > 0$ (i.e., extrapolative expectations), the condition is always verified. If $c < 0$, the condition holds only if $|b| > |b_1 c|$.

13.4. Substituting the definitional equations for supply and demand in the equilibrium condition and rearranging terms, we obtain

$$p(t) = \frac{a_1 - a}{b} + \frac{b_1}{b}\hat{p}(t).$$

Substituting for $p(t)$ in the expectation equation, we have

$$\begin{aligned} \hat{p}'(t) &= \beta\left[\frac{a_1 - a}{b} + \frac{b_1}{b}\hat{p}(t) - \hat{p}(t)\right] \\ &= \beta\frac{b_1 - b}{b}\hat{p}(t) + \beta\frac{a_1 - a}{b}. \end{aligned}$$

A particular solution is obtained setting $\hat{p}'(t) = 0$, which gives $(a - a_1)/(b_1 - b)$. This is the same as p_e, the static equilibrium price. The general solution is

$$\hat{p}(t) = Ae^{\beta\frac{b_1 - b}{b}t} + p_e.$$

The stability condition is

$$\frac{b_1 - b}{b} < 0,$$

which is the same as in exercise 13.2. Given the equation written at the beginning of this exercise, the behaviour through time of $p(t)$ will be the same as that of $\hat{p}(t)$.

13.5. In equilibrium, the increment in output is equal to the increment in aggregate demand

$$Y' = C' + I'.$$

Assuming the simple consumption function $C = (1 - s)Y$, we have

$$Y' = (1 - s)Y' + I',$$

which gives

$$Y' = \frac{1}{s}I'.$$

Assuming that in the initial period productive capacity is fully utilized, for this situation to be continued over time the increase in potential output must equal the increase in output, i.e. $Y' = P'$, where $P' = \sigma I$ is the increase in potential output brought about by investment. From $P' = \sigma I, Y' = \frac{1}{s}I'$ and $Y' = P'$, we have

$$I' = s\sigma I,$$

that expresses the relation that must be satisfied by investment for productive capacity to be fully utilized over time. The solution of the differential equation is

$$I(t) = I_0 e^{s\sigma t},$$

where I_o is the initial level of investment. Investment must grow at the constant rate $s\sigma$. It is easy to see that, if investment grows at that rate, output will grow at the same rate. Integrating with respect to time both sides of

$$Y' = \frac{1}{s}I',$$

we have

$$sY = I + A,$$

where A in an arbitrary constant, and given

$$I(t) = I_0 e^{s\sigma t},$$

we have

$$sY = I_0 e^{s\sigma t} + A,$$

i.e.

$$Y = \frac{1}{s}\, I_0 e^{s\sigma t} + \frac{A}{s}.$$

Assuming that we start from an equilibrium situation, we must have $sY_0 = I_0$, so that $Y_0 = (1/s)I_0$ and, consequently, A must be zero. Thus we have

$$Y(t) = Y_0 e^{s\sigma t},$$

and so Y also grows at the rate $s\sigma$. Therefore, income grows at a rate which is the same as the one which obtains in Harrod's model $(s\sigma = s/k)$.

13.6. To calculate the time required for 90% of the initial deviation of $k(t)$ from its equilibrium value to be eliminated it is possible to apply formula (13.54). We obtain that $t_{k,0.90} = 137.9$. As far as $r(t)$ is concerned, we first have to find the equilibrium value. With the given values of n, s and α, the basic differential equation becomes

$$r' = 0.10r^{1/3} - 0.025r$$

The substitution $z = r^{2/3}$ transforms the above equation into

$$z' + \frac{0.05}{3}z = \frac{0.2}{3},$$

whose solution is

$$z(t) = Ae^{-\frac{0.05}{3}t} + 4.$$

The inverse transformation $r = z^{3/2}$ gives:

$$r(t) = \sqrt{\left(Ae^{-\frac{0.05}{3}t} + 4\right)^3}.$$

As $t \to \infty$, the term $Ae^{-\frac{0.05}{3}t}$ tends to zero, so that $r(t)$ tends to $\sqrt{4^3} = 8$, which is its equilibrium value. Given $r_0 = 54.812$ the solution becomes

$$r(t) = \left[\left(54.812^{2/3} - 4\right)e^{-\frac{0.05}{3}t} + 4\right]^{3/2}.$$

The initial deviation is 46.812, so that we must find the value of t such that

$$\left[10.6e^{-\frac{0.05}{3}t} + 4\right]^{3/2} = 12.6812,$$

from which

$$10.6e^{-\frac{0.05}{3}t} \simeq 5.3467.$$

Using logarithms,

$$t_{r,0.90} = -\frac{3}{0.05}\log 0.5044057 \simeq 41.062.$$

To calculate $t_{y,0.90}$ we need y_0 and y_e. Knowing the values of r_0 and r_e and the relations between y and r (see Sect. 13.2.2), we obtain

$$\begin{aligned} y_0 &= (r_0)^{\frac{1}{3}} = 3.80, \\ y_e &= (r_e)^{\frac{1}{3}} = 2. \end{aligned}$$

At this point we have to find the value of t such that

$$y(t) - 2 = 0.10(3.80 - 2),$$

where $y(t) = \left[10.6e^{-\frac{0.05}{3}t} + 4\right]^{1/2}$. Hence, substituting for $y(t)$ and rearranging terms we obtain

$$10.6e^{-\frac{0.05}{3}t} \simeq 0.7524.$$

Using logarithms,

$$t_{y,0.90} = 158.72$$

13.7. Applying the same procedure as in the previous exercise, we obtain

$$\begin{aligned} t_{k,0.50} &= 26.25, \\ t_{k,0.75} &= 52.51, \\ t_{r,0.50} &= 11.06, \\ t_{r,0.75} &= 20.78, \\ t_{y,0.50} &= 20.63, \\ t_{y,0.75} &= 37.48. \end{aligned}$$

13.8.a) Given the hypothesis on the labour supply, Eq. (13.33) becomes

$$K = rL_0e^{nt}\left(\frac{w}{p}\right)^h.$$

Differentiating both members with respect to time, we obtain

$$K' = r'L_0e^{nt}\left(\frac{w}{p}\right)^h + nrL_0e^{nt}\left(\frac{w}{p}\right)^h + rL_0e^{nt}h\left(\frac{w}{p}\right)^{h-1}\frac{w'}{p},$$

where we have used the fact that, in our one-good real economy, p is a constant. After collecting terms and simple manipulations, the fundamental dynamic equation is now:

$$r' = sf(r,1) - r\left[n + h(w'/w)\right].$$

13.8.b) Adopting a Cobb-Douglas production function, the real wage rate is equal to

$$\frac{\partial Y}{\partial L} = K^\alpha(1-\alpha)L^{-\alpha} = \left(\frac{K}{L}\right)^\alpha(1-\alpha) = r^\alpha(1-\alpha) = \frac{w}{p}.$$

Differentiating $w/p = r^\alpha(1-\alpha)$ with respect to time, we obtain

$$\frac{w'}{p} = \alpha(1-\alpha)r^{\alpha-1}r',$$

and so

$$\frac{w'}{p}/\frac{w}{p} = \frac{w'}{w} = \alpha\frac{r'}{r}.$$

After a little manipulation, the fundamental dynamic equation can be written as

$$r' = [sf(r,1) - nr](1+\alpha h)^{-1},$$

which, as Solow (1956) points out, "gives some insight into how an elastic labour supply changes things. In the first place, an equilibrium state of balanced growth still exists, when the right-hand side becomes zero, and it is stable, approached from any initial condition. Moreover, the equilibrium capital-labour ratio is *unchanged*, since r' becomes zero exactly where it did before. This will not always happen, of course; it is a consequence of the special supply-of-labour schedule. Since r behaves much in the same way, so will all those quantities which depend only on r, such as the real wage rate".

13.9. Starting from the basic differential equation of the neoclassical growth model

$$r' = sf(r,1) - nr,$$

we can rewrite it as

$$\frac{K'L - KL'}{L^2} = sL_0e^{nt} f\left(\frac{K}{L_0e^{nt}}, 1\right) - \frac{L'K}{L^2},$$

which is the same as

$$\frac{K'}{L} = sL_0e^{nt} f\left(K, L_0e^{nt}\right),$$

and so

$$K' = sf(K, L_oe^{nt}).$$

Assuming a Cobb-Douglas production function the last equation becomes

$$K' = sK^{\alpha}\left(L_0e^{nt}\right)^{1-\alpha}.$$

This can be integrated directly and the solution is

$$K(t) = \left(K_0^{1-\alpha} - \frac{s}{n}L_0^{1-\alpha} + \frac{n}{s}L_0^{1-\alpha}e^{n(1-\alpha)t}\right)^{\frac{1}{1-\alpha}}.$$

As t becomes large, $K(t)$ grows essentially like $(s/n)^{\frac{1}{1-\alpha}} L_0e^{nt}$, namely at the same rate of growth as the labour force.

Chapter 14

14.1.

$$\begin{aligned}
(i)\quad y(t) &= A_1e^{0.62t} + A_2e^{-1.62t} + 10;\\
(ii)\quad y(t) &= A_1 + A_2e^{-t} - 5t;\\
(iii)\quad y(t) &= A_1e^{-1.76t} + A_2e^{-10.24t} + 3;\\
(iv)\quad y(t) &= e^{-t}(A_1\cos 7t + A_2\sin 7t) + 20 - 0.4\cos 4\pi t + 0.1\sin 4\pi t.
\end{aligned}$$

14.2.

$$\begin{aligned}
(i)\quad y(t) &= A_1e^{1.73t} + A_2e^{-1.73t};\ A_1 = -0.87,\ A_2 = 0.87.\\
(ii)\quad y(t) &= (A_1 + A_2t)e^{-t} + 6 - 4t + t^2;\ A_1 = -6,\ A_2 = -1.\\
(iii)\quad y(t) &= A_1e^{\frac{1}{2}t} + A_2e^{t} + e^{t}\left(9t - 2t^2 + \frac{1}{3}t^3\right);\ A_1 = 0,\ A_2 = 5.\\
(iv)\quad y(t) &= A_1e^{-t} + A_2e^{-\frac{1}{3}t} + \frac{1}{13}e^{-t}(2\cos t - 3\sin t);\ A_1 = -1,\ A_2 = \frac{24}{13}.
\end{aligned}$$

Chapter 15

15.1. After the introduction of the hypothesis on expectations in the standard multiplier equation, i.e. $Y = [1/(1-b)](K' + a)$, we obtain the following second order differential equation:

$$\left(1 - \alpha\beta k\gamma\frac{1}{1-b}\right)K'' + \left(\beta - \alpha\beta k\frac{1}{1-b}\right)K' + \alpha\beta K = \frac{\alpha\beta ka}{1-b}.$$

The corresponding homogeneous part is

$$K'' + K'\frac{\beta(1-b-\alpha k)}{1-b-\alpha\beta k\gamma} + \frac{\alpha\beta k(1-b)}{1-b-\alpha\beta k\gamma} = 0,$$

whose characteristic equation is

$$\lambda^2 + \frac{\beta(1-b-\alpha k)}{1-b-\alpha\beta k\gamma}\lambda + \frac{\alpha\beta k(1-b)}{1-b-\alpha\beta k\gamma} = 0.$$

The conditions for the roots to be negative if real, and have negative part if complex, are

$$\begin{aligned}
\frac{\beta(1-b-\alpha k)}{1-b-\alpha\beta k\gamma} &> 0,\\
\frac{\alpha\beta k(1-b)}{1-b-\alpha\beta k\gamma} &> 0.
\end{aligned}$$

The conditions are satisfied only if the numerator and the denominator in each fraction have the same sign. Hence, if

$$1 - b - \alpha\beta k\gamma > 0,$$

then

$$\beta(1-b-\alpha k) > 0 \text{ and } \alpha\beta k(1-b) > 0$$

must hold.

Otherwise, if

$$1 - b - \alpha\beta k\gamma < 0,$$

then

$$\beta(1-b-\alpha k) < 0 \text{ and } \alpha\beta k(1-b) < 0$$

have to be satisfied. This is however impossible, since $\alpha\beta k(1-b) < 0$ can never be satisfied given that $0 < b < 1$. Hence in the case in which $1 - b - \alpha\beta k\gamma < 0$ the model is not stable.

In the previous case, the crucial conditions are the first two, being the third one (i.e. $\alpha\beta k(1-b) > 0$) always satisfied.

From the crucial conditions, we obtain

$$\alpha < \frac{s}{k}\frac{1}{\beta\gamma} \text{ and } \alpha < \frac{s}{k}.$$

We can note that the second inequality coincides with Eq. (15.11). Now, the first inequality will be more or less or equally stringent than the second one depending on the value of $\beta\gamma$. We can now analyse the two different cases, namely $\gamma > 0$ (i.e., extrapolative expectations) and $\gamma < 0$ (i.e., regressive expectations).

Case I: $\gamma > 0$

If $\beta\gamma > 1$ the case of extrapolative expectations is more stringent than the case of static expectations (i.e. $\gamma = 0$).

If $\beta\gamma < 1$ the opposite is true, while for $\beta\gamma = 1$, extrapolative and static expectations are equivalent.

Case II: $\gamma < 0$

The condition $1-b-\alpha\beta k\gamma > 0$ is always satisfied, so that the crucial condition is $\alpha < \frac{s}{k}$, as in the presence of static expectations.

We can then conclude that depending on $\beta\gamma \lesseqqgtr 1$, extrapolative expectations are respectively less, equally or more stringent than regressive expectations.

15.2.a) Substituting the process of expectation-formation in the excess demand for foreign exchange and rearranging terms, we obtain the following second order differential equation

$$SR''(t) + \frac{b_1}{b_2}SR'(t) + \frac{a_1}{mb_2}SR(t) = \frac{-a_0 - B\cos\omega t}{mb_2}.$$

The characteristic equation of the homogeneous part is

$$\lambda^2 + \frac{b_1}{b_2}\lambda + \frac{a_1}{mb_2} = 0.$$

The stability conditions are

$$\frac{b_1}{b_2} > 0,$$
$$\frac{a_1}{mb_2} > 0.$$

Given that $m > 0, a_1 < 0$, the inequalities are satisfied only for $b_1 < 0$ and $b_2 < 0$.

15.2.b) To find the particular solution we try

$$\overline{SR}(t) = C_1\cos\omega t + C_2\sin\omega t + C_3,$$

where C_1, C_2, C_3 are undetermined coefficients. We have

$$\begin{aligned}\overline{SR}'(t) &= -\omega C_1\sin\omega t + \omega C_2\cos\omega t,\\ \overline{SR}''(t) &= -\omega^2 C_1\cos\omega t - \omega^2 C_2\sin\omega t.\end{aligned}$$

Substituting in the equation and collecting terms we obtain

$$\begin{aligned}&\left[\left(a_1 - mb_2\omega^2\right)C_1 + mb_1\omega C_2 + B\right]\cos\omega t\\ &+\left[-mb_1\omega C_1 + \left(a_1 - mb_2\omega^2\right)C_2\right]\sin\omega t + \left[a_1C_3 + a_0\right] = 0.\end{aligned}$$

This equation is identically satisfied if, and only if, the three expressions in square brackets are all zero. Solving the system we obtain

$$\begin{aligned}C_1 &= \frac{-B\left(a_1 - mb_2\omega^2\right)}{\left(a_1 - mb_2\omega^2\right)^2 + m^2b_1^2\omega^2},\\ C_2 &= \frac{-Bmb_1\omega}{\left(a_1 - mb_2\omega^2\right)^2 + m^2b_1^2\omega^2},\\ C_3 &= -\frac{a_0}{a_1}.\end{aligned}$$

To compare the particular solution with the element expressing the seasonal influences in $E_n(t)$, i.e. $B\cos\omega t$, we can write the particular solution in the following form. Let us make the substitutions

$$\begin{aligned}C_1 &= C\cos\gamma,\\ C_2 &= -C\sin\gamma,\end{aligned}$$

from which

$$C_1\cos\omega t + C_2\sin\omega t = C\cos(\omega t + \gamma),$$

where

$$C = +(C_1^2 + C_2^2)^{1/2} = \frac{B}{\sqrt{(a_1 - mb_2\omega^2)^2 + m^2b_1^2\omega^2}}.$$

Both express a constant-amplitude oscillation, having the same frequency $\omega/2\pi$, around the same constant value $(-a_0/a_1)$.

15.2.c) In the absence of speculative activity the amplitude of the basic oscillation is $B/-a_1$, which has to be compared with C as found in exercise (2.b). In the case in which $b_2 > 0$ (i.e., the model is unstable), then the denominator of C is clearly greater than $-a_1$, hence $C < B/-a_1$. If, on the contrary, stability obtains, then $b_2 < 0$, and nothing can be said on *a priori* grounds.

15.3.a) After the introduction of $E_G(t)$, the differential equation becomes

$$SR''(t) + \frac{mb_1 + f_2}{mb_2}SR'(t) + \frac{a_1 + f_1}{mb_2}SR(t) = \frac{1}{mb_2}\left[-f_1\frac{a_0}{a_1} - a_0 - B\cos\omega t\right].$$

The characteristic equation of the homogeneous part is

$$\lambda^2 + \frac{mb_1 + f_2}{mb_2}\lambda + \frac{a_1 + f_1}{mb_2} = 0.$$

Hence, the stability conditions are

$$\frac{mb_1 + f_2}{mb_2} > 0,$$
$$\frac{a_1 + f_1}{mb_2} > 0,$$

that can always be satisfied by a suitable choice of f_1, f_2. For example, if $b_1 > 0, b_2 > 0$ (the basic model is unstable, see exercise 2.a), then

$$f_1 > -a_1, \quad f_2 > -mb_1$$

will make the model stable.

15.3.b) As particular solution we try

$$\overline{SR}(t) = C_1 \cos \omega t + C_2 \sin \omega t + C_3.$$

Following the same procedure as in the previous exercise, we get

$$\begin{aligned} &\left[-\omega^2 m b_2 C_1 + \omega\left(m b_1 + f_2\right) C_2 + \left(a_1 + f_1\right) C_1 + B\right] \cos \omega t \\ &+\left[-\omega^2 m b_2 C_2 + \omega\left(m b_1 + f_2\right) C_1 + \left(a_1 + f_1\right) C_2\right] \sin \omega t \\ &+\tfrac{1}{m b_2}\left[C_3\left(a_1 + f_1\right) + a_0\left(1 + \tfrac{f_1}{a_1}\right)\right] = 0 \end{aligned}$$

Again, the equation is identically satisfied if, and only if, the three expressions in square brackets are all zero. Thus, we obtain a system of three equations in the three unknowns C_1, C_2, C_3. By solving it

$$\begin{aligned} C_1 &= \frac{B\left(a_1 + f_1 - b_2 m \omega^2\right)}{\left(f_2 + b_1 m\right)^2 \omega^2 + \left(a_1 + f_1 - b_2 m \omega^2\right)^2}, \\ C_2 &= \frac{B\left(f_2 + b_1 m\right) \omega}{\left(f_2 + b_1 m\right)^2 \omega^2 + \left(a_1 + f_1 - b_2 m \omega^2\right)^2}, \\ C_3 &= -\frac{a_0}{a_1}. \end{aligned}$$

We can write the particular solution in a simplified form. Let us make the substitutions

$$\begin{aligned} C_1 &= C \cos \gamma, \\ C_2 &= -C \sin \gamma, \end{aligned}$$

from which

$$C_1 \cos \omega t + C_2 \sin \omega t = C \cos(\omega t + \gamma)$$

where

$$C = +(C_1^2 + C_2^2)^{1/2} = \frac{B}{\sqrt{(f_2 + b_1 m)^2 \omega^2 + (a_1 + f_1 - b_2 m \omega^2)^2}}.$$

The policy parameters, f_1 and f_2, cause C to decrease, so the amplitude of the fluctuation is smaller than the one obtained in the previous part of the exercise. We can conclude that the presence of the monetary authorities in the foreign exchange market will have a stabilizing influence.

15.3.c) To find an explicit expression for the change in the authorities' foreign exchange reserves we have to evaluate the following definite integral

$$\begin{aligned} \int_0^T E_G(t) \mathrm{d}t &= \int_0^T \left\{ f_1 \left[SR(t) - \left(\frac{a_o}{-a_1}\right) \right] + f_2 SR'(t) \right\} \mathrm{d}t \\ &= f_1 \int_0^T \left[SR(t) - \left(\frac{a_o}{-a_1}\right) \right] \mathrm{d}t + f_2 \left[SR(t)\right]_0^T. \end{aligned}$$

Assuming that the policy parameters have been chosen so as to make the model stable, the homogeneous part will tend to zero, hence we can neglect it and consider only the particular solution. Substituting it in the previous expression we obtain

$$\begin{aligned} \int_0^T E_G(t) \mathrm{d}t &= f_1 \left[C_1 \frac{1}{\omega} \sin \omega t - C_2 \frac{1}{\omega} \cos \omega t \right]_0^T \\ &\quad + f_2 \left[C_1 \cos \omega t + C_2 \sin \omega t - a_0/a_1 \right]_0^T \\ &= f_1 \left[C_1 \frac{1}{\omega} (\sin \omega T - \sin 0) - C_2 \frac{1}{\omega} (\cos \omega T - \cos 0) \right] \\ &\quad + f_2 \left[C_1 (\cos \omega T - \cos 0) + C_2 \sin \omega T - \sin 0 \right)]. \end{aligned}$$

It is now easy to see that, if we take $T = s(2\pi/\omega), s = 1, 2, \ldots$, the last expression is zero. This is also true for any interval t_1, t_2 such that $t_2 - t_1 = T$. All this means that the change in international reserves is zero over a full cycle (or multiple thereof) of the exchange rate.

Chapter 16

16.1.

$$\begin{aligned} &(i) & y(t) &= A_1 e^{-2t} + e^t \left(A_2 \cos t + A_3 \sin t\right); \\ &(ii) & y(t) &= A_1 e^t + A_2 e^{3.8t} + A_3 e^{-0.8t}; \\ &(iii) & y(t) &= A_1 + A_2 t + A_3 e^t + A_4 e^{-3t}. \end{aligned}$$

16.2.

$$\begin{aligned} &(i) & \bar{y}(t) &= \tfrac{1}{8} + \tfrac{3}{4} t + \tfrac{1}{4} t^2; \\ &(ii) & \bar{y}(t) &= \tfrac{158}{9} - \tfrac{13}{3} t + \tfrac{19}{3} t^2 + \tfrac{1}{3} t^4; \\ &(iii) & \bar{y}(t) &= -\tfrac{1}{108} t^2 \left(3t^2 + 8t + 28\right) + \tfrac{3}{20} e^{2t} + -\tfrac{2}{5} \left(\cos t + 2 \sin t\right). \end{aligned}$$

Chapter 17

17.1. In this case we must substitute $G^* = -f_i \int_0^t Y \mathrm{d}t - f_p Y$ in Eq. (17.12), obtaining

$$Y'' + (\alpha l + \beta) Y' + \alpha\beta (l + f_p) Y + \alpha\beta f_i \int_0^t Y \mathrm{d}t = -\alpha\beta.$$

Differentiating

$$Y''' + (\alpha l + \beta) Y'' + \alpha\beta (l + f_p) Y' + \alpha\beta f_i Y = 0.$$

By applying the stability conditions, i.e. Eq. (16.29), we can see that the crucial stability condition is

$$(\alpha l + \beta) \alpha\beta (l + f_p) - \alpha\beta f_i > 0,$$

which is the same as

$$(\alpha l + \beta)(l + f_p) > f_i.$$

It can be seen that the greater f_i is, the more difficult it is to satisfy the inequality. But this *destabilizing* effect of integral policy is counterbalanced by a stabilizing effect of proportional policy. This allows us to say that the joint use of derivative and integral policies reduces the probability of instability, i.e. explosive oscillations.

17.2. In this case we must substitute $G^* = -f_i \int_0^t Y\mathrm{d}t - f_d Y'$ in Eq. (17.12), obtaining

$$Y'' + (\alpha l + \beta + \alpha\beta f_d)\, Y' + \alpha\beta l Y + \alpha\beta f_i \int_0^t Y\mathrm{d}t = -\alpha\beta.$$

Differentiating

$$Y''' + (\alpha l + \beta + \alpha\beta f_d)\, Y'' + \alpha\beta l Y' + \alpha\beta f_i Y = 0.$$

By applying the stability conditions stated in Eq. (16.29), we can see that the crucial stability condition is

$$(\alpha l + \beta + \alpha\beta f_d)\, \alpha\beta l - \alpha\beta f_i > 0,$$

which is the same as

$$(\alpha l + \beta + \alpha\beta f_d)\, l > f_i.$$

It can be seen that the greater is the value of f_i, the more difficult it is to satisfy the inequality. But this *destabilizing* effect of the integral policy is counterbalanced by a stabilizing effect of the derivative policy. This allows us to say that the joint use of derivative and integral policies reduces the probability of instability, i.e. explosive oscillations.

17.3. In this case we must substitute $G^* = -f_i \int_0^t Y\mathrm{d}t - f_d Y' - f_p Y$ in Eq. (17.12), obtaining

$$Y'' + (\alpha l + \beta + \alpha\beta f_d)\, Y' + \alpha\beta\,(l + f_p)\, Y + \alpha\beta f_i \int_0^t Y\mathrm{d}t = -\alpha\beta.$$

Differentiating

$$Y''' + (\alpha l + \beta + \alpha\beta f_d)\, Y'' + \alpha\beta\,(l + f_p)\, Y' + \alpha\beta f_i Y = 0.$$

By applying the stability conditions, i.e. Eq.(16.29), we can see that the crucial stability condition is

$$(\alpha l + \beta + \alpha\beta f_d)\, \alpha\beta\,(l + f_p) - \alpha\beta f_i > 0,$$

which is the same as

$$(\alpha l + \beta + \alpha\beta f_d)\,(l + f_p) > f_i.$$

It can be seen that the bigger is the value of f_i, the more difficult it is to satisfy the inequality. But this *destabilizing* effect of the integral policy is counterbalanced by a stabilizing effect of both the derivative and proportional policy. This allows us to say that the joint use of the three policies reduces the probability of instability, i.e. explosive oscillations.

17.4.a) Let us consider the fundamental equations of the model:

$$\begin{aligned} I' &= \eta\,(kY' - I)\,, \\ Y' &= -\alpha l Y + \alpha I + \alpha G - \alpha\mu, \\ G' &= -\beta G + \beta G^*. \end{aligned}$$

From the second equation we can find an expression for I. We then differentiate it, obtaining an expression for I'. We can now substitute these expressions in the first fundamental equation, obtaining

$$G = -\frac{1}{\alpha\eta} Y'' + \left(\frac{1}{\alpha} + k + \frac{l}{\eta}\right) Y' + lY + \frac{1}{\eta} G' + \mu.$$

Substituting for G' the expression given in the third fundamental equation and rearranging terms, we obtain

$$G\left(1 + \frac{\beta}{\eta}\right) = -\frac{1}{\alpha\eta} Y'' + \left(\frac{1}{\alpha} + k + \frac{l}{\eta}\right) Y' + lY + \frac{\beta}{\eta} G^* + \mu.$$

By differentiating it with respect to time we get an equation for G'. Substituting these values of G and G' in the third fundamental equation and rearranging terms we obtain the basic equation of the model:

$$\begin{aligned} &Y''' + (\alpha l + \eta - \alpha\eta k + \beta)\, Y'' + (\alpha\eta l + \alpha\beta l + \beta\eta - \alpha\beta\eta k)\, Y' \\ &+ \alpha\beta\eta l Y = \alpha\beta G^{*\prime} + \alpha\beta\eta G^* - \alpha\beta\eta\mu. \end{aligned}$$

If we now set $G^* = -f_p Y$ (i.e., proportional stabilisation policy), we get

$$\begin{aligned} &Y''' + (\alpha l + \eta - \alpha\eta k + \beta)\, Y'' + (\alpha\eta l + \alpha\beta l + \beta\eta - \alpha\beta\eta k + \alpha\beta f_p)\, Y' \\ &+ \alpha\beta\eta\,(l + f_p)\, Y = -\alpha\beta\eta\mu. \end{aligned}$$

The corresponding characteristic equation is

$$\lambda^3 + (\alpha l + \eta - \alpha\eta k + \beta)\,\lambda^2 + (\alpha\eta l + \alpha\beta l + \beta\eta - \alpha\beta\eta k + \alpha\beta f_p)\,\lambda + \alpha\beta\eta\,(l + f_p) = 0.$$

Applying stability conditions (16.29) we obtain

$$\begin{aligned} \alpha l + \eta - \alpha\eta k + \beta &> 0, \\ \alpha\eta l + \beta\,(\alpha l + \eta - \alpha\eta k) + \alpha\beta f_p &> 0, \\ \alpha\beta\eta\,(l + f_p) &> 0, \\ (\alpha l + \eta - \alpha\eta k + \beta)\,[\alpha\eta l + \beta\,(\alpha l + \eta - \alpha\eta k) + \alpha\beta f_p] - \alpha\beta\eta\,(l + f_p) &> 0. \end{aligned}$$

We can see that the only condition that is certainly satisfied is the third. Stability is not certain: in particular, from the first condition we get

$$\left(k - \tfrac{1}{\alpha}\right)\eta \;<\; l + \tfrac{\beta}{\alpha},$$

This is satisfied for sufficiently low values of k (i.e., $k < 1/\alpha$). In the opposite case, it will be satisfied for sufficiently low values of η. Since no policy parameter appears in this inequality, we cannot be sure that the model is stable. This shows that the introduction of the accelerator into Phillips' model makes the model itself intrinsically less stable.

If we set $G^* = -f_p Y - f_d Y'$ in the basic equation, we obtain the characteristic equation

$$\lambda^3 + (\alpha l + \eta - \alpha\eta k + \beta + \alpha\beta f_d)\,\lambda^2 + (\alpha\eta l + \alpha\beta l + \beta\eta - \alpha\beta\eta k + \alpha\beta\eta f_d + \alpha\beta\eta f_p)\,\lambda$$
$$+\alpha\beta\eta(l + f_p) = 0.$$

Applying stability conditions (16.29) we obtain

$$\begin{aligned}
\alpha l + \eta - \alpha\eta k + \beta + \alpha\beta f_d &> 0,\\
\alpha\eta l + \alpha\beta l + \beta\eta - \alpha\beta\eta k + \alpha\beta\eta f_d + \alpha\beta f_p &> 0,\\
\alpha\beta\eta(l + f_p) &> 0,\\
(\alpha l + \eta - \alpha\eta k + \beta + \alpha\beta f_d)(\alpha\eta l + \alpha\beta l + \beta\eta - \alpha\beta\eta k + \alpha\beta\eta f_d + \alpha\beta f_p) &\\
-\alpha\beta\eta(l + f_p) &> 0.
\end{aligned}$$

As in the case of pure proportional stabilization policy, the only condition certainly satisfied is the third one. In this case, however, stability is ensured because the other conditions can always be satisfied by choosing sufficiently high values of f_d. This confirms the stabilizing influence of the derivative policy.

17.4.b) We set $G^* = -f_i \int_0^t Y dt$ in the basic equation of the model (found in the previous exercise) and obtain

$$Y''' + (\alpha l + \eta - \alpha\eta k + \beta)\,Y'' + (\alpha\eta l + \alpha\beta l + \beta\eta - \alpha\beta\eta k)\,Y'$$
$$+\alpha\beta\eta l Y = -\alpha\beta f_i Y - \alpha\beta\eta f_i \int_0^t Y dt - \alpha\beta\eta\mu.$$

To get rid of the integral we differentiate with respect to time. Collecting terms we obtain

$$Y'''' + (\alpha l + \eta - \alpha\eta k + \beta)\,Y''' + (\alpha\eta l + \alpha\beta l + \beta\eta - \alpha\beta\eta k)\,Y''$$
$$+\alpha\beta(\eta l + f_i)Y' + \alpha\beta\eta f_i Y = 0,$$

whence the characteristic equation

$$\lambda^4 + (\alpha l + \eta - \alpha\eta k + \beta)\,\lambda^3 + (\alpha\eta l + \alpha\beta l + \beta\eta - \alpha\beta\eta k)\,\lambda^2$$
$$+\alpha\beta(\eta l + f_i)\lambda + \alpha\beta\eta f_i = 0.$$

As in the case of pure proportional policy, stability of the model is uncertain. In fact, necessary stability conditions are the positivity of all coefficients. The coefficient of λ^3 will be positive if, and only if,

$$\left(k - \tfrac{1}{\alpha}\right)\eta \;<\; l + \tfrac{\beta}{\alpha}.$$

This is satisfied for sufficiently low values of k; in the opposite case, it will be satisfied for sufficiently low values of η. Since no policy parameter appears in this inequality, we cannot be sure that the model is stable. The introduction of the accelerator into Phillips' model makes the model itself intrinsically less stable, and pure integral stabilisation policy does not contribute to the stability of the model.

17.5. Differentiating both members of the equation $Y' = \gamma\left[\hat{Q}(t) - Q(t)\right]$, we have

$$Y'' = \gamma\hat{Q}'(t) - \gamma Q'(t).$$

Since $\hat{Q}'(t) = k\hat{Y}'$ where $k > 0$ is the proportionality coefficient, and $Q'(t) = S(t) - I_0 e^{gt}$, we obtain

$$Y'' = \gamma k Y' + \gamma k a_1 Y'' + \gamma k a_2 Y''' - \gamma s Y + \gamma I_0 e^{gt},$$

from which

$$\gamma k a_2 Y''' + (\gamma k a_1 - 1)\,Y'' + \gamma k Y' - \gamma s Y = \gamma I_0 e^{gt}.$$

As a particular solution we try

$$\overline{Y} = c e^{gt}.$$

After substituting in the equation and solving for c we obtain

$$c = \frac{\gamma I_0}{(\gamma k a_2 g^3 + (\gamma k a_1 - 1)\,g^2 + \gamma k g - \gamma s)}.$$

The characteristic equation of the corresponding homogeneous part is

$$\gamma k a_2 \lambda^3 + (\gamma k a_1 - 1)\,\lambda^2 + \gamma k \lambda - \gamma s = 0.$$

It can immediately be seen that the stability conditions (16.29) are not satisfied since $-\gamma s < 0$, so that the model is unstable.

Chapter 18

18.1.

$$\begin{cases} y_1(t) &= e^{-\frac{t}{2}}\left(A_1\alpha_1^1 \cos\frac{\sqrt{7}}{2}t + A_2\alpha_1^2 \sin\frac{\sqrt{7}}{2}t\right),\\ y_2(t) &= e^{-\frac{t}{2}}\left(A_1\alpha_2^1 \cos\frac{\sqrt{7}}{2}t + A_2\alpha_2^2 \sin\frac{\sqrt{7}}{2}t\right).\end{cases}$$

The movement is damped oscillatory since the roots are complex conjugate with negative real part.

18.2.

$$\begin{cases} y_1(t) &= A_1\alpha_1^1 e^{-0.41t} + A_2\alpha_1^2 e^{2.41t},\\ y_2(t) &= A_1\alpha_2^1 e^{-0.41t} + A_2\alpha_2^2 e^{2.41t}.\end{cases}$$

18.3.

(*i*)

$$\begin{cases} y_1(t) &= e^{-t}\left(A_1\alpha_1^1 \cos t + A_2\alpha_1^2 \sin t\right),\\ y_2(t) &= e^{-t}\left(A_1\alpha_2^1 \cos t + A_2\alpha_2^2 \sin t\right),\end{cases}$$

where $\quad A_1 = 1;\ A_2 = 0;\ \alpha_1^1 = \frac{2-\sqrt{1}}{5};\ \alpha_1^2 = \frac{2+\sqrt{1}}{5};\ \alpha_2^1 = 1; \alpha_2^2 = 1.$

(ii)

$$\begin{cases} y_1(t) = A_1\alpha_1^1 e^{-t} + A_2\alpha_1^2 e^{4t}, \\ y_2(t) = A_1\alpha_2^1 e^{-t} + A_2\alpha_2^2 e^{4t}, \end{cases}$$

where $A_1 = 3$; $A_2 = 2$; $\alpha_1^1 = 2$; $\alpha_1^2 = -3$; $\alpha_2^1 = 1; \alpha_2^2 = 1$.

(iii)

$$\begin{cases} y_1(t) = (A_1\alpha_1^1 + A_2\alpha_1^2 t)\, e^t, \\ y_2(t) = (A_1\alpha_2^1 + A_2\alpha_2^2 t)\, e^t, \end{cases}$$

where $A_1 = 0$; $A_2 = 0$; $\alpha_1^1 = 2$; $\alpha_1^2 = 0$; $\alpha_2^1 = 1; \alpha_2^2 = 0$.

(iv)

$$\mathbf{y} = \begin{bmatrix} \cos t + 10\sin t \\ \cos t - 5\sin t \end{bmatrix} e^t + \begin{bmatrix} 2(\cos t + 3\sin t) \\ -4\sin t \end{bmatrix} t e^t.$$

(v)

$$\mathbf{y} = \begin{bmatrix} t - t^2 - \frac{1}{6}t^3 \\ \frac{1}{2}t^2 \\ t^2 + \frac{1}{6}t^3 \end{bmatrix} e^t.$$

18.4.

$$\begin{cases} y_1(t) = e^{-\frac{t}{2}}\left(A_1\alpha_1^1 \cos\frac{\sqrt{15}}{2}t + A_2\alpha_1^2 \sin\frac{\sqrt{15}}{2}t\right) - \frac{1}{6}e^t, \\ y_2(t) = e^{-\frac{t}{2}}\left(A_1\alpha_2^1 \cos\frac{\sqrt{15}}{2}t + A_2\alpha_2^2 \sin\frac{\sqrt{15}}{2}t\right) + \frac{2}{3}e^t. \end{cases}$$

18.5.

$$\begin{cases} y_1(t) = A_1\alpha_1^1 e^t + A_2\alpha_1^2 e^{-t} + (A_3\alpha_1^3 \cos t + A_4\alpha_1^4 \sin t) + \frac{2}{5}e^{2t}, \\ y_2(t) = A_1\alpha_2^1 e^t + A_2\alpha_2^2 e^{-t} + (A_1\alpha_2^3 \cos t + A_4\alpha_2^4 \sin t) - \frac{1}{15}e^{2t}, \end{cases}$$

where $A_1 = \frac{3}{4}$; $A_2 = \frac{7}{4}$; $A_3 = -\frac{19}{10}$; $A_4 = \frac{1}{5}$; $\alpha_1^1, \alpha_1^2, \alpha_1^3, \alpha_1^4 = 1; \alpha_2^1, \alpha_2^2 = -\frac{1}{3}$; $\alpha_2^3, \alpha_2^4 = -1$.

18.6. The characteristic equation of the system is

$$\lambda^2 - (a_{11} + a_{22})\,\lambda + (a_{11}a_{22} - a_{12}a_{21}) = 0.$$

Hence, the derivatives of the characteristic roots with respect to the parameter α are:

a) using the standard formula

$$\begin{aligned} \frac{\partial\lambda_1}{\partial\alpha} &= \frac{1}{2}\left(\frac{\partial a_{11}}{\partial\alpha} + \frac{\partial a_{22}}{\partial\alpha}\right) + \frac{1}{2\sqrt{(a_{11}+a_{22})^2 + 4(a_{12}a_{21} - a_{11}a_{22})}}\left[(a_{11}+a_{22})\left(\frac{\partial a_{11}}{\partial\alpha} + \frac{\partial a_{22}}{\partial\alpha}\right)\right. \\ &\quad \left. +2\left(a_{12}\frac{\partial a_{21}}{\partial\alpha} + a_{21}\frac{\partial a_{12}}{\partial\alpha} - a_{22}\frac{\partial a_{11}}{\partial\alpha} - a_{11}\frac{\partial a_{22}}{\partial\alpha}\right)\right], \\ \frac{\partial\lambda_2}{\partial\alpha} &= \frac{1}{2}\left(\frac{\partial a_{11}}{\partial\alpha} + \frac{\partial a_{22}}{\partial\alpha}\right) - \frac{1}{2\sqrt{(a_{11}+a_{22})^2 + 4(a_{12}a_{21} - a_{11}a_{22})}}\left[(a_{11}+a_{22})\left(\frac{\partial a_{11}}{\partial\alpha} + \frac{\partial a_{22}}{\partial\alpha}\right)\right. \\ &\quad \left. +2\left(a_{12}\frac{\partial a_{21}}{\partial\alpha} + a_{21}\frac{\partial a_{12}}{\partial\alpha} - a_{22}\frac{\partial a_{11}}{\partial\alpha} - a_{11}\frac{\partial a_{22}}{\partial\alpha}\right)\right]. \end{aligned}$$

b) using the following relations between the coefficients and the roots

$$\begin{aligned} \lambda_1 + \lambda_2 &= a_{11} + a_{22}, \\ \lambda_1\lambda_2 &= a_{11}a_{22} - a_{12}a_{21}, \end{aligned}$$

whence

$$\begin{aligned} \frac{\partial\lambda_1}{\partial\alpha} &= \left(\frac{\partial a_{11}}{\partial\alpha} + \frac{\partial a_{22}}{\partial\alpha}\right) - \frac{\partial\lambda_2}{\partial\alpha}, \\ \frac{\partial\lambda_2}{\partial\alpha} &= -\left(\frac{\partial a_{11}}{\partial\alpha} + \frac{\partial a_{22}}{\partial\alpha}\right)\frac{\lambda_2}{(\lambda_1 - \lambda_2)} - \left(a_{12}\frac{\partial a_{21}}{\partial\alpha} + a_{21}\frac{\partial a_{12}}{\partial\alpha} - a_{22}\frac{\partial a_{11}}{\partial\alpha} - a_{11}\frac{\partial a_{22}}{\partial\alpha}\right)\frac{1}{(\lambda_1 - \lambda_2)}. \end{aligned}$$

Chapter 19

19.1. After the introduction of price expectations, the linearised dynamic model is

$$\bar{p}_j' = k_j\left(\sum_{i=1}^{m} a_{ji}\,\bar{p}_j + b_j\,\bar{p}_j + \rho_j\,\bar{p}_j\right).$$

Let us consider the matrix of the system with static expectations

$$\mathbf{A} = \begin{bmatrix} k_1(a_{11}+b_1) & k_1 a_{12} & \dots & k_1 a_{1n} \\ k_2 a_{21} & k_2(a_{22}+b_2) & \dots & k_2 a_{2n} \\ \dots & \dots & \dots & \dots \\ k_m a_{m1} & k_m a_{m2} & \dots & k_m(a_{mm}+b_m) \end{bmatrix}$$

which is stable, since it fulfils the Metzler conditions by assumption. When $\rho \neq 0$, the system can be written as

$$\bar{\mathbf{p}}' = \mathbf{DA}\bar{\mathbf{p}},$$

where $\mathbf{D}$ is a diagonal matrix with elements

$$d_j = (1 - k_j b_j \rho_j)^{-1}.$$

Hence, by letting $d_j > 0$ or

$$1/k_j > b_j\rho_j,$$

we can apply the results on D-stability (see Sect. 18.2.2.1) and conclude that, if $\mathbf{A}$ is Metzlerian-stable, $\mathbf{DA}$ is also stable.

19.1.a-b-c) In the case of extrapolative expectations, i.e. $\rho_j > 0$ and a negative effect of $\hat{p}_j$ on E_j, i.e. $b_j < 0$, we have $b_j\rho_j < 0$. So, the inequality is certainly satisfied. It is also satisfied in the case of conservative expectations ($\rho_j < 0$) coupled with a positive effect of $\hat{p}_j$ on E_j ($b_j > 0$).

In all other cases we shall have $b_j\rho_j > 0$, and the inequality will be satisfied when the speed of adjustment is sufficiently low ($k_j < 1/b_j\rho_j$) or when the expectations coefficient is sufficiently low in absolute value ($|\rho_j| < 1/|b_j|\,k_j$).

19.2. The dynamic system is

$$\begin{aligned} p' &= cbp - cq + ca, \\ q' &= kp - \frac{k}{b_1}q + a\frac{k}{b_1}. \end{aligned}$$

The characteristic roots of the system are obtained by solving

$$\begin{vmatrix} cb-\lambda & -c \\ k & -\frac{k}{b_1}-\lambda \end{vmatrix} = \lambda^2 + \left(\frac{k}{b_1} - cb\right)\lambda + 4ck\left(1-\frac{b}{b_1}\right) = 0.$$

We can check the stability of the system without solving the characteristic equation. We know that the necessary and sufficient conditions for the roots to be negative if real, and have negative real parts if complex, are that the coefficient of the characteristic equation are positive. We can then distinguish four different cases, depending on the values of b_1 and b, that are listed in the following table:

	Parameter values		Movement
Case I	$b_1 > 0 \quad b < 0$		Stable
Case II	$b_1 > 0 \quad b > 0$		
a)		$\frac{k}{b_1} > cb \quad 1 > \frac{b}{b_1}$	Stable
b)	All other cases		Unstable
Case III	$b_1 < 0 \quad b < 0$		
a)		$\frac{k}{b_1} < cb \quad 1 > \frac{b}{b_1}$	Stable
b)	All other cases		Unstable
Case IV	$b_1 < 0 \quad b > 0$		Unstable

19.3. Considering Eq. (19.63), we can see that Eq. (19.62) is the same as Eq. (19.51), i.e. the characteristic equation of the physical system. Thus, since $\eta = r - \lambda$, it follows that the characteristic roots of the price system are obtained subtracting the characteristic roots of the physical system from r. This allows to state that the prices corresponding to the balanced growth path are constant if, and only if, $r = \lambda_1$.

Let $\eta_1 = r - \lambda_1$. Then the prices corresponding to the balanced growth path are those obtained when $p_i(0)/p_1(0) = \alpha_i^{(1)}$. It can be checked that when $r = \lambda_1$ these prices are the same as the static prices. The system for the static prices is homogeneous so that its determinant must be zero for it to give a non- trivial solution. This determinant is the same as the determinant appearing in the characteristic equation of the output system, with r instead of λ; thus if $r = \lambda_1$, then $p_i/p_1 = \alpha_i^{(1)}$. This proves that the dynamic price system (19.61) does not necessarily give rise to a changing price level on the balanced growth path. But if $r \neq \lambda_1$ (and this is surely the most probable case), then the prices corresponding to the balanced growth path will be changing over time, converging or diverging according as $r \lessgtr \lambda_1$.

19.4.a) The dynamics of the model can be analysed taking into account that the output of sector 1 is equal to net investment, so that $K' = X_1$. Differentiating the equation

$$K = a_1 N_1 + a_2 N_2$$

with respect to time, we obtain

$$K' = a_1 N_1' + a_2 N_2',$$

and, since, by definition, $b_1 X_1 = N_1$ (where b_1 is the fixed labour-input coefficient), we have that $K' = X_1 = (1/b_1)\,N_1$, so that the differential equation can be rewritten as

$$a_1 N_1' + a_2 N_2' = \frac{1}{b_1} N_1.$$

Differentiating $N = N_1 + N_2$ with respect to time, we have

$$N' = N_1' + N_2'.$$

Remembering that the growth rate of N is n, i.e. $N' = nN$, and $nN = n\,(N_1 + N_2)$, the equation can be rewritten as

$$N_1' + N_2' = nN_1 + nN_2.$$

Thus we obtain the system

$$\begin{aligned} a_1 N_1' + a_2 N_2' - \frac{1}{b_1} N_1 &= 0, \\ N_1' + N_2' - nN_1 - nN_2 &= 0. \end{aligned}$$

19.4.b) To find a solution of the system we can either put it in normal form (for this to be possible the determinant $|\mathbf{A}| = a_1 - a_2$ must not vanish), or apply the method explained in Sect. 18.3.1, which gives the characteristic equation

$$\begin{vmatrix} a_1\lambda - \frac{1}{b_1} & a_2\lambda \\ \lambda - n & \lambda - n \end{vmatrix} = (a_1 - a_2)\,\lambda^2 - \left[(a_1 - a_2)\,n + \frac{1}{b_1}\right]\lambda + \frac{n}{b_1} = 0.$$

Solving it we obtain the roots:

$$\lambda_1 = n; \quad \lambda_2 = \frac{1}{(a_1 - a_2)\,b_1}.$$

If $a_1 - a_2 = 0$, the characteristic equation has the only root $\lambda = n$. Both sectors grow at the same rate as the labour force, the same result as in the aggregate growth model (see Sect. 13.2). This is not surprising because, when $a_1 = a_2$, the model loses the characteristics of a two-sector model. In fact, if in our fixed-coefficient framework the capital/labour ratio is the same in both sectors, then the capital good and the consumption good can be considered from the production point of view as the same good but for a scale factor. Thus to have a true two-sector model, we must assume that $a_1 \neq a_2$.

19.4.c) Once found the roots, we can write the general solution of the system

$$\begin{aligned} N_1 &= A_{11} e^{\lambda_1 t} + A_{12} e^{\lambda_2 t}, \\ N_2 &= A_{21} e^{\lambda_1 t} + A_{22} e^{\lambda_2 t}. \end{aligned}$$

Using the initial conditions, we can now determine the values of the arbitrary constants $A_{11}, A_{12}, A_{21}, A_{22}$:

$$\begin{cases} A_{11} &= cN_0, \\ A_{12} &= \left(\frac{K_0}{c} - \frac{N_0}{b_1 n}\right) d, \\ A_{21} &= (1-c)\, N_0, \\ A_{22} &= \left(\frac{N_0}{b_1 n} - \frac{K_0}{c}\right) d = -A_{12}, \end{cases}$$

where

$$\begin{aligned} c &= \tfrac{a_2 b_1 n}{1-n(a_1-a_2)b_1}, \\ d &= \tfrac{c}{a}. \end{aligned}$$

19.4.d) From the solutions for the arbitrary constants it follow that, if K_0 is equal to $cN_0/b_1 n$, then A_{12} and A_{22} turn out to be zero. Thus in the solution only the exponential term containing n remains, and this means that N_1, N_2 (as well as X_1, X_2, K_1, K_2, given the assumption of fixed technical coefficients), grow at the same proportionate rate n (the growth rate of the labour force).

19.4.e) If $(a_1 - a_2) < 0$, then the second exponential in the solution tends to zero as t increases, so that the solution approaches the balanced growth path. On the other hand, if $(a_1 - a_2) > 0$, then the opposite conclusion holds. Since $(a_1 - a_2) \lessgtr 0$ means $a_1 \lessgtr a_2$, it follows that the balanced growth path is stable (unstable) if the consumption good sector is more (less) capital intensive than the capital good sector. This concerns stability in the absolute sense. For relative stability it is enough that $1/(a_1 - a)b_1 < n$, and this does not require $(a_1 - a_2) < 0$.

19.5. The stability of the model depends on the roots of the matrix kA. If the matrix A is stable on its own, by using the results on D-stability (see Sect. 18.2.2.1) it will be possible to choose economically meaningful (i.e., positive) policy parameters such that the matrix kA remains stable. More generally, it is possible to apply the Fisher and Fuller theorem; this describes the conditions that a matrix A must fulfil for a diagonal matrix k to exist capable of making the product matrix kA stable (see Sect.18.2.2.1). The diagonal matrix k may further be chosen so that the matrix kA has strictly negative real eigenvalues. It follows that, if A is unstable but satisfies the required conditions, it is possible to find a matrix k that not only stabilizes the system but also eliminates possible oscillations.

In our case, we want $k_{ii} > 0$, hence the matrix A must satisfy conditions (18.80).

19.6. The pairing fiscal policy-internal equilibrium and monetary policy-external equilibrium gives rise to the following system of differential equations

$$\begin{aligned} G' &= v_1(Y_F - Y), \\ r' &= v_2\left[m(Y) - x_0 - E_0 - K(r)\right], \\ Y' &= v_3\left[A(Y,r) + x_0 - m(Y) + G - Y\right]. \end{aligned}$$

The first two equations are the formal counterpart of the adjustment rules: government expenditure increases (decreases) if income falls short of (exceeds) the full employment level and the interest rate increases (decreases) if there is a deficit

(surplus) in the balance of payments. The third equation describes the usual process of adjustment of national income in response to excess demand; the constants $v_i > 0$ indicate the adjustment velocities. Expanding in Taylor's series at the point of equilibrium, we have

$$\mathbf{x}' = \mathbf{A}\mathbf{x},$$

where $\mathbf{x}$ is the column vector of the deviations $(\bar{G}, \bar{r}, \bar{Y})$ and

$$\mathbf{A} \equiv \begin{bmatrix} 0 & 0 & -v_1 \\ 0 & v_2 K_r & v_{2m_y} \\ v_3 & v_3 A_r & v_3(A_y - m_y - 1) \end{bmatrix}$$

is the matrix of the system of differential equations. The characteristic equation of system is

$$\begin{aligned} &\lambda^3 + \left[v_2 K_r + v_3(1 - A_y + m_y)\right]\lambda^2 \\ &+ \left[v_2 v_3 K_r(1 - A_y + m_y) - v_2 v_3 m_y A_r + v_1 v_3\right]\lambda + v_1 v_2 v_3 K_r = 0, \end{aligned}$$

and the necessary and sufficient conditions of stability are

$$\begin{aligned} v_2 K_r + v_3(1 - A_y + m_y) &> 0 \\ v_2 v_3 K_r(1 - A_y + m_y) - v_2 v_3 m_y A_r + v_1 v_3 &> 0 \\ v_1 v_2 v_3 K_r &> 0 \\ \left[v_2 K_r + v_3(1 - A_y + m_y)\right]\left[v_2 v_3 K_r(1 - A_y + m_y) - v_2 v_3 m_y A_r + v_1 v_3\right] & \\ -v_1 v_2 v_3 K_r & \\ = \left[v_2 K_r + v_3(1 - A_y + m_y)\right]\left[v_2 v_3 K_r(1 - A_y + m_y) - v_2 v_3 m_y A_r\right] & \\ +v_1 v_3^2(1 - A_y + m_y) &> 0. \end{aligned}$$

Given the signs of the various derivatives, it follows that the system will be stable if, and only if,

$$(1 - A_y + m_y) > 0.$$

Let us now consider the assignment of monetary policy to internal equilibrium and fiscal policy to external equilibrium, which gives rise to the following system of differential equations ($k_i > 0$ are the adjustment speeds):

$$\begin{aligned} G' &= k_1\left[x_0 - m(Y) + E_0 + K(r)\right], \\ r' &= k_2(Y - Y_F), \\ Y' &= k_3\left[A(Y,r) + x_0 - m(Y) + G - Y\right]. \end{aligned}$$

from which, by linearizing,

$$\mathbf{x}' = \mathbf{A}_1\mathbf{x},$$

where, as before, $\mathbf{x}$ indicates the vector of the deviations and

$$\mathbf{A}_1 \equiv \begin{bmatrix} 0 & k_1 K_r & -k_1 m_y \\ 0 & 0 & k_2 \\ k_3 & k_3 A_r & k_3(A_y - m_y - 1) \end{bmatrix}$$

is the matrix of the system of differential equations. If we expand the characteristic equation, we have

$$\lambda^3 + k_3(1 - A_y + m_y)\lambda^2 + (k_1 k_3 m_y + k_2 k_3 A_r)\lambda - k_1 k_2 k_3 K_r = 0.$$

One can see immediately that, as the constant term is negative, one of the stability conditions is violated, and thus the assignment in question gives rise to a movement that diverges from equilibrium.

Chapter 20

20.1 The dynamic behaviour assumptions made in Sect. 20.7 plus the assumption on the dynamics of money supply, give rise to the following system of differential equations:

$$\begin{aligned}
\frac{\mathrm{d}Y}{\mathrm{d}t} &= \varphi_1 \left[I(Y,R) - S(Y,R) - M(Y) + X_0 \right], \\
\frac{\mathrm{d}R}{\mathrm{d}t} &= \varphi_2 \left[L(Y,R) - L_s \right], \\
\frac{\mathrm{d}L_s}{\mathrm{d}t} &= \varphi_3 \left[X_0 - M(Y) + K(R) \right],
\end{aligned}$$

from which, by linearizing at the equilibrium point and denoting with an overbar the deviations from equilibrium, we have

$$\begin{aligned}
\frac{\mathrm{d}\bar{Y}}{\mathrm{d}t} &= c_1 \left[\frac{\partial I}{\partial Y} - \frac{\partial S}{\partial Y} - \frac{\partial M}{\partial Y} \right] \bar{Y} + c_1 \left[\frac{\partial I}{\partial R} - \frac{\partial S}{\partial R} \right] \bar{R}, \\
\frac{\mathrm{d}\bar{R}}{\mathrm{d}t} &= c_2 \frac{\partial L}{\partial Y} \bar{Y} + c_2 \frac{\partial L}{\partial R} \bar{R} - c_2 \bar{L}_s, \\
\frac{\mathrm{d}\bar{L}_s}{\mathrm{d}t} &= -c_3 \frac{\partial M}{\partial Y} \bar{Y} + c_3 \frac{\partial K}{\partial R} \bar{R}.
\end{aligned}$$

Stability depends on the roots of the characteristic equation

$$\begin{aligned}
&\begin{vmatrix}
c_1 \left[\frac{\partial I}{\partial Y} - \frac{\partial S}{\partial Y} - \frac{\partial M}{\partial Y} \right] - \lambda & c_1 \left[\frac{\partial I}{\partial R} - \frac{\partial S}{\partial R} \right] & 0 \\
c_2 \frac{\partial L}{\partial Y} & c_2 \frac{\partial L}{\partial R} - \lambda & -c_2 \\
-c_3 \frac{\partial M}{\partial Y} & c_3 \frac{\partial K}{\partial R} & 0 - \lambda
\end{vmatrix} \\
&= \lambda^3 + \left[-c_1 \left(\frac{\partial I}{\partial Y} - \frac{\partial S}{\partial Y} - \frac{\partial M}{\partial Y} \right) - c_2 \frac{\partial L}{\partial R} \right] \lambda^2 \\
&+ \left\{ c_2 c_3 \frac{\partial K}{\partial R} + c_1 c_2 \left[\left(\frac{\partial I}{\partial Y} - \frac{\partial S}{\partial Y} - \frac{\partial M}{\partial Y} \right) \frac{\partial L}{\partial R} - \left(\frac{\partial I}{\partial R} - \frac{\partial S}{\partial R} \right) \frac{\partial L}{\partial Y} \right] \right\} \lambda \\
&+ c_1 c_2 c_3 \left[- \left(\frac{\partial I}{\partial Y} - \frac{\partial S}{\partial Y} - \frac{\partial M}{\partial Y} \right) \frac{\partial K}{\partial R} - \left(\frac{\partial I}{\partial R} - \frac{\partial S}{\partial R} \right) \frac{\partial M}{\partial Y} \right] \\
&= 0.
\end{aligned}$$

Assuming, for simplicity, $c_i = 1$, the stability conditions are

$$\begin{aligned}
\left[\frac{\partial I}{\partial Y} - \frac{\partial S}{\partial Y} - \frac{\partial M}{\partial Y} \right] &< -\frac{\partial L}{\partial R}, \\
\left[\frac{\partial I}{\partial Y} - \frac{\partial S}{\partial Y} - \frac{\partial M}{\partial Y} \right] &< \frac{\frac{\partial K}{\partial R} - \left(\frac{\partial I}{\partial R} - \frac{\partial S}{\partial R} \right) \frac{\partial L}{\partial Y}}{-\frac{\partial L}{\partial R}}, \\
\left[\frac{\partial I}{\partial Y} - \frac{\partial S}{\partial Y} - \frac{\partial M}{\partial Y} \right] &< \frac{-\left(\frac{\partial I}{\partial R} - \frac{\partial S}{\partial R} \right) \frac{\partial M}{\partial Y}}{\frac{\partial K}{\partial R}},
\end{aligned}$$

and

$$\begin{aligned}
&\left[\left(\frac{\partial M}{\partial Y} - \frac{\partial I}{\partial Y} - \frac{\partial S}{\partial Y} \right) - \frac{\partial L}{\partial R} \right] \left[\frac{\partial K}{\partial R} \left(\frac{\partial I}{\partial Y} - \frac{\partial S}{\partial Y} - \frac{\partial M}{\partial Y} \right) \frac{\partial L}{\partial R} - \left(\frac{\partial I}{\partial R} - \frac{\partial S}{\partial R} \right) \frac{\partial L}{\partial Y} \right] \\
&\qquad - \left[\frac{\partial K}{\partial R} \left(\frac{\partial M}{\partial Y} - \frac{\partial I}{\partial Y} - \frac{\partial S}{\partial Y} \right) - \left(\frac{\partial I}{\partial R} - \frac{\partial S}{\partial R} \right) \frac{\partial M}{\partial Y} \right] > 0.
\end{aligned}$$

Note that the first three inequalities are necessary and sufficient for each real root of the characteristic equation to be negative and therefore exclude monotonic (but not oscillatory) instability; the fourth inequality, on the other hand, is necessary and sufficient, together with the previous ones, for the complex roots to have a negative real part, and therefore to exclude oscillatory instability.

Before performing the comparative static exercise we have to check that

$$J = \begin{vmatrix}
\left(\frac{\partial I}{\partial Y} - \frac{\partial S}{\partial Y} \right) - \frac{\partial M}{\partial Y} & \left(\frac{\partial I}{\partial R} - \frac{\partial S}{\partial R} \right) & 0 \\
\frac{\partial L}{\partial Y} & \frac{\partial L}{\partial R} & -1 \\
\frac{\partial M}{\partial Y} & -\frac{\partial K}{\partial R} & 0
\end{vmatrix} \neq 0.$$

Provided that this condition holds, there exist the functions

$$\begin{aligned}
Y &= (X_0, \alpha_1, G), \\
R &= (X_0, \alpha_1, G), \\
L_s &= (X_0, \alpha_1, G),
\end{aligned}$$

where α_1 is the parameter indicating an exogenous increase in imports ($\partial M / \partial \alpha_1 > 0$), and G is government expenditure, that additively enters the first equilibrium relationship.

20.1.a) To study the consequences an exogenous increase in exports we differentiate the equilibrium relationship with respect to X_0:

$$\begin{aligned}
\left(\frac{\partial I}{\partial Y} - \frac{\partial S}{\partial Y} - \frac{\partial M}{\partial Y} \right) \frac{\partial Y}{\partial X_0} + \left(\frac{\partial I}{\partial R} - \frac{\partial S}{\partial R} \right) \frac{\partial R}{\partial X_0} &= -1, \\
\frac{\partial L}{\partial Y} \frac{\partial Y}{\partial X_0} + \frac{\partial L}{\partial R} \frac{\partial R}{\partial X_0} - \frac{\partial L_s}{\partial X_0} &= 0, \\
\frac{\partial M}{\partial Y} \frac{\partial Y}{\partial X_0} - \frac{\partial K}{\partial R} \frac{\partial R}{\partial X_0} &= 1.
\end{aligned}$$

The solution of this system gives the values for the unknowns $\frac{\partial Y}{\partial X_0}, \frac{\partial R}{\partial X_0}, \frac{\partial L_s}{\partial X_0}$, which turn out to be

$$\begin{aligned}
\frac{\partial Y}{\partial X_0} &= \left[\frac{\partial K}{\partial R} - \left(\frac{\partial I}{\partial R} - \frac{\partial S}{\partial R} \right) \right] / J, \\
\frac{\partial R}{\partial X_0} &= \left[\left(\frac{\partial I}{\partial Y} - \frac{\partial S}{\partial Y} \right) \right] / J, \\
\frac{\partial L_s}{\partial X_0} &= \left\{ \frac{\partial L}{\partial R} \left(\frac{\partial I}{\partial Y} - \frac{\partial S}{\partial Y} \right) + \frac{\partial L}{\partial Y} \left[\frac{\partial K}{\partial R} - \left(\frac{\partial I}{\partial R} - \frac{\partial S}{\partial R} \right) \right] \right\} / J,
\end{aligned}$$

where $J \equiv \frac{\partial M}{\partial Y}\left[\frac{\partial K}{\partial R} - \left(\frac{\partial I}{\partial R} - \frac{\partial S}{\partial R}\right)\right] - \left(\frac{\partial I}{\partial Y} - \frac{\partial S}{\partial Y}\right)\frac{\partial K}{\partial R}$.

Given the third stability condition, J is positive. Once this has been ascertained, we find that $\partial Y/\partial X_0 > 0$, while the signs of $\partial R/\partial X_0$, $\partial L_s/\partial X_o$ remain uncertain. Note, in fact, that $\left(\frac{\partial I}{\partial Y} - \frac{\partial S}{\partial Y}\right) < 0$ is a necessary stability condition in a *closed* economy, but need not hold in an open economy.

20.1.b) Differentiating the equilibrium relationships with respect to α_1 we obtain a system whose solution is

$$\begin{aligned}
\frac{\partial Y}{\partial \alpha_1} &= -\frac{\partial M}{\partial \alpha_1}\left[\frac{\partial K}{\partial R} - \left(\frac{\partial I}{\partial R} - \frac{\partial S}{\partial R}\right)\right]/J\,, \\
\frac{\partial R}{\partial \alpha_1} &= -\frac{\partial M}{\partial \alpha_1}\left(\frac{\partial I}{\partial Y} - \frac{\partial S}{\partial Y}\right)/J\,, \\
\frac{\partial L_s}{\partial \alpha_1} &= \frac{\partial M}{\partial \alpha_1}\left\{\frac{\partial L}{\partial Y}\left[\left(\frac{\partial I}{\partial R} - \frac{\partial S}{\partial R}\right) - \frac{\partial K}{\partial R}\right] - \frac{\partial L}{\partial R}\left(\frac{\partial I}{\partial Y} - \frac{\partial S}{\partial Y}\right)\right\}/J\,.
\end{aligned}$$

Given $J > 0, \partial M/\partial \alpha_1 > 0$, it can be checked that $\partial Y/\partial \alpha_1 < 0$ while the signs of $\partial R/\partial X_0$, $\partial L_s/\partial X_o$ remain uncertain.

20.1.c) To study the consequences of an exogenous variation of government expenditure we differentiate with respect to G and obtain:

$$\begin{aligned}
\frac{\partial Y}{\partial \alpha_2} &= \frac{\partial K}{\partial R}/J, \\
\frac{\partial R}{\partial \alpha_2} &= \frac{\partial M}{\partial Y}/J, \\
\frac{\partial L_s}{\partial \alpha_2} &= \left(\frac{\partial L}{\partial Y}\frac{\partial K}{\partial R} + \frac{\partial L}{\partial R}\frac{\partial M}{\partial Y}\right)/J.
\end{aligned}$$

Given $J > 0$, we can immediately see that $\partial Y/\partial \alpha_2$ and $\partial R/\partial \alpha_2$ are positive, but the sign of $\partial L_s/\partial \alpha_2$ is undefined.

20.2. In the Walrasian case the excess demand function is

$$E(p, \alpha_1, \alpha_2) = D(p, \alpha_1) - S(p, \alpha_2)\,.$$

Provided that

$$\frac{\partial D}{\partial p} - \frac{\partial S}{\partial p} \neq 0,$$

there exists a function

$$p = p(\alpha_1, \alpha_2)$$

of which we want to know the partial derivatives, that give us the reaction of the equilibrium price to a shift in the parameters. Let us first examine the effect of a shift in the tastes parameter, i.e. α_1. Totally differentiating the excess demand function with respect to α_1 we have

$$\frac{\partial D}{\partial p}\frac{\partial p}{\partial \alpha_1} - \frac{\partial S}{\partial p}\frac{\partial p}{\partial \alpha_1} = -\frac{\partial D}{\partial \alpha_1},$$

from which

$$\frac{\partial p}{\partial \alpha_1} = -\frac{\partial D}{\partial \alpha_1}/\left(\frac{\partial D}{\partial p} - \frac{\partial S}{\partial p}\right).$$

Given that the sign of $(\partial D/\partial p - \partial S/\partial p)$ (see Chap. 13) is negative and that $\partial D/\partial \alpha_1$ is positive, we can conclude that a shift in tastes has the effect of increasing the equilibrium price.

We can similarly calculate

$$\frac{\partial p}{\partial \alpha_2} = \frac{\partial S}{\partial \alpha_2}/\left(\frac{\partial D}{\partial p} - \frac{\partial S}{\partial p}\right).$$

Given that $\partial S/\partial \alpha_2 > 0$, we can conclude that a reduction in production costs has the effect of reducing price.

To study the effects of shifts in the two parameters in the Marshallian case, we start from the excess demand price equation

$$E^{-1}(q, \alpha_1, \alpha_2) = p_d(q, \alpha_1) - p_s(q, \alpha_2)\,.$$

Provided that

$$\frac{\partial p_d}{\partial q} - \frac{\partial p_s}{\partial q} \neq 0,$$

there exists a function

$$q = q(\alpha_1, \alpha_2)\,,$$

of which we want to know the partial derivatives, that give us the reaction of the equilibrium quantity to a shift in the parameters. Let us first examine the effect of a shift in the tastes parameter, i.e. α_1. Totally differentiating the excess demand price function with respect to α_1 we have

$$\frac{\partial p_d}{\partial q}\frac{\partial q}{\partial \alpha_1} - \frac{\partial p_s}{\partial q}\frac{\partial q}{\partial \alpha_1} = -\frac{\partial p_d}{\partial \alpha_1},$$

whence

$$\frac{\partial q}{\partial \alpha_1} = -\frac{\partial p_d}{\partial \alpha_1}/\left(\frac{\partial p_d}{\partial q} - \frac{\partial p_s}{\partial q}\right).$$

Knowing that $\partial p_d/\partial \alpha_1 > 0$ and that $(\partial p_d/\partial q - \partial p_s/\partial q) < 0$ (see Chap. 13), we can say that a shift in α_1 causes an increase in q.

We similarly calculate

$$\frac{\partial q}{\partial \alpha_2} = \frac{\partial p_s}{\partial \alpha_2}/\left(\frac{\partial p_d}{\partial q} - \frac{\partial p_s}{\partial q}\right).$$

The effect of a shift in α_2 is an increase in q, given that $\partial p_s/\partial \alpha_2$ is negative.

20.3. To determine the effects of a shift in s_k and s_h on the steady state values of k^*, h^* and y^*, we need to totally differentiate the solution functions (Eqs. 19.27) with respect to s_k and s_h, account being taken that $\mathrm{d}s_h > 0$, $\mathrm{d}s_k + \mathrm{d}s_h = 0$, namely

$$\begin{aligned}
\mathrm{d}k^* &= \frac{\partial k^*}{\partial s_k}\mathrm{d}s_k + \frac{\partial k^*}{\partial s_h}\mathrm{d}s_h = \left(\frac{\partial k^*}{\partial s_h} - \frac{\partial k^*}{\partial s_k}\right)\mathrm{d}s_h \\
\mathrm{d}h^* &= \frac{\partial h^*}{\partial s_k}\mathrm{d}s_k + \frac{\partial h^*}{\partial s_h}\mathrm{d}s_h = \left(\frac{\partial h^*}{\partial s_h} - \frac{\partial h^*}{\partial s_k}\right)\mathrm{d}s_h, \\
\mathrm{d}y^* &= \frac{\partial y^*}{\partial s_k}\mathrm{d}s_k + \frac{\partial y^*}{\partial s_h}\mathrm{d}s_h = \left(\frac{\partial y^*}{\partial s_h} - \frac{\partial y^*}{\partial s_k}\right)\mathrm{d}s_h.
\end{aligned}$$

Thus we get

$$\frac{dk^*}{ds_h} = \frac{1}{(1-\alpha-\beta)(n+g+\delta)}\left(\frac{s_k^{1-\beta}s_h^{\beta}}{n+g+\delta}\right)^{\frac{\alpha+\beta}{1-\alpha-\beta}}\left(\frac{s_k}{s_h}\right)^{1-\beta}\left[\beta+\beta\frac{s_h}{s_k}-\frac{s_h}{s_k}\right] \gtreqless 0,$$

$$\frac{dh^*}{ds_h} = \frac{1}{(1-\alpha-\beta)(n+g+\delta)}\left(\frac{s_k^{\alpha}s_h^{1-\alpha}}{n+g+\delta}\right)^{\frac{\alpha+\beta}{1-\alpha-\beta}}\left(\frac{s_k}{s_h}\right)^{\alpha}\left[1-\alpha+\alpha\frac{s_h}{s_k}\right] > 0,$$

$$\frac{dy^*}{ds_h} = \alpha(k^*)^{\alpha-1}\frac{dk^*}{ds_h}+\beta(h^*)^{\beta-1}\frac{dh^*}{ds_h} \gtreqless 0.$$

Thus h^* will increase, according to economic intuition; it is easy to see that dk^*/ds_h will also be positive if the initial values of s_h, s_k satisfy the inequality

$$\frac{s_h}{s_k} < \frac{\beta}{1-\beta}.$$

Finally, $dy^*/ds_h > 0$ if both dk^*/ds_h and dh^*/ds_h are positive, while its sign remains uncertain when h^* increases but k^* decreases. A shift in investment from physical to human capital is probably (but not certainly) favourable to growth.

Chapter 21

21.1. The Jacobians of the two differential equation systems are

$$\begin{bmatrix} 3y_1^2+y_2^2 & 1+2y_1y_2 \\ -1+2y_1y_2 & y_1^2+3y_2^2 \end{bmatrix},$$

and

$$\begin{bmatrix} -3y_1^2-y_2^2 & 1-2y_1y_2 \\ -1-2y_1y_2 & -y_1^2-3y_2^2 \end{bmatrix}.$$

At the equilibrium point $y_1 = y_2 = 0$, both become

$$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix},$$

with pure imaginary characteristic roots

$$\lambda_{1,2} = \pm i.$$

Hence, according to the Hartman-Grobman theorem, nothing can be said on the local stability of either system.

Let us then consider the original non-linear systems. Starting from the first one, if we follow the hint (multiply through the first equation by y_1, the second equation by y_2, and add), we immediately get

$$y_1y_1' + y_2y_2' = (y_1^2+y_2^2)^2.$$

By a change of variable $r = y_1^2 + y_2^2$ (hence $\frac{dr}{dt} \equiv r' = 2y_1y_1' + 2y_2y_2'$) we have

$$\frac{dr}{dt} - 2r^2 = 0$$

or

$$-r^{-2}dr + 2dt = 0.$$

This is a differential equation with separated variables (see Chap. 24, Eqs. (24.8) and (24.9)) whose solution is

$$-\int r^{-2}dr + 2\int dt = r^{-1} + 2t = A,$$

where A is an arbitrary constant. Hence we get

$$r = \frac{1}{A-2t} = \frac{c}{1-2ct},$$

where $c \equiv 1/A$ is an arbitrary constant. When $t = 0, r_0 = y_1^2(0) + y_2^2(0) = c$, which determines the arbitrary constant.
Since c is a positive value, starting from $t = 0$ we see that $(1-2ct) \to 0$ as t increases toward $1/2c$. Hence $r(t) = y_1^2(t) + y_2^2(t) \to \infty$, and the equilibrium point is not stable.

The second nonlinear system gives rise to

$$y_1y_1' + y_2y_2' = -(y_1^2+y_2^2)^2,$$

from which

$$\frac{dr}{dt} + 2r^2 = 0,$$

that can be solved as above, yielding

$$r = \frac{c}{1+2ct}.$$

This fraction tends to zero as $t \to \infty$, hence $r(t) \to 0$, and the equilibrium point is asymptotically stable.

21.2. With the given signs, we have $\Delta \equiv (a_{11}+a_{22})^2 - 4(a_{11}a_{22} - a_{12}a_{21}) > 0$ and $a_{11}a_{22} - a_{12}a_{21} < 0$. Hence by Table 21.1 the equilibrium point is a saddle.

21.3. We first observe that this non-linear system is autonomous (the function $\mathbf{h}(\mathbf{y})$ does not contain t explicitly), hence its linear part is a uniformly good approximation to the whole system. The characteristic equation of the matrix $\mathbf{A}$ is

$$(-\lambda-1)(-2-\lambda)(-6-\lambda) = 0,$$

hence the roots are $-1, -2, -6$. The system is locally asymptotically stable.

21.4. The null solution is clearly $r_e = 0, \theta_e = \theta_0$. Since $\theta(t; r_0, \theta_0, t_0) = \theta_0$ is a constant, let us examine the solution for r, that is

$$r(t; r_0, \theta_0, t_0) = \frac{r_0}{g(\theta_0, t_0)} \left[\frac{\sin^2 \theta_0}{\sin^4 \theta_0 + (1 - t \sin^2 \theta_0)^2} + \frac{1}{1+t^2} \right].$$

It is immediately evident that this expression tends to zero as $t \to \infty$. Consider now a $t = t_1 = \sin^{-2} \theta_0$ and calculate the value of the expression. We have

$$\begin{aligned} r(t_1; r_0, \theta_0, t_0) &= \frac{r_0}{g(\theta_0, t_0)} \left[\frac{\sin^2 \theta_0}{\sin^4 \theta_0 + (1 - t_1 \sin^2 \theta_0)^2} + \frac{1}{1+t_1^2} \right] \\ &= \frac{r_0}{g(\theta_0, t_0)} \left[\frac{\sin^2 \theta_0}{\sin^4 \theta_0 + (1 - \sin^{-2} \theta_0 \sin^2 \theta_0)^2} + \frac{1}{1+\sin^{-4} \theta_0} \right] \\ &= \frac{r_0}{g(\theta_0, t_0)} \left[\frac{1}{\sin^2 \theta_0} + \frac{1}{1+\sin^{-4} \theta_0} \right]. \end{aligned}$$

Now, for $\theta_0 \to \pm\pi$, we know that $\sin \theta_0 \to 0$, and hence $\sin^2 \theta_0 \to 0$, $\sin^{-4} \theta_0 \to \infty$. It follows that $r(t_1; r_0, \theta_0, t_0) \to \infty$, which means that the condition of Definition 21.1 is not satisfied and the system is not stable. This is due to the fact that the function $T(\mu, \mathrm{y}_0, t_0)$ is not continuous along the ray $\theta_0 = \pm\pi$.

Chapter 22

22.1. Before solving the exercise we define the variables. Following Ramsey, we denote by $x(t)$ and $a(t)$ the total rate of consumption and labour in the economy, and by $c(t)$ its capital. The production function is $y = f(a, c)$. Knowing that $U(x)$ and $V(a)$ are respectively the total rate of utility of consumption and the total rate of disutility of labour, $U(x) - V(a)$ denotes the net enjoyment to be maximized. So, the optimal control problem is

$$\begin{aligned} &\max W = \textstyle\int_0^\infty \{B - [U(x) - V(a)]\} \, \mathrm{d}t, \\ &\text{subject to} \\ &\tfrac{\mathrm{dc}}{\mathrm{d}t} + x = f(a, c). \end{aligned}$$

By performing a change of variable, we can rewrite the functional W as

$$\max W = \int_{c_0}^{\infty} \left\{ \frac{B - [U(x) - V(a)]}{\frac{\mathrm{dc}}{\mathrm{d}t}} \right\} \mathrm{dc},$$

hence

$$\max W = \int_{c_0}^{\infty} \left\{ \frac{B - [U(x) - V(a)]}{f(a, c) - x} \right\} \mathrm{dc}.$$

In this way we have simplified the problem by incorporating the constraint in the objective functional, and the Hamiltonian coincides with the intermediate function of the objective functional. Hence considering x as the control variable the maximum principle gives

$$\frac{\partial H}{\partial x} = \frac{\partial I}{\partial x} = \frac{-u(x)[f(a, c) - x] + [B - U(x) + V(a)]}{[f(a, c) - x]^2} = 0,$$

where $u(x) \equiv \mathrm{d}U/\mathrm{d}x$ is the marginal utility of consumption. Thus we get

$$u(x) \frac{\mathrm{dc}}{\mathrm{d}t} = B - [U(x) - V(a)],$$

that is the Ramsey rule.

In the original Ramsey model the discounting factor is not present, so that current and future consumption have the same weight. Differently, in the model with discounting, consumption increases, remains constant or decreases, depending on whether the marginal product of capital exceeds, is equal to or is less than the rate of preference. This is very important because the higher is the marginal product of capital relative to the rate of time preference, the more it pays to depress the current level of consumption in order to enjoy higher consumption later.

22.2.a) The optimal growth problem with final terminal time is the same as Eqs. (22.18), except that the upper limit of the welfare integral is t_1, a given finite parameter, and there is an additional condition that the terminal stock of capital per worker, i.e. $k(t_1)$, must be no less than some given level, k_1. The problem is solved as in the infinite horizon case and the canonical equations (22.25) are still applicable. In this case a terminal condition on the costate variable has to be added. Applying the transversality condition (22.6), where $x(t_1) = k(t_1) - k_1$, we get the relation stated in the exercise.

22.2.b) The *turnpike theorem* was originally elaborated by DOSSO (Dorfman, Samuelson, Solow, 1958, *Linear Programming and Economic Analysis*, New York, McGraw-Hill, Chap.12) in essentially pure-production multisector models with linear technology. This exercise is concerned with its extension to the properties of optimum growth paths in the neoclassical aggregate model where the value of consumption as such is accounted for, an extension due to David Cass. For the (rather lengthy) proof we refer the reader to the paper by Cass (D. Cass, 1966, Optimum Growth in an Aggregative Model of Capital Accumulation: A Turnpike Theorem, *Econometrica* 34, 833-850).

22.3. We know from Sect. 13.2.1 that

$$y = f(r, 1).$$

Considering the deviation from the equilibrium point, we can write

$$\bar{y} = f(r, 1) - f(r_e, 1) = [f(r_0, 1) - f(r_e, 1)] \, e^{-\beta t},$$

where β is the stable root of Eq. (22.29).

Dividing it by $f(r, 1)$ we have

$$\frac{f(r, 1) - f(r_e, 1)}{f(r_e, 1)} \simeq \ln \left[\frac{f(r, 1)}{f(r_e, 1)} \right] = \ln \left[\frac{f(r_0, 1)}{f(r_e, 1)} \right] e^{-\beta t}.$$

Assuming the difference $f(r, 1) - f(r_e, 1)$ very small, the approximation is quite good. So, we have

$$\ln f(r, 1) = \left(1 - e^{-\beta t}\right) f(r_e, 1) + \ln f(r_0, 1) \, e^{-\beta t}.$$

The convergence coefficient β for y is the same as that for r that we found in Sect. 13.2.4.1. Its size depends on the behaviour of the capital share in the production function and it indicates how rapidly an economy's output per worker, y, approaches its steady-state value, y_e.

22.4. We can adopt an utility function with constant substitution elasticity of the type

$$u(c(t)) = \frac{c(t)^{1-\theta}-1}{1-\theta},$$

where $c(t)$ is given by

$$c(t) = c_0 e^{(1/\theta)(A-\delta-\rho)t}.$$

By substituting these functions in the objective functional, we obtain

$$\int_0^{\infty} \frac{1}{1-\theta}\left[c_0 e^{-(\rho-n)t+\frac{1-\theta}{\theta}(A-\delta-\rho)t}-1\right] \mathrm{d}t.$$

For the integral to converge, the following condition must hold

$$-(\rho-n)t+\frac{1-\theta}{\theta}(A-\delta-\rho)t<0,$$

whence, for $t>0$

$$-(\rho-n)+\frac{1-\theta}{\theta}(A-\delta-\rho)<0,$$

from which

$$\rho>\frac{1-\theta}{\theta}(A-\delta-\rho)+n.$$

By adding δ to both sides, we obtain

$$\rho+\delta>\frac{1-\theta}{\theta}(A-\delta-\rho)+\delta+n.$$

To show that this inequality ensures $\varphi>0$, we can rewrite it as follows

$$\rho+\frac{1-\theta}{\theta}\rho>\frac{1-\theta}{\theta}(A-\delta)+n.$$

Considering Eq. (22.46), we can rewrite this last inequality, obtaining

$$\frac{\rho}{\theta}>\frac{1-\theta}{\theta}(A-\delta)+n,$$

that implies

$$0>\frac{1-\theta}{\theta}(A-\delta)-\frac{\rho}{\theta}+n=-\varphi.$$

Hence

$$\varphi>0.$$

22.5. From the definition of u,p we get $u'=m'-p', c'=e'-p'$. Substituting m',p',e' from the dynamic equations of the model, we have

$$\begin{aligned} u' &= -\theta(y-\overline{y}),\\ c' &= r-r^*-\theta(y-\overline{y})-\pi. \end{aligned}$$

The variable $(y-\overline{y})$ is already given in terms of c and the exogenous variables by the second equation (the IS schedule), hence

$$\begin{aligned} u' &= -\theta\delta c-\theta\psi R_{\infty},\\ c' &= r-r^*-\theta\delta c-\theta\psi R_{\infty}-\pi. \end{aligned}$$

We now substitute $y=\overline{y}+\delta c+\psi R_{\infty}$ from the IS into the LM schedule to obtain r in terms of u, c and the exogenous variables

$$r=[-u+k\delta c+k(\overline{y}+R+\psi R_{\infty})]/\lambda.$$

Finally we can substitute this expression into the differential equation for c', and obtain the system

$$\begin{aligned} u' &= -\theta\delta c-\theta\psi R_{\infty},\\ c' &= -\lambda^{-1}u+(k\lambda^{-1}-\theta)\delta c+[\psi R_{\infty}(k\lambda^{-1}-\theta)-\pi-r^*+k(\overline{y}+R)\lambda^{-1}]. \end{aligned}$$

By letting $u'=c'=0$ we determine the equilibrium point $c_e=-\psi R_{\infty}/\delta, u_e=k(\overline{y}+R)-\lambda(\pi+r^*)$.

The matrix of the homogeneous part of this differential system is

$$\begin{bmatrix} 0 & -\theta\delta \\ -\lambda^{-1} & (k\lambda^{-1}-\theta)\delta \end{bmatrix}$$

with characteristic equation

$$\mu^2+(k\lambda^{-1}-\theta)\delta\mu-\theta\delta\lambda^{-1}=0.$$

Since the succession of the coefficient signs is $+?-$, the roots will be real and with opposite sign, hence a saddle point.

22.6. Before the oil discovery, the equilibrium point is at $c_e=0, u_e=k\overline{y}-\lambda(\pi+r^*)$. After the oil discovery, the new equilibrium point is as found in the previous exercise, namely $c_e=-\psi R_{\infty}/\delta, u_e=k(\overline{y}+R)-\lambda(\pi+r^*)$. The pre-oil equilibrium point can be taken as the initial point of the post-oil system. Given the saddle-point stability of the model, the real exchange rate will jump on the stable arm of the saddle and move along it toward the new equilibrium point.

To determine the nature of this jump we have to determine the slope of the stable arm. This can be done along the lines treated in Chap. 21, Sect. 21.3.2.4, Eq. (21.17). Letting $\mu_1<0$ be the stable root, the slope of the stable arm referred to the new equilibrium point will be

$$\frac{c}{u}=\frac{\mu_1}{-\theta\delta}>0.$$

The real exchange rate, to go from the initial value of zero to the new (negative) equilibrium value will have to drop to a value on the stable arm which is below the new equilibrium value, after which it will start increasing towards equilibrium. This is a case of undershooting.

Chapter 23

23.1. From the suggested Liapunov function $V = \frac{1}{2c}(z_1 - 1)^2 + \frac{1}{2c}(z_2 - 1)^2$ we have

$$\frac{\mathrm{d}V}{\mathrm{d}t} = \frac{1}{c}(z_1 - 1)z_1' + \frac{1}{c}(z_2 - 1)z_2'$$

hence—see Chap. 19, Eqs. (19.30)—

$$\frac{\mathrm{d}V}{\mathrm{d}t} = (z_1 - 1)(z_1^\alpha z_2^\beta - z_1) + (z_2 - 1)(z_1^\alpha z_2^\beta - z_2).$$

It is now easy to show that this expression is always negative (except of course at the equilibrium point $z_1 = z_2 = 1$, where it is zero) by using Fig. 19.1. Consider, for example, points like A in that figure. In Sect. 19.2 we have shown that

$$z_1^\alpha z_2^\beta - z_1 > 0, z_1^\alpha z_2^\beta - z_2 > 0$$

at point A, where we also have $z_1 < 1, z_2 < 1$ or $z_1 - 1 < 0, z_2 - 1 < 0$. It follows that

$$(z_1 - 1)(z_1^\alpha z_2^\beta - z_1) < 0, (z_2 - 1)(z_1^\alpha z_2^\beta - z_2) < 0,$$

hence $\mathrm{d}V/\mathrm{d}t < 0$.

Similarly it can be shown that at points like B we have $z_1 > 1, z_2 < 1$, $z_1^\alpha z_2^\beta - z_1 < 0, z_1^\alpha z_2^\beta - z_2 > 0$, hence $\mathrm{d}V/\mathrm{d}t < 0$. Other points can be examined in like manner.

23.2 Given the mathematical structure of the model, in which both y_{1t} and y_{2t} depend on y_{1t-1}, it is sufficient to examine the difference equation $y_{1t} = h(y_{1t-1})$. In fact, if y_1 (the leader's output) tends to a stationary point, so will y_2 (the follower's output). This is consistent with the economics of the leader-follower model.

Let us first compute the derivative of the function h with respect to its argument, i.e.

$$h' \equiv \mathrm{d}y_{1t}/\mathrm{d}y_{1t-1}.$$

It is easier to compute $\mathrm{d}y_{1t}/\mathrm{d}y_{1t-1}$ as $\mathrm{d}y_{1t}/\mathrm{d}y_{1t-1} = (\mathrm{d}y_{2t}/\mathrm{d}y_{1t-1})(\mathrm{d}y_{1t}/\mathrm{d}y_{2t})$, since these two derivatives can be obtained by applying the implicit function rule to the equations defining the first-order conditions. Totally differentiating the first-order condition for the follower we get

$$\frac{\mathrm{d}^2 f}{\mathrm{d}Y^2}\left(1 + \frac{\mathrm{d}y_{2t}}{\mathrm{d}y_{1t-1}}\right) y_{2t} + \frac{\mathrm{d}f}{\mathrm{d}Y}\frac{\mathrm{d}y_{2t}}{\mathrm{d}y_{1t-1}} + \frac{\mathrm{d}f}{\mathrm{d}Y}\left(1 + \frac{\mathrm{d}y_{2t}}{\mathrm{d}y_{1t-1}}\right) - \frac{\mathrm{d}^2 C_2}{\mathrm{d}y_{2t}^2}\frac{\mathrm{d}y_{2t}}{\mathrm{d}y_{1t-1}} = 0,$$

which gives

$$\frac{\mathrm{d}y_{2t}}{\mathrm{d}y_{1t-1}} = -\frac{y_{2t}\frac{\mathrm{d}^2 f}{\mathrm{d}Y^2} + \frac{\mathrm{d}f}{\mathrm{d}Y}}{\left(y_{2t}\frac{\mathrm{d}^2 f}{\mathrm{d}Y^2} + \frac{\mathrm{d}f}{\mathrm{d}Y}\right) + \left(\frac{\mathrm{d}f}{\mathrm{d}Y} - \frac{\mathrm{d}^2 C_2}{\mathrm{d}y_{2t}^2}\right)}.$$

Totally differentiating the first-order conditions for the leader we similarly obtain

$$\frac{\mathrm{d}y_{1t}}{\mathrm{d}y_{2t}} = \frac{y_{1t}\frac{\mathrm{d}^2 f}{\mathrm{d}Y^2} + \frac{\mathrm{d}f}{\mathrm{d}Y}}{\left(y_{1t}\frac{\mathrm{d}^2 f}{\mathrm{d}Y^2} + \frac{\mathrm{d}f}{\mathrm{d}Y}\right) + \left(\frac{\mathrm{d}f}{\mathrm{d}Y} - \frac{\mathrm{d}^2 C_1}{\mathrm{d}y_{1t}^2}\right)}.$$

Under 'normal' conditions—namely $\mathrm{d}f/\mathrm{d}Y < 0$ (downward-sloping demand curve), $\mathrm{d}^2C_i/\mathrm{d}y_i > 0$ (increasing marginal cost), $y_{it}(\mathrm{d}^2f/\mathrm{d}Y^2)+\mathrm{d}f/\mathrm{d}Y > 0$ (the marginal revenue of any firm is decreasing with respect to the other firm's output)—we see that in the expressions above both numerator and denominator are negative, the denominator being greater in absolute value than the numerator.

Hence both $\mathrm{d}y_{2t}/\mathrm{d}y_{1t-1}$ and $\mathrm{d}y_{1t}/\mathrm{d}y_{2t}$ are always smaller than one in absolute value, which means that

$$|h'| < 1$$

everywhere.

We can now pass to the study of global stability. Using the mean value theorem (Okuguchi, 1976) we have

$$|y_{1t} - y_{1t-1}| = |h'|\,|y_{1t-1} - y_{1t-2}|,$$

where h' is computed at an intermediate value between y_{1t-1} and y_{1t-2}. Since we have just shown that $|h'| < 1$, it follows from the Wu and Brown theorem—see Eq. (23.9), where we are using the absolute value norm—that global stability obtains.

Chapter 24

24.1. With $p = \frac{1}{2}$, the given production function becomes

$$Y = a^2K + L + 2a(KL)^{\frac{1}{2}}.$$

Consequently, the basic dynamic equation of the model—see Eq. (13.34)—becomes

$$r' = s\left(a\sqrt{r} + 1\right)^2 - nr,$$

where the square root has to be taken with the positive sign. This equation can be rewritten as

$$r' = s\left[\left(a^2 - \frac{n}{s}\right)r + 2a\sqrt{r} + 1\right].$$

24.1.a) If we define

$$A \equiv a - \left(\frac{n}{s}\right)^{\frac{1}{2}},$$

$$B \equiv a + \left(\frac{n}{s}\right)^{\frac{1}{2}},$$

we can rewrite the equation as

$$r' = s\left(A\sqrt{r} + 1\right)\left(B\sqrt{r} + 1\right).$$

24.1.b) To solve it, we simply change the notation of the time derivative from r' to $\mathrm{d}r/\mathrm{d}t$ and rewrite the equation as

$$\frac{1}{s\left(A\sqrt{r} + 1\right)\left(B\sqrt{r} + 1\right)}\mathrm{d}r - \mathrm{d}t = 0,$$

which has separated variables. To simplify the integration, let us multiply through by $(ns)^{\frac{1}{2}}$, obtaining

$$\frac{(ns)^{\frac{1}{2}}}{s(A\sqrt{r}+1)(B\sqrt{r}+1)}\mathrm{d}r-(ns)^{\frac{1}{2}}\,\mathrm{d}t=0,$$

whose solution is

$$-\int (ns)^{\frac{1}{2}}\,\mathrm{d}t+\int\frac{(ns)^{\frac{1}{2}}}{s(A\sqrt{r}+1)(B\sqrt{r}+1)}\mathrm{d}r=C.$$

The second integral can also be explicitly solved, so that we obtain

$$-(ns)^{\frac{1}{2}}t+\left[\frac{1}{A}\ln(A\sqrt{r}+1)-\frac{1}{B}\ln(B\sqrt{r}+1)\right]=C,$$

from which

$$(A\sqrt{r}+1)^{\frac{1}{A}}(B\sqrt{r}+1)^{-\frac{1}{B}}=D\exp\left[(ns)^{\frac{1}{2}}t\right],$$

where $D\equiv e^{-C}$ is the new arbitrary constant. Assuming $r=r_0$ known for $t=0$, from this last equation we have

$$D=(A\sqrt{r_0}+1)^{\frac{1}{A}}(B\sqrt{r_0}+1)^{-\frac{1}{B}},$$

and so,

$$\left(\frac{A\sqrt{r}+1}{A\sqrt{r_0}+1}\right)^{\frac{1}{A}}\left(\frac{B\sqrt{r}+1}{B\sqrt{r_0}+1}\right)^{-\frac{1}{B}}=\exp\left[(ns)^{\frac{1}{2}}t\right],$$

which is the final form of the solution. From this equation it is not possible to make r explicit as a function of t, so that the solution is not very useful for obtaining in a simple way a clear picture of the behaviour of r over time. This, however, can be obtained easily by means of the phase diagram (see the next exercise).

24.2. We first determine the equilibrium point(s) by setting $r'=0$, which gives $s(a\sqrt{r}+1)^2-nr=0$. Rearranging terms we have

$$\left(a^2-\frac{n}{s}\right)r+2a\sqrt{r}+1=0.$$

Let $x\equiv+\sqrt{r}>0$ (remember that the square root has to be taken with the positive sign) and solve the quadratic

$$\left(a^2-\frac{n}{s}\right)x^2+2ax+1=0.$$

It can immediately be seen that if $a^2-n/s\geq 0$, there is no positive value of x (hence no real value of r) that satisfies this equation. If, on the other hand, $a^2-n/s<0$, then the equation has one positive and one negative root. The positive root is $x=1/[(n/s)^{1/2}-a]$, which gives

$$r_e=\left[\left(\frac{n}{s}\right)^{\frac{1}{2}}-a\right]^{-2}.$$

To build the phase diagram, consider the function

$$f(r)=s\left[\left(a^2-\frac{n}{s}\right)r+2a\sqrt{r}+1\right].$$

We have

$$\begin{aligned} f(0) &= s>0,\\ \frac{\mathrm{d}f}{\mathrm{d}r} &= s\left[(a^2-n/s)+a/\sqrt{r}\right],\\ \frac{\mathrm{d}^2f}{\mathrm{d}r^2} &= -as/2\sqrt{r^3}. \end{aligned}$$

The second derivative is always negative for $r\geq 0$. As regards the first derivative, we can distinguish two cases:

(a) it is always positive for $r\geq 0$ if $a^2-n/s>0$. This means that the function is monotonically increasing from s, hence it never intersects the horizontal axis (r–axis) at positive values of r. This confirms what we already know, that in this case there is no real positive root.

(b) In the economically meaningful case $a^2-n/s<0$, the first derivative is positive, zero, negative, respectively, for $r \lesseqgtr a^2/\left(n/s-a^2\right)^2$, $r\geq 0$. The critical value $a^2/\left(n/s-a^2\right)^2$ is smaller than r_e[1], so that we have a diagram in which the function $f(r)$ starts at the ordinate s for $r=0$, increases up to its maximum for $r=a^2/\left(n/s-a^2\right)^2$, then starts decreasing and cuts the r–axis from above at $r=r_e$, after which it decreases toward $-\infty$.

From this phase diagram it can easily be seen that in case (a) r increases beyond all bounds, whereas in case (b) it tends monotonically from both sides to the equilibrium point r_e , which is therefore monotonically stable.

Let us note that this exercise, together with exercise 1, provide an example of those cases in which, although the differential equation is explicitly integrable, qualitative methods yield better results. This is due to the fact that the explicit function $r=r(t)$ cannot be obtained from the integral , as we remarked above (see also Chap. 21, Sect. 21.3.1 remark (2)).

24.3. From the model we obtain the equations

$$\begin{aligned} Y &= \frac{1}{1-b}K'+\frac{a}{1-b},\\ K^* &= \frac{k}{1-b}K'+\frac{ka}{1-b}, \end{aligned}$$

[1] Consider the inequality

$$\frac{a^2}{(a^2-n/s)^2}<\frac{1}{\left[(n/s)^{\frac{1}{2}}-a\right]^2},$$

which gives $a^2\left[(n/s)^{\frac{1}{2}}-a\right]^2<(a^2-n/s)^2$. This can be written as $a^2\left[(n/s)^{\frac{1}{2}}-a\right]^2<\left\{\left[(n/s)^{\frac{1}{2}}-a\right]\left[(n/s)^{\frac{1}{2}}+a\right]\right\}^2$, that is, $a^2<\left[(n/s)^{\frac{1}{2}}+a\right]^2$, which is certainly satisfied.

where K' is as given in the exercise.

We can immediately note that the model has an equilibrium point for $K' = 0$, so that $Y_e = a/(1-b)$ and $K_e = kY_e$. This equilibrium, however, is not stable, although it cannot be considered unstable either. In other words, any deviation from equilibrium gives rise to a constant-amplitude oscillation. The phase diagram of the model is drawn in Fig. 24.10. Suppose that the initial point is A. The capital stock in existence is smaller then the desired capital stock, so that investment takes place at the rate K_1; from the equations written above we have that income is

$$Y = \frac{1}{1-b}K_1 + \frac{a}{1-b},$$

and consequently the desired capital stock is

$$K^* = \frac{k}{1-b}K_1 + \frac{ka}{1-b}.$$

The actual capital stock increases and passes through K_e, but cannot stop there because we still have $K' = K_1 > 0$ (point E). Thus the capital stock goes on increasing up to the point in which it reaches its desired level (corresponding to point B), where K' falls sharply to zero given the assumptions. But then income falls to $a/(1-b)$ and the desired capital stock to $ka/(1-b)$, which is now smaller than the existing capital stock. Therefore K' falls to K_2 (point C); consequently, income falls to $(a+K_2)/(1-b)$ and the desired capital stock to $k(a+K_2)/(1-b)$.

The desired capital stock goes on decreasing: it passes through K_e, but it cannot remain there since K' is still negative ($K' = K_2 < 0$: point F), so that it decreases more and more until it reaches its desired value $k(a+K_2)/(1-b)$, corresponding to point D. Here K' falls to zero, Y increases to $a/(1-b)$ and the desired capital stock is $ka/(1-b)$, so that K' rises immediately to K_1, and we are again at point A.

This shows that in the phase diagram we have a continuous movement in the segments $\overline{AB}$ and $\overline{CD}$, with a discontinuous jump from B to C and from D to A.

24.4

24.4.a) The equilibrium point is obtained letting $p' = q' = 0$, whence

$$\frac{\varepsilon}{\varepsilon-1}q - \delta q^3 = q(\frac{\varepsilon}{\varepsilon-1} - \delta q^2) = 0.$$

Since $0 < \varepsilon < 1$, the only real root is $q = 0$, whence $p = 0$.

The Jacobian of the system is

$$\begin{bmatrix} -1+\varepsilon & -1 \\ 1 & 1-3\delta q^2 \end{bmatrix}.$$

Evaluating its characteristic equation at the origin we have

$$\lambda^2 - \varepsilon\lambda + \varepsilon = 0,$$

with roots

$$\lambda_{1,2} = \frac{\varepsilon \pm \sqrt{\varepsilon(\varepsilon-4)}}{2}.$$

These are complex conjugate with positive real part, hence the origin is an unstable focus (see Table 21.1).

24.4.b) Applying remark (ii) of the Poincaré-Bendixson positive criterion, it only remains to find C_2. Setting it as suggested in the hint, we then consider the Liapunov function $V = \frac{1}{2}(p^2+q^2)$, from which

$$\begin{aligned} \frac{dV}{dt} &= p\frac{dp}{dt} + q\frac{dq}{dt} = p[(-1+\varepsilon)p - q] + q[p - (\delta q^3 - q)] = q^2 - \delta q^4 - (1-\varepsilon)p^2 \\ &= \frac{1}{4\delta} - \left[\delta\left(q^2 - \frac{1}{2\delta}\right)^2 + (1-\varepsilon)p^2\right]. \end{aligned}$$

We begin by observing that both terms in square brackets are non-negative, hence if either of them is greater than $1/4\delta$, then $dV/dt < 0$. Now, $\delta\left(q^2 - \frac{1}{2\delta}\right)^2 > 1/4\delta$ if $q^2 > 1/\delta$, and $(1-\varepsilon)p^2 > 1/4\delta$ if $p^2 > 1/4\delta(1-\varepsilon)$.

It follows that a sufficient condition for $dV/dt < 0$ is $p^2+q^2 > 1/4\delta(1-\varepsilon)+1/\delta$ or

$$p^2 + q^2 > \frac{5-4\varepsilon}{4\delta(1-\varepsilon)} = \rho_c.$$

Hence, if we take C_2 as the circle given in the exercise, it follows from the Liapunov function that all trajectories coming from outside C_2 pass through the points of C_2 for t increasing and penetrate in the circular region D included between C_2 and the (arbitrarily small) circle C_1 drawn around the origin. Since there is no singular point in D (the only singular point is the origin), at least one limit cycle exists in D.

24.4.c) Application of Nemitzky's theorem gives rise to the expression

$$N = p[(-1+\varepsilon)p - q] + q[p - (\delta q^3 - q)].$$

It is immediately obvious that this expression coincides with dV/dt examined in 4.b). Hence $N < 0$ for $p^2 + q^2 = \rho_2 > \rho_c$. Similarly it can be shown that $N > 0$ for $p^2 + q^2 = \rho_1$, where ρ_1 is taken sufficiently smaller than ρ_c, for example $\frac{1}{2}\rho_c$. It is sufficient to observe that if we take $q^2 < 1/2\delta$ and $p^2 < 1/8\delta(1-\varepsilon)$, then $\delta\left(q^2 - \frac{1}{2\delta}\right)^2 < (1/2)(1/4\delta)$ and $(1-\varepsilon)p^2 < (1/2)(1/4\delta)$, hence

$$\left[\delta\left(q^2 - \frac{1}{2\delta}\right)^2 + (1-\varepsilon)p^2\right] < \frac{1}{4\delta},$$

which means that N is positive. Thus a sufficient condition for $N > 0$ is $p^2 + q^2 < (1/2)[1/4\delta(1-\varepsilon) + 1/\delta]$ or

$$p^2 + q^2 < \frac{1}{2}\frac{5-4\varepsilon}{4\delta(1-\varepsilon)} = \frac{1}{2}\rho_c.$$

Since there is no singular point in the circular ring D between the two circles of radii ρ_1 and ρ_2, there is a stable limit cycle in D.

24.5. The Jacobian of the Lotka-Volterra system is

$$\begin{bmatrix} a_1 - b_1 u & -b_1 v \\ b_2 u & -(a_2 - b_2 v) \end{bmatrix}.$$

This becomes, when evaluated at the singular point $u = a_1/b_1, v = a_2/b_2$,

$$\begin{bmatrix} 0 & -b_1 a_2/b_2 \\ b_2 a_1/b_1 & 0 \end{bmatrix}.$$

whose characteristic equation is

$$\lambda^2 + a_1 a_2 = 0.$$

The roots are pure imaginary, hence the singular point is locally a centre.

Chapter 25

25.1. Let $v[D(r,\alpha) - S(r)] = f(r,\alpha)$, and assume that the demand and supply curves become tangent for the value α_0 of the parameter (with a corresponding value r_e of the equilibrium exchange rate), hence $\partial D(r_{e,0},\alpha_0)/\partial r = \partial S(r_{e,0})/\partial r$. Then we compute:

(1.a) $\partial f(r_{e,0},\alpha_0)/\partial r = v[\partial D(r_{e,0},\alpha_0)/\partial r - \partial S(r_{e,0})/\partial r] = 0,$

(1.b) $\partial^2 f(r_{e,0},\alpha_0)/\partial r^2 = v[\partial^2 D(r_{e,0},\alpha_0)/\partial r^2 - \partial^2 S(r_{e,0})/\partial r^2]$
$= -v\partial^2 S(r_{e,0})/\partial r^2 \neq 0,$

(1.c) $\partial f(r_{e,0},\alpha_0)/\partial \alpha > 0,$

where we have used the data of the problem. It follows that the conditions of definition 1 in Sect. 25.2.1 are satisfied.

25.2. With the introduction of depreciation, the characteristic equation of the model (see Sect. 25.2.4) becomes

$$\begin{aligned} &\lambda^2 + a_1\lambda + a_2 = 0, \\ &\quad \text{where} \\ &a_1 \equiv -[\alpha(I_Y - S_Y) + I_K - \delta], \\ &a_2 \equiv \alpha[(I_Y - S_Y)(I_K - \delta) - I_Y(I_K - S_K)]. \end{aligned}$$

A necessary condition for complex roots is $a_2 > 0$. which gives

$$\alpha(1-\delta)(I_Y - S_Y) > 0.$$

Since $(I_Y - S_Y) > 0$, $a_2 > 0$ requires $\delta < 1$. This is the critical value of δ (note that it is not obvious that $\delta < 1$, as this partly depends on the time unit chosen).

The rest of the analysis proceeds as in the text.

25.3. If we consider the characteristic equation (25.40), the condition for the roots to be complex is

$$\Delta = \left[\left(\frac{\partial y_1}{\partial k_1} - \sigma\right) + \left(\frac{\partial y_2}{\partial k_2} - \sigma\right)\right]^2 - 4\left[\left(\frac{\partial y_1}{\partial k_1} - \sigma\right)\left(\frac{\partial y_2}{\partial k_2} - \sigma\right) - \frac{\partial y_1}{\partial k_2}\frac{\partial y_2}{\partial k_1}\right] < 0.$$

This can be written as

$$\left[\left(\frac{\partial y_1}{\partial k_1} - \sigma\right) - \left(\frac{\partial y_2}{\partial k_2} - \sigma\right)\right]^2 + 4\frac{\partial y_1}{\partial k_2}\frac{\partial y_2}{\partial k_1} < 0,$$

whence

$$4\frac{\partial y_1}{\partial k_2}\frac{\partial y_2}{\partial k_1} < -\left[\frac{\partial y_1}{\partial k_1} - \frac{\partial y_2}{\partial k_2}\right]^2.$$

25.4. Performing a linear approximation at the equilibrium point (where $I = S, L = L_s$) we obtain the Jacobian

$$\begin{bmatrix} \alpha(I_Y - S_Y) & \alpha(I_R - S_R) \\ \beta L_Y & \beta L_R \end{bmatrix},$$

whose characteristic equation is

$$\begin{aligned} &\lambda^2 + a_1\lambda + a_2 = 0, \\ &\quad \text{where} \\ &a_1 \equiv -[\alpha(I_Y - S_Y) + \beta L_R], \\ &a_2 \equiv \alpha\beta[(I_Y - S_Y)L_R - (I_R - S_R)L_Y]. \end{aligned}$$

The signs of a_1, a_2 are uncertain, since—given the assumption $I_Y - S_Y > 0$— they both contain positive as well as negative terms. Hence the equilibrium point may be stable as well as unstable.

A necessary condition for the presence of complex roots is $a_2 > 0$, namely

$$-(I_R - S_R)/(I_Y - S_Y) > -L_R/L_Y,$$

which means that the IS schedule is not only positively sloped (due to the Kaldor assumption) but has a slope greater than the slope of the LM schedule.

As regards a_1, we have two degrees of freedom (the parameters α, β). Let us for example take $\gamma = \alpha/\beta$ as a parameter. Then

$$a_1 = -\beta[\gamma(I_Y - S_Y) + L_R] \gtreqless 0 \text{ according as } \gamma \lesseqgtr \gamma_0, \quad \gamma_0 \equiv \frac{-L_R}{I_Y - S_Y} > 0.$$

Let us then consider the roots of the Jacobian,

$$\lambda_{1,2} = -\frac{1}{2}a_1 \pm \frac{1}{2}\sqrt{a_1^2 - 4a_2}.$$

Since $a_1 = 0$ for the critical value γ_0 of the parameter, the Jacobian has a pair of pure imaginary roots there, $\lambda_{1,2} = \pm i\sqrt{a_2}$. The roots will remain complex conjugate for a_1 sufficiently small, namely for γ sufficiently near to γ_0. This satisfies assumption (H.1) of the Hopf bifurcation theorem.

It is also easy to check that

$$\left.\frac{\mathrm{d}\left(-\frac{1}{2}a_1\right)}{\mathrm{d}\gamma}\right|_{\gamma=\gamma_0} = \frac{1}{2}\beta(I_Y - S_Y) > 0,$$

which satisfies assumption (H.2). The existence of a closed orbit is proved.

25.5. The Jacobian of the system is

$$\begin{bmatrix} -\gamma x_2^2 & -2\gamma x_1 x_2 + \gamma\phi \\ x_2^2 & 2x_1x_2 - 1 \end{bmatrix}.$$

This Jacobian, evaluated at the stationary point of the system $x_1 = x_2 = 1$, becomes

$$\begin{bmatrix} -\gamma & \gamma(\phi - 2) \\ 1 & 1 \end{bmatrix}$$

with characteristic equation

$$\lambda^2 + a_1\lambda + a_2 = 0,$$
where
$$a_1 \equiv 1 - \gamma,$$
$$a_2 \equiv \gamma(1 - \phi).$$

Given the definition of $\phi = \beta/\delta = \beta/(\varepsilon + \beta) < 1$, it follows that $a_2 > 0$, hence the necessary condition for the roots to be complex is satisfied. As regards a_1, it is easy to see that

$$a_1 \gtreqless 0 \text{ according as } \gamma \lesseqgtr \gamma_0, \quad \gamma_0 \equiv 1.$$

Let us then consider the roots of the Jacobian,

$$\lambda_{1,2} = -\frac{1}{2}a_1 \pm \frac{1}{2}\sqrt{a_1^2 - 4a_2}.$$

Since $a_1 = 0$ for the critical value γ_0 of the parameter, the Jacobian has a pair of pure imaginary roots there, $\lambda_{1,2} = \pm \mathrm{i}\sqrt{a_2}$. The roots will remain complex conjugate for a_1 sufficiently small, namely for γ sufficiently near to 1. This satisfies assumption (H.1) of the Hopf bifurcation theorem.

It is also easy to check that

$$\left.\frac{\mathrm{d}\left(-\frac{1}{2}a_1\right)}{\mathrm{d}\gamma}\right|_{\gamma=\gamma_0} = \frac{1}{2} > 0,$$

which satisfies assumption (H.2). The existence of a closed orbit is proved.

Chapter 26

26.1. When the function $f(k)$ is well behaved, the function $sf(k)$ will have the same form as the function $sf(.)$ in Fig. 13.2 (with k in the place of r). Hence:

1.a) the intersection between this function and the straight line $(1+n)k$, which determines the equilibrium point $k_{t+1} = k_t = k_e$, exists and is unique in the positive quadrant.

1.b) Let us take a linear approximation at the equilibrium point, which gives the first-order linear difference equation

$$\overline{k}_{t+1} = \frac{s\varphi}{1+n}\overline{k}_t,$$

where $\varphi \equiv (\mathrm{d}f/\mathrm{d}k)$ evaluated at the equilibrium point, and the overbar denotes the deviations from equilibrium. From the diagram recalled in 1.a it is easy to see that, at the equilibrium point, the slope of the function $sf(k)$ is smaller than the slope of the straight line, hence $s\varphi < 1+n$, i.e., $[s\varphi/(1+n] < 1$. It follows that the solution to the linear difference equation (Chap. 3, Sect. 3.1) is monotonically stable.

1.c) Consider now the original nonlinear difference equation

$$k_{t+1} = \frac{s}{1+n}f(k_t).$$

The function $[s/(1+n)]f(k_t)$ will have the same aspect as the function $sf(k_t)$. By using a phase diagram (see, for example, Fig. 21.14) it is easy to check global stability.

26.2. The linear approximation to the Lorenz model (26.24) gives rise to the following linear system

$$\begin{aligned} \overline{x}' &= -\sigma\overline{x} + \sigma\overline{y}, \\ \overline{y}' &= (\rho - z_e)\overline{x} - \overline{y} - x_e\overline{z}, \\ \overline{z}' &= y_e\overline{x} + x_e\overline{y} - \beta\overline{z}, \end{aligned}$$

where x_e, y_e, z_e are evaluated at the equilibrium point. At the origin, $x_e = y_e = z_e = 0$, which much simplifies the linear system, whose Jacobian becomes

$$\begin{bmatrix} -\sigma & \sigma & 0 \\ \rho & -1 & 0 \\ 0 & 0 & -\beta \end{bmatrix}.$$

The characteristic equation is easily seen to be

$$(-\beta - \lambda)\begin{vmatrix} -\sigma - \lambda & \sigma \\ \rho & -1 - \lambda \end{vmatrix} = 0.$$

One of the characteristic roots is clearly given by $\lambda = -\beta$, which is negative, while the remaining two are given by the equation

$$\lambda^2 + (1 + \sigma)\lambda + \sigma(1 - \rho) = 0.$$

When $\rho < 1$, the succession of the signs of the coefficients in this quadratic equation is $+++$, hence (see Sect. 14.1.4) both its roots will be of stable type. This shows that the origin is locally stable when $\rho < 1$.

26.3. By using the results of the previous exercise, we see that the succession of the signs of the coefficients in the quadratic equation is $++-$ when $\rho > 1$, hence this equation will have one positive and one negative real root. Thus the characteristic equation has one positive and two negative real roots, hence the origin is a saddle point with a one-dimensional unstable manifold and a two-dimensional stable manifold (see Chap. 18, Sect. 18.2.2.3).

26.4. Let us first determine the two non-trivial equilibrium points., which arise when $\rho > 1$. If we let $x' = y' = z' = 0$ in system (26.24) we have

$$\begin{aligned} 0 &= \sigma(-x+y), \\ 0 &= \rho x - y - zx, \\ 0 &= -\beta z + xy. \end{aligned}$$

From the first equation we have $x = y$ and consequently $z = (1/\beta)x^2$ from the third. Substitution into the second gives

$$(1/\beta)x^3 + (1-\rho)x = 0.$$

A root is $x = 0$ (the trivial solution) while the remaining two are given by

$$x = \pm\sqrt{\beta(\rho-1)},$$

which are real and different from the trivial solution if, and only if, $\rho > 1$. Thus we have

$$(x_e, y_e, z_e) = (\pm\sqrt{\beta(\rho-1)}, \pm\sqrt{\beta(\rho-1)}, \rho-1).$$

At these points the characteristic equation of the Jacobian of the system is

$$\begin{vmatrix} -\sigma-\lambda & \sigma & 0 \\ \rho - z_e & -1-\lambda & -x_e \\ y_e & x_e & -\beta-\lambda \end{vmatrix} = 0,$$

whence

$$\lambda^3 + (\sigma+\beta+1)\lambda^2 + \beta(\sigma+\rho)\lambda + 2\sigma\beta(\rho-1) = 0.$$

Since we are considering $\rho > 1$, all coefficients are positive, hence there can be no non-negative real root. The crucial stability condition is (16.27) in Chap. 16, Sect. 16.4, which gives

$$(\sigma+\beta+1)\beta(\sigma+\rho) - 2\sigma\beta(\rho-1) > 0,$$

from which

$$(-\sigma+\beta+1)\rho + \sigma(\sigma+\beta+3) > 0.$$

Since $\sigma > 1 + \beta$ by assumption, this yields

$$\rho < \frac{\sigma(\sigma+\beta+3)}{\sigma-\beta-1}.$$

26.5. We know (see Sect. 16.4, p. 221) that, when the last inequality in the previous exercise turns into an equality, namely when

$$\rho = \frac{\sigma(\sigma+\beta+3)}{\sigma-\beta-1},$$

then there is certainly a pair of complex roots with zero real part. The third root must be real. Now, given a real root, say λ_1, and two pure imaginary roots, say $\lambda_{2,3} = \pm i\omega$, and using the fact that the sum of the roots equals minus the coefficient of λ^2 in the characteristic equation, we immediately obtain $\lambda_1 = -(\sigma+\beta+1)$. Since the product of the roots equals minus the constant term, we get

$$-(\sigma+\beta+1)\omega^2 = -2\sigma\beta(\rho-1),$$

from which, substituting the value of ρ and solving for ω, we obtain (after simple manipulations)

$$\omega = \sqrt{2\sigma(\sigma+1)/(\sigma-\beta-1)}.$$

As regards the Hopf bifurcation, condition (H1) of the theorem given in Sect. 25.2.2 is satisfied. To check condition (H2), we have to compute the derivative of the real part of the pair of complex roots (evaluated at the point where this real part becomes zero, namely $(\partial\theta/\partial\rho)_{\theta=0}$). For this we can use sensitivity analysis, in particular Eqs. (18.101) in Sect. 18.2.2.2. The parameter is ρ, hence we have

$$\begin{array}{llll} -\dfrac{\partial\lambda_1}{\partial\rho} & -2\dfrac{\partial\theta}{\partial\rho} & & = 0, \\ & -2(\sigma+\beta+1)\dfrac{\partial\theta}{\partial\rho} & +2\sqrt{\frac{2\sigma(\sigma+1)}{\sigma-\beta-1}}\dfrac{\partial\omega}{\partial\rho} & = \beta, \\ \frac{-2\sigma(\sigma+1)}{\sigma-\beta-1}\dfrac{\partial\lambda_1}{\partial\rho} & & +2(\sigma+\beta+1)\sqrt{\frac{2\sigma(\sigma+1)}{\sigma-\beta-1}}\dfrac{\partial\omega}{\partial\rho} & = 2\sigma\beta. \end{array}$$

Solving this system for $\partial\theta/\partial\rho$ we have

$$\left(\frac{\partial\theta}{\partial\rho}\right)_{\theta=0} = \frac{2\beta(\sigma-\beta-1)}{4(\sigma+\beta+1)^2 + 8\sigma(\sigma+1)/(\sigma-\beta-1)} > 0.$$

Condition (H2) is satisfied, hence we have a Hopf bifurcation.

26.6. When $\rho > \sigma(\sigma+\beta+3)/(\sigma-\beta-1)$, the crucial stability condition (see exercise 26.4) is not satisfied, which means that the characteristic equation will have one negative real root and a pair of complex conjugate roots with positive real part. The same result can be reached by observing that, since $(\partial\theta/\partial\rho)_{\theta=0} > 0$

(see the previous exercise), when ρ increases beyond the critical point, the real part of the complex roots becomes positive.

26.7. Following the hint, to maximize the given utility function in any period t we form the Lagrangian (the time subscript has been dropped for notational simplicity)

$$L = x_1^a x_2^{1-a} + \lambda(1 - p_1 x_1 - p_2 x_2),$$

and apply the first-order conditions

$$\begin{aligned} \partial L/\partial x_1 &= a x_1^{a-1} x_2^{1-a} - \lambda p_1 = 0, \\ \partial L/\partial x_2 &= (1-a) x_1^a x_2^{-a} - \lambda p_2 = 0, \\ \partial L/\partial \lambda &= M - p_1 x_1 - p_2 x_2 = 0. \end{aligned}$$

By simple manipulations we obtain the demand functions

$$x_1^* = aM/p_1, x_2^* = (1-a)M/p_2,$$

where the star denotes the optimal value. We now introduce the variability of the utility weight a

$$a_t = \alpha x_{1t-1}^* x_{2t-1}^*.$$

From this and the results of the optimization procedure we get

$$a_t = \alpha(a_{t-1}M/p_1)(1-a_{t-1})M/p_2) = \frac{M^2}{p_1 p_2}\alpha a_{t-1}(1-a_{t-1}).$$

By an appropriate choice of the units of measurement we can set $p_1 = p_2 = M = 1$ (or, if we prefer, we can introduce a new parameter $\beta \equiv \alpha M^2/p_1 p_2$). Thus we have

$$a_t = \alpha a_{t-1}(1 - a_{t-1}),$$

which is the logistic map (26.1), known to give rise to chaos for $\alpha > 3.5699...$ (see Sect. 26.2.1). Given the one-to-one correspondence between a and the quantities demanded, if a_t is chaotic, x_{1t}^*, x_{2t}^* will also be.

This exercise shows that rational consumer decisions in the traditional micro setting may give rise to chaotic movements when preferences depend on past experience. Another interesting conclusion (Benhabib and Day, 1981), is that, in a comparative dynamics framework (hence without normalizing income to one), if the consumer's income is increased sufficiently for any given set of prices, chaos is bound to appear no matter how small the "experience parameter" α. This erratic behaviour explains, according to the authors (p. 463), "the whimsical behavior of the very rich".

26.8.a) From the given the utility function it is easy to calculate

$$U_0 = e^{k-c_0} = e^{a\{1-[(c_0-w_0)/a]\}}, U_1 = 1.$$

Hence, from $U_0/U_1 = [w_1 - c_1(t+1)]/[c_0(t) - w_0]$ we have

$$e^{a\{1-[(c_0-w_0)/a]\}} = \frac{w_1 - c_1(t+1)}{c_0(t) - w_0}.$$

To reduce this to a difference equation involving solely c_0 we first obtain (from the market-clearing condition)

$$c_1(t) - w_1 = (1+\gamma)[w_0 - c_0(t)],$$

hence, letting $\gamma = 0$, shifting the time subscripts, and rearranging terms,

$$w_1 - c_1(t+1) = c_0(t+1) - w_0.$$

It follows that

$$e^{a\{1-[(c_0-w_0)/a]\}} = \frac{c_0(t+1) - w_0}{c_0(t) - w_0},$$

hence

$$c_0(t+1) - w_0 = e^{a\{1-[(c_0-w_0)/a]\}}[c_0(t) - w_0].$$

26.8.b) Following the hint, we get

$$y_{t+1} = e^{a\{1-[y_t/a]\}} y_t = \alpha y_t e^{-y_t},$$

where $y_t \equiv c_0(t) - w_0$ and $\alpha \equiv e^a$. This map is known to give rise to chaos for $\alpha > 14.675$ (or $a > 2.692$), see Eq. (26.10).

Chapter 27

27.1. At the steady-state equilibrium point we have $r'(t) = 0$ and, of course, $r(t-1) = r(t)$. Thus this point is the same as in the standard neoclassical growth model. Linearizing the dynamic equation at the equilibrium point we have

$$\bar{r}'(t) + n\bar{r}(t) - s(\mathrm{d}f/\mathrm{d}r)\bar{r}(t-1) = 0,$$

where a superscript bar denotes the deviations from equilibrium, and $(\mathrm{d}f/\mathrm{d}r)$ is evaluated at the equilibrium point. The characteristic equation is

$$\lambda + n - s(\mathrm{d}f/\mathrm{d}r)e^{-\lambda} = 0,$$

namely

$$\lambda = -n + s(\mathrm{d}f/\mathrm{d}r)e^{-\lambda},$$

which is in the form (27.19). We can then apply Burger's theorem: observing that $s(\mathrm{d}f/\mathrm{d}r) > -1$, the necessary and sufficient condition (1) requires $-n < -s(\mathrm{d}f/\mathrm{d}r)$, that is

$$n > s(\mathrm{d}f/\mathrm{d}r).$$

Inspection of Fig. 13.2 shows that, at the equilibrium point, $s(\mathrm{d}f/\mathrm{d}r) < n$; hence the stability condition is satisfied.

We know from the general treatment of mixed equations of retarded type that their characteristic equation has at most two real roots, while all the others are complex. Hence the approach to equilibrium will be oscillatory.

27.2. By straightforward substitutions we obtain the basic differential equation of the model

$$p'(t) + gp(t) - g\beta\mu p(t-1) = 0,$$

whose characteristic equation is

$$\lambda + g - g\beta\mu e^{-\lambda} = 0.$$

27.2.a) To avoid the possibility of prices converging to zero there must be one zero real root, for in this case the solution will be of the type

$$p(t) = c_1 + \sum_{r=2}^{\infty} c_r e^{\lambda_r t},$$

where c_1 is the arbitrary constant corresponding to the zero real root. Letting $\lambda = 0$ in the characteristic equation we obtain $\beta\mu = 1$ as the condition for this to occur.

27.2.b) Let us then consider the equation

$$\lambda + g - ge^{-\lambda} = 0.$$

Since $g > 0$, according to general principles this equation will have only one real root which is a simple root; hence there will be no real root other than zero.

27.2.c) The remaining roots will be complex. Letting $\alpha \pm i\beta$ be the typical pair of complex roots and applying the procedure described in Sect. 27.4, we obtain

$$\alpha = -g + ge^{-\alpha} \cos\beta,$$

$$\beta = ge^{-\alpha} \sin\beta.$$

The second equation implies $\beta > 0$, namely $2k\pi < \beta < (2k+1)\pi$, where $k = 0, 1, 2, \ldots$ Therefore, there will be a major cycle for $k = 0$ and (an infinite number of) minor cycles for $k = 1, 2, \ldots$

In any of these intervals we have $0 < \sin\beta < 1$. It is then easy to see that, in these intervals, α must be negative. Let us consider, for example, the interval, $2\pi < \beta < 3\pi$, and let us take an arbitrary value of g, say $g = 1$. Although we do not know β (this would require the numerical solution of the system, which we are not interested in), we know that it is greater than one (between 2π and 3π) and that $0 < \sin\beta < 1$. The second equation then requires $e^{-\alpha} > 1$, which in turn implies $\alpha < 0$. This proves that the oscillations will be damped. For a more general proof see Padoan and Petit (1978, Sect. 3).

27.3. Let us consider (see exercise 27.2.c) the expression $\alpha = -g + ge^{-\alpha} \cos\beta$ or $\alpha + g - ge^{-\alpha} \cos\beta = 0$. This can be considered as an implicit function $\varphi(\alpha, \beta, g) = 0$ in the three variables α, β, g. From the implicit function theorem we have

$$\frac{\partial\beta}{\partial g} = -\frac{\partial\varphi/\partial g}{\partial\varphi/\partial\beta} = -\frac{1 - e^{-\alpha}\cos\beta}{ge^{-\alpha}\sin\beta}.$$

Now, $ge^{-\alpha}\sin\beta = \beta$, and $-(1 - e^{-\alpha}\cos\beta) = \alpha/g$. Thus we have

$$\frac{\partial\beta}{\partial g} = \frac{\alpha}{\beta g}.$$

Since we know that $\beta > 0$ and $\alpha < 0$ (exercise 2.c), it follows that $\partial\beta/\partial g < 0$, which implies that the frequency of the oscillations $(\beta/2\pi)$ decreases as g increases.

Similarly, from the implicit function $\beta - ge^{-\alpha}\sin\beta = 0$, we obtain

$$\frac{\partial\alpha}{\partial g} = \frac{e^{-\alpha}\sin\beta}{ge^{-\alpha}\sin\beta} = \frac{1}{g} > 0,$$

which shows that the amplitude of the oscillations is an increasing function of g.