

# **Introduction to Metabolic Control Analysis**

---

*Herbert M. Sauro  
University of Washington  
Seattle, WA*



Ambrosius Publishing

Copyright ©2013-2018 Herbert M. Sauro. All rights reserved.

First Edition, version 0.99b

Published by Ambrosius Publishing and Future Skill Software

[books.analogmachine.org](http://books.analogmachine.org)

Typeset using L<sup>A</sup>T<sub>E</sub>X 2<sub>&</sub>, TikZ (3.0.1a), PGFPlots (1.15), WinEdt (10)  
and 11pt Math Time Professional 2 Fonts

**Limit of Liability/Disclaimer of Warranty:** While the author has used his best efforts in preparing this book, he makes no representations or warranties with respect to the accuracy or completeness of the contents of this book and specifically disclaim any implied warranties of merchantability or fitness for a particular purpose. The advice and strategies contained herein may not be suitable for your situation. Neither the author nor publisher shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages. No part of this book may be reproduced by any means without written permission of the author.

ISBN 10: 0-9824773-6-8 (paperback)

ISBN 13: 978-0-9824773-6-6 (paperback)

Printed in the United States of America.

Mosaic image modified from Daniel Steger's Tikz image (<http://www.texexample.net/tikz/examples/mosaic-from-pompeii/>)

Front-Cover: Metabolic pathway image from the Interactive Pathways Explorer v3, iPath 3 <https://pathways.embl.de/> and KEGG Pathway Atlas (with kind permission from iPath and KEGG).

The image depicts part of the glycolytic and Krebs cycle. Ref: Ivica Letunic, Takuji Yamada, Minoru Kanehisa, and Peer Bork (2008), iPath: interactive exploration of biochemical pathways and networks. Trends in Biochemical Sciences 33(3), p101-103

**University Disclaimer:** Any views, opinions, data, documentation and other information presented in this book are solely those of the author and do not represent those of the University of Washington.

# ***Contents***

Preface . . . . .	x
<b>I Preliminary Topics</b>	<b>1</b>
<b>1 Stoichiometry</b>	<b>5</b>
1.1 Preliminaries . . . . .	5
1.2 Reaction Stoichiometry . . . . .	7
1.3 Rates of Change . . . . .	8
1.4 Reaction Rates . . . . .	11
1.5 Elementary Kinetics . . . . .	15
1.6 The Equilibrium Constant . . . . .	18
1.7 Reaction Order . . . . .	24
Further Reading . . . . .	25
Exercises . . . . .	25
<b>2 Enzyme Kinetics in a Nutshell</b>	<b>27</b>
2.1 Enzymes . . . . .	27
2.2 Michaelis-Menten Kinetics . . . . .	27
2.3 Multisubstrate Kinetics . . . . .	32
2.4 Cooperativity . . . . .	33
2.5 Allostery . . . . .	36
Further Reading . . . . .	38
<b>3 Stoichiometric Networks</b>	<b>39</b>
3.1 Stoichiometric Networks . . . . .	39
3.2 Standard Visualization Notation . . . . .	46
3.3 Mass-Balance Equations . . . . .	46

3.4 Stoichiometry Matrix . . . . .	53
3.5 Reversibility . . . . .	55
3.6 Signaling Networks . . . . .	57
3.7 Gene Regulatory Networks . . . . .	59
3.8 Moiety Conserved Cycles . . . . .	60
3.9 The System Equation . . . . .	63
3.10 Tellurium . . . . .	64
Further Reading . . . . .	65
Exercises . . . . .	65
<b>4 How Systems Behave</b>	<b>69</b>
4.1 System Behavior . . . . .	69
4.2 Equilibrium . . . . .	69
4.3 Boundary and Floating Species . . . . .	71
4.4 Steady State . . . . .	72
4.5 Transients . . . . .	75
4.6 Setting up a Model in Software . . . . .	76
4.7 Effect of Different Kinds of Perturbations . . . . .	76
4.8 Sensitivity Analysis . . . . .	78
4.9 Robustness and Homeostasis . . . . .	82
Further Reading . . . . .	82
Exercises . . . . .	83
Tellurium Scripts . . . . .	83
<b>5 Traditional Concepts in Metabolic Regulation</b>	<b>89</b>
5.1 Introduction . . . . .	89
5.2 Early Quantitative Efforts . . . . .	90
5.3 Prevailing Ideas . . . . .	90
5.4 A Modern Understanding of Metabolism . . . . .	92
Further Reading . . . . .	92
<b>II Metabolic Control Analysis</b>	<b>93</b>
<b>6 Elasticities</b>	<b>97</b>

6.1	Introduction . . . . .	97
6.2	Elasticity Coefficients . . . . .	100
6.3	Mass-action Kinetics . . . . .	108
6.4	Enzyme Catalyzed Reactions . . . . .	111
6.5	Cooperativity . . . . .	112
6.6	Local Equations . . . . .	114
6.7	General Elasticity Rules . . . . .	116
6.8	Summary . . . . .	119
	Further Reading . . . . .	121
	Exercises . . . . .	121
<b>7</b>	<b>Introduction to Biochemical Control</b>	<b>123</b>
7.1	What do we mean by Control? . . . . .	123
7.2	Control Coefficients . . . . .	124
7.3	Distribution of Control . . . . .	128
7.4	Predicting Flux and Concentration Changes . . . . .	131
7.5	Summation Theorems . . . . .	133
	Further Reading . . . . .	144
	Exercises . . . . .	144
	Appendix 7.A Python/Tellurium Scripts . . . . .	145
<b>8</b>	<b>Linking the Parts to the Whole</b>	<b>149</b>
8.1	Control Coefficients in Terms of Elasticities . . . . .	149
8.2	Connectivity Theorems . . . . .	151
8.3	Response Coefficients . . . . .	158
8.4	Canonical Control Coefficients . . . . .	160
8.5	How to Derive Control Equations . . . . .	161
8.6	Relationship to S-Systems . . . . .	167
	Further Reading . . . . .	170
	Exercises . . . . .	170
	Appendix 8.A Python/Tellurium Scripts . . . . .	171
<b>9</b>	<b>Experimental Methods</b>	<b>173</b>
9.1	Introduction . . . . .	173

9.2 Using Classical Genetics . . . . .	174
9.3 Genetic Engineering . . . . .	174
9.4 Titration by Inhibitors . . . . .	176
9.5 Double Modulation Technique . . . . .	180
9.6 Reconstituted Systems . . . . .	182
9.7 By Computer Simulation . . . . .	183
9.8 By Calculation - Serine Pathway . . . . .	184
9.9 Leveraging CRISPR Technology . . . . .	186
Further Reading . . . . .	187
Exercises . . . . .	187
<b>10 Linear Pathways</b>	<b>189</b>
10.1 Basic Properties . . . . .	189
10.2 Product Insensitive Steps and Fast Reactions . . . . .	192
10.3 Saturable Enzyme Kinetics . . . . .	196
10.4 Front Loading . . . . .	201
10.5 Optimal Allocation of Protein . . . . .	206
Further Reading . . . . .	207
Exercises . . . . .	208
Appendix 10.A Appendix . . . . .	209
<b>11 Branched and Cyclic Systems</b>	<b>211</b>
11.1 Branched Pathways . . . . .	211
11.2 Futile or Substrate Cycles . . . . .	218
Exercises . . . . .	221
Tellurium Scripts . . . . .	222
<b>12 Negative Feedback</b>	<b>225</b>
12.1 Historical Background . . . . .	225
12.2 Simple Quantitative Analysis . . . . .	229
12.3 Negative Feedback in Biochemical Systems . . . . .	233
12.4 The PFK Paradox . . . . .	238
12.5 Robustness and Supply/Demand . . . . .	244
12.6 Instability . . . . .	245

Further Reading . . . . .	245
Exercises . . . . .	245
<b>13 Stability</b>	<b>247</b>
13.1 Jacobian for Biochemical Systems . . . . .	253
13.2 External Stability . . . . .	255
13.3 Phase Portraits . . . . .	256
13.4 Bifurcation Plots . . . . .	262
Further Reading . . . . .	272
Exercises . . . . .	272
13.5 Appendix . . . . .	275
<b>14 Stability of Negative Feedback Systems</b>	<b>279</b>
14.1 Introduction . . . . .	279
14.2 Stability using MCA . . . . .	280
14.3 Effect of Pathway Length . . . . .	286
Further Reading . . . . .	290
Exercises . . . . .	290
<b>15 Moiety Conservation Laws</b>	<b>291</b>
15.1 Moiety Constraints . . . . .	291
15.2 Moiety Conserved Cycles . . . . .	294
15.3 Basic Theory . . . . .	296
15.4 Computational Approaches . . . . .	300
15.5 Additional Theory – Advanced . . . . .	311
15.6 Numerical Methods . . . . .	315
15.7 Design of Simulation Software . . . . .	322
Further Reading . . . . .	325
Exercises . . . . .	325
<b>16 Moiety Conserved Cycles</b>	<b>327</b>
16.1 Moiety Conserved Cycles . . . . .	327
16.2 MCA of Conserved Cycles . . . . .	329
16.3 Using MCA to Understand Ultrasensitivity . . . . .	333
16.4 Cycle Connectivity Theorems . . . . .	338

16.5 Sequestration . . . . .	343
16.6 Cascades . . . . .	349
Further Reading . . . . .	356
Exercises . . . . .	356
Appendix 16.A Python/Tellurium Scripts . . . . .	356
<b>17 Flux Constraints</b>	<b>359</b>
17.1 Flux Constraints . . . . .	359
17.2 Flux Balance Laws . . . . .	359
17.3 Determined Systems . . . . .	362
Further Reading . . . . .	370
<b>18 Steady State Flux Patterns</b>	<b>371</b>
18.1 The Null Space . . . . .	371
18.2 Elementary Modes . . . . .	375
18.3 Definition of a Pathway . . . . .	380
18.4 Maximum Yield Predictions . . . . .	381
18.5 Pathway Engineering . . . . .	386
Further Reading . . . . .	386
<b>III Appendices</b>	<b>389</b>
<b>A List of Symbols and Abbreviations</b>	<b>391</b>
<b>B Math Fundamentals</b>	<b>395</b>
B.1 Notation . . . . .	395
B.2 Short Table of Derivatives . . . . .	396
B.3 Logarithms . . . . .	396
B.4 Partial Derivatives . . . . .	396
B.5 Differential Equations . . . . .	398
B.6 Polynomials . . . . .	399
B.7 Taylor Series . . . . .	401
B.8 Total Derivative . . . . .	403
B.9 Eigenvalues and Eigenvectors . . . . .	403
Further Reading . . . . .	407

<b>C Control Equations</b>	<b>409</b>
C.1 Linear Pathways . . . . .	409
C.2 Cycles . . . . .	412
C.3 Branches . . . . .	412
<b>D Modeling with Python</b>	<b>415</b>
D.1 Introduction . . . . .	415
D.2 Introduction to Python . . . . .	416
D.3 Describing Reaction Networks using Antimony . . . . .	420
D.3.1 Initialization of Model Values . . . . .	423
D.4 Using libRoadRunner in Python . . . . .	423
D.4.1 Time Course Simulation . . . . .	424
D.4.2 Plotting Simulation Results . . . . .	425
D.4.3 Applying Perturbations to a Simulation . . . . .	426
D.4.4 Steady State and Metabolic Control . . . . .	427
D.4.5 Other Model Properties of Interest . . . . .	429
D.5 Generating SBML and Matlab Files . . . . .	429
D.6 Exercise . . . . .	430
<b>References</b>	<b>433</b>
<b>History</b>	<b>443</b>
<b>Index</b>	<b>445</b>

## Preface

---

This book is an introduction to control in biochemical pathways. The book should be suitable for undergraduates in their early (Junior, USA, second year UK) to mid years at college. The book can also serve as a reference guide for researchers and teachers.

The latest edition together with free software and other material can be found at [www.books.analogmachine.org](http://www.books.analogmachine.org), [tellurium.analogmachine.org](http://tellurium.analogmachine.org) and the research site, [www.sys-bio.org](http://www.sys-bio.org).

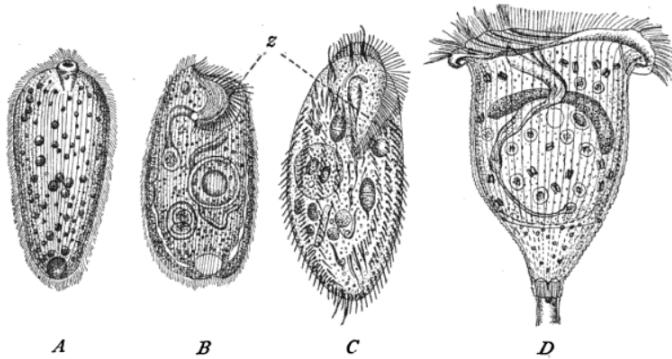
This is a book I've wanted to write for many years. It covers the basics of metabolic control analysis, that is the framework developed originally by Heinrich, Kacser, Burns and Savageau that helps us think about how biochemical networks operate. There are however a number of topics missing from this edition. These include control in complex branched systems, a detailed look at the effect of sequestration, including metabolic channeling, and hierarchical control analysis that include genetic regulation as part of the analysis. These topics will be included in subsequent editions. This edition also lack examples taken from the research literature, this again will be remedied in a subsequent edition.

As with my earlier text books I have decided to publish this book myself via a service called Createspace that is part of Amazon. Over the years I've considered publishing text books via bonafide publishers but have found the contracts they offer to be far too restrictive. Two restrictions in particular stand out, the loss of copyright on the text as well as any figures and the inability to rapidly update the text when either errors are found or new material needs to be added.

There are many people and organizations who I should thank but foremost must be my infinitely patient wife, Holly, who has put up with the many hours I have spent working alone in our basement or long hours at the department and who contributed significantly to editing this book. I am also most grateful to the National Science Foundation and the National Institutes of Health who paid my summer salary so that I could allocate the time to write, edit and research. I would also like to thank the many undergraduates, graduates and colleagues who have directly or indirectly contributed to this work. In particular I want to thank my two teachers, David Fell and the late Henrik Kacser who I had the privilege to work with as a graduate student and postdoctoral fellow respectively. I had many hours of fruitful and stimulating conversations with Jannie Hofmeyr, Athel Cornish-Bowden, Michael Savageau, Reinhart Heinrich, Frank Bruggeman, Bas Tugesink, Hans Westerhoff, Boris Kholodenko, Stefan Schuster, Jim Burns and last but not least Luis Azerenza. In early 2000s I had the good fortune to be introduced to control theory by Brian Ingalls, John Doyle and Mustafa Khammash, all three of which had a significant influence on my understanding of control theory and had that rare knack of explaining complex ideas. More recently I should thank my graduate students, in particular Frank Bergmann (author of SBW) who was (and is) a brilliant programmer, and Deepak Chandran (author of TinkerCell) who developed a very deep understanding of how networks operate. I thank them for their dedication and steadfast enthusiasm while they worked in my lab.

I wish to sincerely thank the authors of the  $\text{\TeX}$  system, MikTeX (2.9), TikZ (3.0.1a), PGFPLOTS (1.15) and WinEdt (9.0) for making available such amazing tools to technical authors.

Finally, I should thank Michael Corral (<http://www.mecmath.net/>) and Mike Hucka ([sbml.org](http://sbml.org)) whose  $\text{\LaTeX}$  work inspired some of the styles I used in the text.



**Figure 1** Some of my favourite animals. Types of Ciliates: From The Protozoa, Gary Nathan Calkins, Macmillan, 1910.

July 2018  
Seattle, WA

Herbert M. Sauro

**Other topics to add:**

1. Competitive inhibitor and MCA
2. Controlled Comparison
3. Limitations of MCA/BST
4. ATP/ADP/AMP see Interface Perspective for Moiety cycle chapter
5. Solutions to linear odes in math appendix
6. Use Tellurium for bifurcation plots
7. Integral control of cycle?
8. Adair model for multisite phos?
9. More on supply and demand

# **Part I**

# **Preliminary Topics**



## Preface to Part I

---

The first few chapters provides a foundation for part II of the book which is where we start to discuss metabolic control analysis. For those already familiar with many of these topics Part I can be omitted. However, it does emphasizes some important concepts which even seasoned practitioners might be unaware of.

### **Notation:**

Following other authors, an upper case italic letter such as  $S$  indicates the name of a molecular species while the lower case italic variant,  $s$ , indicates the concentration of species  $S$ . This is to avoid the use of square brackets, as in  $[S]$  which can result in unnecessary clutter in equations.

Bold text will tend to be used either for emphasis or when introducing new terminology.



# 1

## *Stoichiometry*

The word stoichiometry originates from the Greek words ‘element’ and ‘measure’, and literally means the measurement of the elements. The word is commonly used with reference to calculations involved in the composition of a chemical system. Thus, stoichiometry could refer to the composition of atoms in a molecule, or the composition of subunits in a protein. More commonly, it’s used to represent the proportions of reactants and products in a chemical reaction.

Stoichiometry deals with static information about the amounts of substances involved in a chemical transformation, whereas kinetics refers to the rates of change that occur in these amounts. To paraphrase a statement made by Aris [3], stoichiometry provides the framework within which chemical change takes place irrespective of the forces that bring them about, whereas kinetics addresses the speed of chemical change. Aris then went on to state, “Just as the latter can only be built on a proper understanding of the kinematics, so the analysis of stoichiometry must precede that of kinetics”. We will do the same here.

### 1.1 Preliminaries

---

#### The Mole

The mole is the fundamental SI unit of amount and refers to the number of constituent entities in a substance. Such entities may include atoms, molecules, ions etc. The actual number of entities in one mole is given by Avogadro’s number,  $6.022 \times 10^{23}$  with units  $\text{mol}^{-1}$ . Living cells are known to contain all manner of different molecules, from small ions such as  $\text{Fe}^{2+}$ , to macromolecules such as DNA. A typical *E. coli* cell has roughly five

to seven thousand different kinds of molecules, most of these being proteins (about four thousand). Some molecules such as  $\text{Fe}^{2+}$  are present in huge numbers at approximately seven million ions per cell. Other molecules such as the LacI repressor is present in very small quantities, perhaps no more than ten copies per cell [90, 64].

Examples 1.1 and 1.2 illustrate two calculations that estimate the number of molecules in biological cells. Small molecules such as ATP that are involved in mainstream metabolism tend to be present in large numbers; in the case of ATP, a few million. In contrast, other molecules such as some proteins, can be present in very small numbers. In the case of CheB, it's roughly 250 molecules.

### Example 1.1

How many molecules of Adenosine Triphosphate (ATP) are present in a single *E. coli* cell if the concentration of ATP is 2.5 mM? Assume that the average *E. coli* is a cylinder of length  $2\mu\text{m}$  and diameter of  $1\mu\text{m}$ .

Assuming that an *E. coli* cell can be approximated by a simple cylinder, the volume of an *E. coli* can be computed from the standard formula for a cylinder:  $\pi r^2 h$ , where  $r$  is the radius and  $h$  the length. That is, the volume of a single *E. coli* cell is  $\pi \times (0.25 \times 10^{-12}) \times (2 \times 10^{-6}) = 1.57 \times 10^{-18} \text{ m}^3$ .

The concentration of ATP is 2.5 mM which represents  $2.5 \times 10^{-3}$  moles per liter or 2.5 moles per  $\text{m}^3$  ( $1 \text{ m}^3 = 1000 \text{ L}$ ).

The number of moles of ATP in one *E. coli* volume is then  $1.57 \times 10^{-18} \times 2.5 = 3.93 \times 10^{-18}$  moles. Given Avogadro's number of  $6.022 \times 10^{23}$ , the number of ATP molecules must then equal  $(3.93 \times 10^{-18}) \times (6.022 \times 10^{23}) = 2,366,646$  molecules.

Consider for example the chemotactic circuit in *E. coli*. This circuit is used to detect and process information on external nutrient and repellent gradients. Without going into great detail, the *E. coli* chemotactic circuit samples external concentrations over time. If the circuit records an increase in nutrient concentration, the cell flagellum motor is modulated so that there is an increased likelihood of moving in a straight line. Otherwise, it increases its random tumbling frequency so the cell can reorientate to a potentially more favorable direction [6, 133]. The circuit that accomplishes this computation is composed of at least seven protein components. Table 1.1 shows the number of molecules of each type and highlights the wide range in the number of protein molecules, ranging from 140 for CheR to over 15,000 for the receptor molecules themselves.

### Example 1.2

The enzyme CheB (a methylesterase) is involved in the chemotaxis circuit of *E. coli*. The enzyme is present in *E. coli* at a concentration of about  $0.28\mu\text{m}$ . How many copies of the enzyme are present in a single *E. coli* cell?

Assuming that the volume of a single *E. coli* is the same as previously computed, we can estimate the

Protein	Protein Description	Number of Molecules
Tsr + Tar	Total Receptors	15,000
CheA	Autophosphorylase	6700
CheW	Adaptor Protein	6700
CheY	Motor Transducer	8200
CheZ	Phosphatase of CheY	3200
CheB	Receptor Demethylase	240
CheR	Receptor Methylase	140

**Table 1.1** Table of proteins involved in the *E. coli* chemotaxis detection circuit and the number of proteins of each type present in the cell. There is almost two orders of magnitude difference between the most common and rarest protein. Adapted from <http://www.pdn.cam.ac.uk/groups/comp-cell/Rates.html>

number of CheB molecules by simple proportion. If a concentration of 2.5 mM equates to 2,366,646 molecules of ATP, then by proportion,  $0.28\mu\text{m}$  of CheB equates to  $2,366,646/(2.5 \times 10^3/0.28)$  or approximately **262** molecules per *E. coli*.

## 1.2 Reaction Stoichiometry

One of the most fundamental processes in living organisms is the chemical reaction. This is where molecules combine, decompose, change configuration or exchange molecular sub-units. It is the variety and sheer number of reactions that gives rise to the richness of living systems. A chemical reaction is usually depicted in the form of a chemical equation which describes the transformation of one or more reactants into one or more products. The reactants appear on the left of the equation and the products on the right. Both sides are separated by an arrow indicating the positive direction of the transformation. For example, the equation:



means that two molecules of ADP are transformed into one molecule of ATP and one molecule of AMP with the positive rate from left to right. A double arrow is sometimes used to indicate that a reaction is reversible, as in:



although in general all reactions are in principle reversible to a greater or lesser extent.

We now need to define a number of terms: the stoichiometric amount, rate of change, stoichiometric coefficient, and reaction rate.

## Stoichiometric Amount

The **stoichiometric amount** is defined as the number of molecules of a particular reactant or product taking part in a reaction. Stoichiometric amounts will **always** be **positive** numbers. For example, in the reaction:



ADP has a stoichiometric amount of two, ATP a stoichiometric amount of one, and AMP also with a stoichiometric amount of one. If the same species occurs on the reactant and product side of a reaction, then it must be treated separately. For example, in the reaction:



The stoichiometric amounts on the reactant side include: A with two, B with one and C with one. On the product side the stoichiometric amounts include: A with three, D with one and B with two.

The **stoichiometric amount** is the number of molecules of a particular reactant or product taking part in a reaction.

**Notation Reminder:** Molecular species will be represented using upper case Roman letters, e.g.  $A$ , concentrations by lower case italic Roman letters, e.g.  $a$ , and stoichiometric amounts using the lower case italic  $s$  with an appropriate subscript, e.g.  $s_A$  is the stoichiometric amount for species  $A$ . Stoichiometric coefficients will be represented using  $c$  with a suitable subscript, e.g.  $c_i, c_a$ .

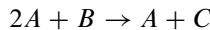
On rare occasions the traditional use of square brackets will also be used to represent concentration, e.g.  $[A]$ .

---

### Example 1.3

---

List the stoichiometric amounts in the following reaction:



On the reactant side the stoichiometric amount for  $A$  is two and for  $B$  is one. On the product side, the stoichiometric amount for  $A$  is one and for  $C$ , one.

---

---

## 1.3 Rates of Change

---

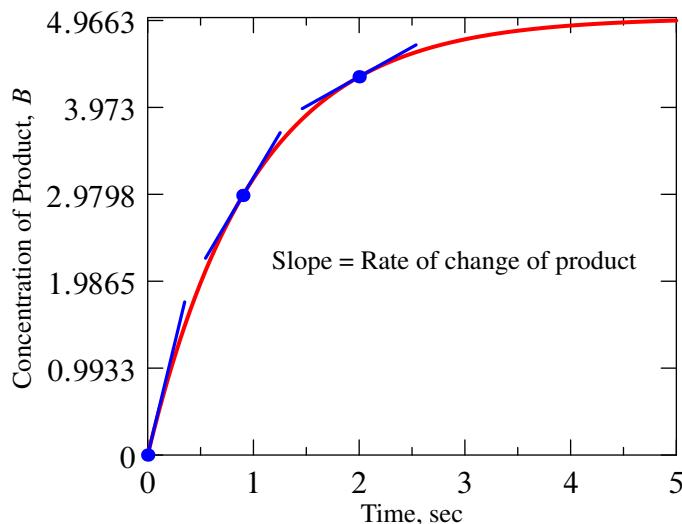
The rate of change can be defined as the rate of change in concentration or amount (depending on units) of a designated species. If  $A$  is the species, then the rate of change is given by:

$$\text{Rate} = \frac{\Delta a}{\Delta t}$$

where the lower case  $a$  refers to the concentration of species  $A$ . Because rates change as reactants are consumed and products made, the rate of change is better defined as the instantaneous change in concentration, or a derivative:

$$\text{Rate} = \frac{da}{dt}$$

If the rate of product formation,  $b$ , is plotted as a function of time, the rate of reaction would be given by the slope of the curve (Figure 1.1). If concentrations are measured in moles per liter (L) and time in seconds (sec), then the rate of reaction is expressed in  $\text{mol L}^{-1} \text{ sec}^{-1}$ .



**Figure 1.1** Progress curve for a simple irreversible reaction,  $A \rightarrow B$ . Initial reactant concentration,  $a$ , is set at 5 units. The plot shows the accumulation of product,  $B$ , as the reaction proceeds. The rate of change of product is given by the slope of the curve which changes over time.

When reporting a rate of change, it is important to give the name of the species that was used to make the measurement. For example, in the reaction  $2A \rightarrow B$ , the rate of change of  $a$  is twice the rate of change of  $b$ . In addition, the rate of change of  $a$  is negative because it is consumed, whereas the rate of change of  $b$  is positive because it is being made.

### Stoichiometric Coefficients

The stoichiometry coefficient refers to the **relative** amount of substance that is consumed and/or produced by a reaction. Given a reaction such as:

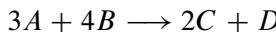


the stoichiometric amount of *A* is 2 and for *B*, 1. The species stoichiometry or **stoichiometric coefficient** however, is the difference between the stoichiometric amounts of a given species on the product side and the stoichiometric amount of the same species on the reactant side. The definition below summarizes this more clearly.

The **stoichiometric coefficient**,  $c_i$ , for a molecular species  $A_i$ , is the difference between the stoichiometric amount of the species on the product side and the stoichiometric amount of the same species on the reactant side, that is:

$$c_i = \text{Stoichiometric Amount of Product, } A_i \\ - \text{Stoichiometric Amount of Reactant, } A_i$$

In the reaction,  $2A \rightarrow B$ , the stoichiometric amount of *A* on the product side is **zero**, while on the reactant size it is two. Therefore the stoichiometric coefficient of *A* is given by  $0 - 2 = -2$ . In many cases a particular species will only occur on the reactant or product side, and it is relatively uncommon to find situations where a species occurs simultaneously as a product and a reactant. As a result, reactant stoichiometric coefficients tend to be **negative** and product stoichiometric coefficients **positive**. To illustrate this further consider the more complex reaction:



Since *A* only appears on the reactant side, its stoichiometric coefficient will be  $-3$ , similarly *B* will have a stoichiometric coefficient of  $-4$ . Species *C* only occurs on the product side, therefore its stoichiometric coefficient is  $+2$ , and similarly *D* will have a stoichiometric coefficient of  $+1$ . In these cases the stoichiometric amounts and the stoichiometric coefficients are the same except for the sign difference on the reactant stoichiometric coefficients.

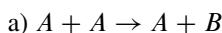
Finally, consider the following reaction where a species occurs on both the reactant and product side:



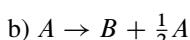
The stoichiometric coefficient of *A* must take into account the fact that *A* appears both as a reactant and a product. The overall stoichiometric coefficient of *A* is therefore  $+1 - 2$  which gives  $-1$ .

#### **Example 1.4**

Write the stoichiometric coefficients for the following reactions:



The stoichiometric amount of *A* on the reactant side is 2 and on the product side, 1. Therefore the stoichiometric coefficient for *A* is  $1 - 2 = -1$ . The stoichiometric amount of *B* on the product side is 1 and on the reactant side, 0, therefore the stoichiometric coefficient for *B* is  $1 - 0 = 1$ .



The stoichiometric amount of  $A$  on the reactant side is 1 and on the product side  $1/2$ , therefore the stoichiometric coefficient for  $A$  is  $1/2 - 1 = -1/2$ . The stoichiometric amount of  $B$  on the reactant side is 0 and on the product side, 1, therefore the stoichiometric coefficient for  $B$  is  $1 - 0 = 1$ .

---

The last example highlights how information can be lost when computing stoichiometric coefficients. It is not possible to recreate the original reaction equation from the stoichiometric coefficients alone, and this underscores the danger of just supplying stoichiometric coefficients when communicating information on reaction equations to other researchers. One option is to store the stoichiometric amounts together with the associated reactant or product. Computer exchange formats, such as the Systems Biology Markup Language (SBML) [59] are specifically designed to preserve complete reaction equation information for this very reason.

Example 1.4 (b) highlights another fact about stoichiometric coefficients. The coefficients can be fractional amounts, often represented as rational fractions.

## 1.4 Reaction Rates

---

In the majority of chemistry and biochemistry texts, it is implicitly assumed that for a given reaction, the number of molecules undergoing transformation in a unit time is very large. Millimolar quantities of ATP for example in an *E. coli* cell represents a million or more molecules and a typical rate of ATP production by pyruvate kinase under anaerobic conditions is approximately 700,000 ATP molecules per second per cell [32]. On the other hand a protein such as the LacI repressor in *E. coli* may only be present in numbers less than ten [90, 64] copies per cell. The huge range in molecule numbers has implications for the way we describe kinetics mathematically.

On the one hand, when dealing with a million molecules, it is reasonable to treat the concentration of that molecule as a continuous real value. For example with the concentration of ATP at 2.5 mM [2], ATP can be regarded as a continuous variable. However, when the number of molecules gets to a point where it is possible to count the number on two hands, the idea of using a continuous real value becomes unrealistic. Instead, it is better to actually quote the number of molecules. Mathematically, the representation of concentration should either be continuous when dealing with large numbers of molecules, or discrete in the case of small numbers of molecules. In addition, a further problem emerges. The molecular model upon which much of chemistry is based is a statistical one that assumes random collisions of molecules and random reaction events.

Again, when dealing with large numbers of molecules and a corresponding large number of events, the assumption is that such random events at the molecular level are averaged out. It is therefore a reasonable assumption to use a deterministic, non-statistical model to describe the reaction rates. However when dealing with situations such as the LacI repressor, which is present in only a handful of copies, the stochastic nature of the molecular world becomes a dominant feature, meaning that the deterministic model is not a reasonable model.

In either case however, be it deterministic, stochastic, discrete or continuous, the rules of stoichiometry will **always** apply because the conservation of mass is not affected by these details.

In the next few sections a brief discussion of deterministic reaction rates will be given and its relationship to stoichiometry. A discussion of stochastic reaction rates is however beyond the scope of this book, but interested readers can refer to the book by Wilkinson [151] and the excellent review by Gillespie [44] and the companion book on kinetics [109] which gives both an intuitive and formal discussion of stochastic kinetics.

## Deterministic Kinetics

The deterministic **rate of reaction** can be defined in terms of the **rate of change** in concentration of a designated species, for example a product,  $P$ . An obvious way to estimate the rate of change is to measure the concentration of  $P$  at two different times, say  $t_1$  and  $t_2$  during the progress of a reaction. The rate of change in  $P$  is then given by:

$$\text{Rate of Change} = \frac{p_2 - p_1}{t_2 - t_1} = \frac{\Delta p}{\Delta t}$$

where  $\Delta p$  is the change in molar concentration of the species  $P$  during the time interval,  $\Delta t$ . However, because most reaction rates show nonlinear behavior with respect to reactant and product concentrations, the rate changes depending on the size of the time interval,  $\Delta t$ . This problem can be easily remedied by defining the rate of change as the instantaneous change in concentration.

$$\text{Rate of Change} = \frac{dp}{dt} \quad (1.1)$$

If the rising concentration of product,  $p$ , is plotted as a function of time, the rate of reaction is given by the slope of the curve as shown in Figure 1.1. If concentration is measured in moles per liter ( $L$ ) and time in seconds (sec), then the rate of change is expressed as  $\text{mol } L^{-1} \text{ sec}^{-1}$ .

Is the rate of change the same as the rate of reaction? Sometimes yes, but not always. To illustrate this consider the reaction,  $2A \rightarrow B$ , where twice as much  $A$  disappears compared to how much  $B$  is formed. If the rate of change in  $a$  and  $b$  are measured at the same time, two different rates of change will be recorded, one twice ( $da/dt$ ) the size of the other ( $db/dt$ ), that is:

$$-2 \frac{da}{dt} = \frac{db}{dt}$$

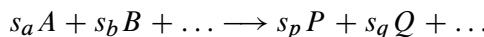
Not only that, the signs will also be different,  $da/dt$  is negative while  $db/dt$  is positive. The difference between  $da/dt$  and  $db/dt$  can be reconciled through stoichiometric coefficients.

To account for differences in the measured rates of change between different species, the rates of change are normalized by the stoichiometric coefficients. For the simple reaction  $2 A \rightarrow$

$B$ , the stoichiometric coefficient for  $A$  is  $-2$  and  $B$ ,  $+1$ . To obtain a species independent reaction rate,  $v$ , the rate of change is divided by the corresponding stoichiometric coefficients, thus:

$$v = \frac{1}{-2} \frac{da}{dt} = \frac{1}{+1} \frac{db}{dt}$$

Note that the  $-2$  and  $+1$  are the stoichiometric coefficients for  $A$  and  $B$ , respectively. For a reaction of the form:



where  $s_a, s_b, \dots$  and  $s_p, s_q, \dots$  represent the stoichiometric amounts, the reaction rate is given by:

$$\text{Rate} = v \equiv \frac{1}{c_a} \frac{da}{dt} = \frac{1}{c_b} \frac{db}{dt} \dots = \frac{1}{c_p} \frac{dp}{dt} = \frac{1}{c_q} \frac{dq}{dt} \dots \quad (1.2)$$

where  $c_i$  is the stoichiometric coefficient. Defined this way, a reaction rate is independent of the species used to measure it. The same applies if a given species appears on both sides of a reaction. For example, in the reaction  $A \rightarrow 2A$ , the stoichiometric coefficient is  $+1$  and the reaction rate,  $v$ , is:

$$v = \frac{1}{+1} \frac{da}{dt}$$

To make the definition of the reaction rate more formal, let us introduce the **extent of reaction**, indicated by the symbol,  $\xi$ . For a given species  $a_i$ , we define a change from  $\xi$  to  $\xi + d\xi$  in time  $dt$  to mean that  $c_i d\xi$  moles of  $a_i$  react. By this definition we can state that for a molecular species  $i$ , the following is true for the time interval,  $dt$ :

$$da_i = c_i d\xi \quad (1.3)$$

or

$$\frac{da_i}{dt} = c_i \frac{d\xi}{dt}$$

From this relation we **define** the **extensive rate of reaction**,  $v_E$ , to be:

$$v_E \equiv \frac{d\xi}{dt}$$

In other words:

$$\frac{da_i}{dt} = c_i v_E \quad (1.4)$$

For the moment we will use  $v_E$  and  $v_I$  to distinguish the extensive and intensive reaction rates. Note that  $\xi$  has units of **amount** and  $v_E$  has units of **amount per unit time** and is therefore an **extensive property**, being dependent on the size of the system. The advantage of introducing the extent of reaction is that it allows us to formally define the rate of reaction independently of the species we use to measure the rate. This convenient property can be expressed as:

$$v_E \equiv \frac{d\xi}{dt} = \frac{1}{c_1} \frac{da_i}{dt}$$

**Example 1.5**

Express the rate of reaction and the rates of change for the following biochemical reaction:



. The rate of reaction is given by:

$$\begin{aligned} v &= \frac{d\xi}{dt} = \frac{d(\text{atp})}{dt} = \frac{d(\text{amp})}{dt} \\ &= -\frac{1}{2} \frac{d(\text{adp})}{dt} \end{aligned}$$

If the volume,  $V$ , of the system is constant, we can also express the rate in terms of concentration, for example  $a_i = A_i / V$  where  $A_i$  is the amount of species  $A$ .

We can therefore rewrite the rate of reaction in the form:

$$\frac{v_E}{V} = \frac{1}{c_1} \frac{da_1}{dt} = \dots$$

where  $v_E$  has units of amount per unit time ( $\text{mol s}^{-1}$ ).

**Definitions:**

*Intensive Property:* A property that does not depend on the quantity of substance. Examples include temperature, density and concentration.

*Extensive Property:* A property that does depend on the quantity of substance. Examples include mass and volume.

The relation  $v_E / V$  is the intensive version of the rate,  $v_I$ , with units of concentration per unit time ( $\text{mol L}^{-1} \text{s}^{-1}$ ) and is the most commonly used form in biochemistry.

$$v_I = \frac{v_E}{V} = \frac{1}{c_i} \frac{da_i}{dt}$$

or:

$$\frac{da_i}{dt} = c_i v_I \quad (1.5)$$

where  $a_i$  is the concentration of species  $i$  and  $v_I$  is the **intensive rate of reaction**. For constant volume, single compartment systems, this is a commonly encountered equation in models of cellular networks. The above equation may also be expressed as:

$$\frac{1}{V} \frac{dA_i}{dt} = c_i v_I \quad (1.6)$$

where  $A_i$  is the amount of species  $A$  to emphasize the change in mass that accompanies a reaction. Recall that  $v_I$  is expressed as  $\text{mol L}^{-1} \text{s}^{-1}$ . If an  $E$  or  $I$  subscript is not used on  $v$ , then the specific form should be clear from the context. In this book, where we use  $v$ , we will generally mean  $v_I$ , the intensive form.

In some simulation situations, for example those involving multiple compartments of different volumes or where there are specific mass conservation laws at work, the intensive rate is not appropriate. This is because the intensive version is unable to keep track of the total number of moles undergoing transformation. In these situations it is necessary to deal explicitly with the extensive rate of reaction, in other words:

$$\frac{dn_i}{dt} = V c_i v_I$$

In most reaction network software simulators, internally the rates are represented using the extensive rate of change in order to keep track of change in the total number of moles of substances.

## 1.5 Elementary Kinetics

---

Chemical reactions that involve no reaction intermediates are called **elementary reactions**. Such reactions often have simple kinetic properties and empirical studies have shown that the rate of reaction is often proportional to the product of the molar concentration of the reactants raised to some power.

For a simple monomolecular reaction such as:



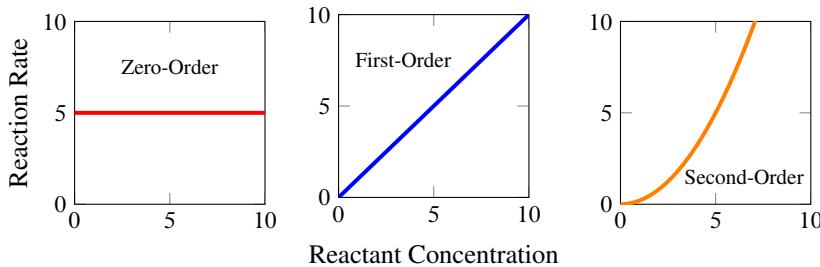
the rate of reaction,  $v$ , is often found to be proportional to the concentration of species A, or:

$$v = ka$$

The proportionality constant,  $k$ , is called the **rate constant**.  $a$  is the concentration of reactant and  $v$  is the rate of reaction. Recall that the rate of change of  $A$  is the reaction rate times the stoichiometry coefficient (Equation (1.5)), since the stoichiometry coefficient of  $A$  is  $-1$ , the rate of change is given by:

$$\frac{da}{dt} = -ka \quad (1.7)$$

If the units for  $k$  are  $\text{t}^{-1}$  and for the concentration of  $A$ , moles  $\text{L}^{-1}$ , then the rate of change of  $A$  has units of moles  $\text{L}^{-1}\text{t}^{-1}$ . By convention, a rate that is proportional to a reactant raised to a first power is called **first-order** with respect to the reactant. Similarly, reactants raised to the zeroth power are called **zero-order**, and reactants raised to the power of two are called **second-order** (see Figure 1.2).

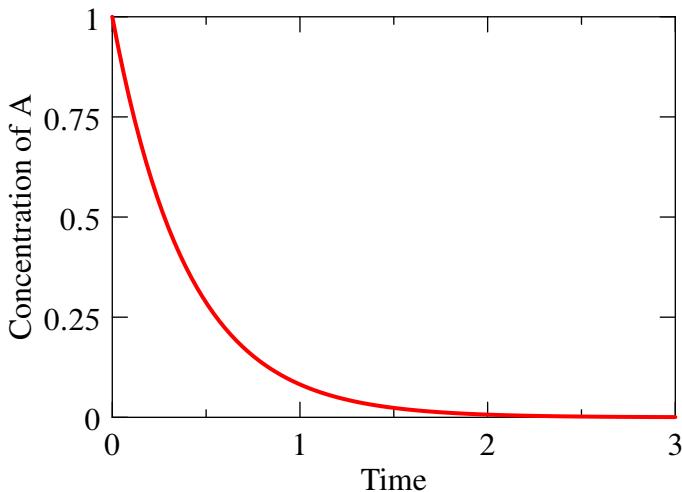


**Figure 1.2** Curves illustrating zero-order, first-order and second-order kinetics.

Equation (1.7) is a differential equation that can be solved using standard methods in differential calculus to describe the change in concentration of  $A$  over time. This solution is shown in Figure 1.3 and is described by the equation:

$$a(t) = a(0) e^{-kt} \quad (1.8)$$

where  $a(0)$  is the initial concentration of  $A$  and  $t$ , the time.



**Figure 1.3** Progress curve for species  $A$  in the irreversible reaction  $A \rightarrow B$ , computed using  $a(t) = a(0) e^{-kt}$  where  $a(0)$  is the initial concentration equal to 1.0,  $k$ , the rate constant, is equal to 2.5, and  $t$  the time. Due to conservation of mass, the change in  $B$  is given by  $b(t) = a(0) - a(t)$ .

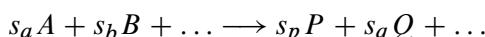
A common metric that is used to judge the rate of different first-order reactions is the half-life. This quantity measures the time taken for half the level of substance ( $A$ ) to be transformed into product. When half of  $A$  has been consumed, we can set  $a(t)$  in equation (1.8) to  $\frac{1}{2}a(0)$

so that:

$$\frac{1}{2} = e^{-kt_{1/2}}$$

The time,  $t_{1/2}$ , is called the **half-life** and by suitable rearrangement is given by,  $t_{1/2} = \ln(2)/k$ . For example if  $k = 0.5 \text{ sec}^{-1}$ , the half life is equal to  $\ln(2)/0.5 \simeq 1.4 \text{ sec}$ , that is after 1.4 sec half the concentration of substance has been consumed.

For the general reaction such as:



where  $s_x$  are the stoichiometric amounts, the rate law has been found through empirical observation to often have the form:

$$v = k a^{s_a} b^{s_b} \dots$$

where each reactant is raised to the power of its stoichiometric amount,  $s_i$ . For example, the reaction rate for the following reaction could be written as:



$$v = k \text{ adp}^2$$

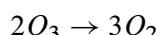
If the reaction is reversible then the rate law is appended with the reverse rate:

$$v = k_1 a^{s_a} b^{s_b} \dots - k_2 p^{s_p} q^{s_q} \dots \quad (1.9)$$

$$v = k_1 \text{ adp}^2 - k_2 \text{ atp amp}$$

The units for the reactant and product terms must be in concentration. The units for the rate constants,  $k$  will depend on the exact form of the rate law, but must be set to ensure that the rate of reaction is expressed in the units moles  $L^{-1}t^{-1}$ .

Although in many cases one will often assume a rate law of the form (1.9), this may not always be the case. For example, the gas reaction for the conversion of ozone to oxygen:



has been found experimentally to follow the rate law:

$$v = \frac{1}{2}k \frac{(O_3)^2}{O_2}$$

rather than the expected,  $v = k (O_3)^2$ . The reason for the discrepancy is that the decomposition of ozone into oxygen occurs via a series of elementary reactions, and it is the combination of these elementary reactions that gives rise to the non-elementary rate law. In biochemistry this effect is readily seen in enzyme kinetics where the rate laws are far from elementary.

## 1.6 The Equilibrium Constant

### Equilibrium

In principle, all reactions are reversible, that is, transformations can occur from reactant to product or product to reactant. The net rate of a reversible reaction is the difference between the forward and reverse rates. For a simple unimolecular reaction, such as:



we can write down the forward rate and reverse rates assuming elementary reaction kinetics as:

$$\begin{aligned} v_f &= k_1 a \\ v_r &= k_2 b \end{aligned} \tag{1.10}$$

The net rate of reaction,  $v$ , is given by the difference between the forward and reverse rates:

$$v = v_f - v_r$$

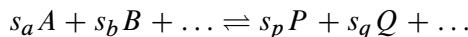
Furthermore, all reactions in an isolated system will tend to a state called **thermodynamic equilibrium**, that is when the forward and reverse rates are equal and the net rate is zero:

$$v_f - v_r = 0$$

Inserting equations (1.10) into the above yields the ratio:

$$\frac{k_1}{k_2} = \frac{b}{a} = K_{eq} \tag{1.11}$$

The ratio  $b/a$ , when the forward and reverse rates are equal, is called the **equilibrium constant**, often denoted by the symbol,  $K_{eq}$ . We can also see that the equilibrium constant is related to the ratio of the rate constants,  $k_1/k_2$ . For a general reversible reaction such as:



the equilibrium constant can be shown to be:

$$K_{eq} = \frac{p^{s_p} q^{s_q} \dots}{a^{s_a} b^{s_b} \dots} \tag{1.12}$$

where  $x_x$  in the exponent are the stoichiometric amounts for each species.

#### **Example 1.6**

Write out the equilibrium relationship for the reaction:



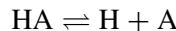
The equilibrium constant for this reaction is given by:

$$K_{eq} = \frac{ab}{a^2}$$

The equilibrium constant, after simplification, is therefore given by:

$$K_{eq} = \frac{b}{a}$$

Chemists and biochemists will often distinguish in bimolecular reactions between two kinds of equilibrium constants called association and dissociation constants. Thus the equilibrium constant for the reaction:



is often called the **dissociation constant**,  $K_d = [H][A]/[HA]$ <sup>1</sup> to indicate the degree that a complex will dissociate into its component molecules. The **association constant**,  $K_a$ , though less commonly used, describes the equilibrium constant for the reverse process, that is, the formation of a complex from component molecules:  $K_a = [HA]/([H][A])$ . It should be evident that  $K_b = 1/K_a$ .

While reactions tend to equilibrium in isolated systems, reactions occurring in living cells are generally out of equilibrium. The ratio of the products to the reactants *in vivo* is called by the **mass-action ratio**,  $\Gamma$ . At equilibrium, the mass-action ratio will be equal to the equilibrium constant. The ratio of the mass-action ratio to the equilibrium constant is often called the **disequilibrium ratio**,  $\Gamma/K_{eq}$ .

### Mass-action and Disequilibrium Ratio

The ratio of the products to the reactants *in vivo* is called the mass-action ratio,  $\Gamma$ . For the system,  $A \rightarrow B$ :

$$\Gamma = \frac{b_{in\ vivo}}{a_{in\ vivo}}$$

at equilibrium,  $\Gamma = K_{eq}$ . The ratio of the mass-action ratio to the equilibrium constant is often called the **disequilibrium ratio** and denoted by the symbol,  $\rho$ .

$$\rho = \frac{\Gamma}{K_{eq}}$$
(1.13)

At equilibrium, the mass-action ratio is equal to the equilibrium constant and  $\rho = 1$ . If the reaction is far from equilibrium ( $b/a < K_{eq}$ ), then  $\rho < 1$ .

<sup>1</sup>Square brackets indicates concentration.

**Example 1.7**

Given the reaction:



with *in vivo* concentrations of  $a = 5$  and  $b = 4$ , compute the mass-action ratio,  $\Gamma$ . If the equilibrium constant  $K_{eq}$  is 7.5, is the reaction at equilibrium or not?

$$\Gamma = \frac{b_{in\ vivo}}{a_{in\ vivo}} = \frac{4}{5} = 0.8$$

Since  $\Gamma$  does not equal  $K_{eq}$ , the reaction is not at equilibrium.

For the general reaction, the mass-action ratio is given by:

$$\Gamma = \frac{p_{in\ vivo}^{s_p} q_{in\ vivo}^{s_q} \cdots}{a_{in\ vivo}^{s_a} b_{in\ vivo}^{s_b} \cdots} \quad (1.14)$$

where the exponents are the stoichiometric **amounts** for each species. As stated before, if the concentrations are all at equilibrium, then  $\Gamma = K_{eq}$ . For a simple unimolecular reaction it was shown previously that the equilibrium ratio of product to reactant,  $b/a$ , is equal to the ratio of the forward and reverse rate constants,  $k_1/k_2$ . Substituting this into the disequilibrium ratio gives:

$$\rho = \frac{\Gamma}{K_{eq}} = \Gamma \frac{k_2}{k_1} = \frac{b}{a} \frac{k_2}{k_1}$$

But  $bk_2 = v_f$  and  $ak_1 = v_r$  (1.10), therefore:

$$\rho = \frac{v_r}{v_f} \quad (1.15)$$

Thus the disequilibrium ratio is the ratio of the reverse and forward rates. This relationship clearly shows how the disequilibrium ratio tells us whether the reaction is going forward, is at equilibrium, or whether it is in reverse. If  $\rho < 1$ , then the net reaction must be in the direction of product formation since  $v_f > v_r$ . If  $\rho = 1$ , then  $v_r = v_f$ , and the system is at equilibrium. Finally if  $\rho > 1$  then  $v_r > v_f$ , the reaction must be going in reverse.

If we take the natural log of equation (1.13) on both sides we get:

$$\ln(\rho) = \ln(\Gamma) - \ln(K_{eq}) \quad (1.16)$$

With this transformation, if  $\ln(\rho)$  is negative, the reaction must be in the forward direction, zero if the reaction is at equilibrium, and greater than zero if the reaction is in the reverse direction.

$\ln(\rho)$	Direction of Reaction	$v$	$\Delta G$
< 0	Forward Direction	$v > 0$	$\Delta G < 0$
= 0	Equilibrium	$v = 0$	$\Delta G = 0$
> 0	Reverse Direction	$v < 0$	$\Delta G > 0$

**Table 1.2** Relationship between  $\rho$  and  $\Delta G$ .

Those who are already familiar with the concept of free energy ( $\Delta G$ ) may realize that equation (1.16) is closely related to the free energy equation:

$$\Delta_r G = \Delta_r G^\circ + RT \ln \Gamma$$

where  $\Delta_r G = RT \ln(\rho)$  and  $\Delta_r G^\circ = -RT \ln K_{eq}$ . Because all the rate information has been lost in the derivation of the equation (1.13), the value of  $\ln(\rho)$  tells us nothing about how fast the reaction will proceed, only the direction it proceeds. The free energy can be expressed as follows:

$$\begin{aligned} \Delta_r G &= -RT \ln K_{eq} + RT \ln \Gamma \\ &= RT \ln \Gamma - RT \ln K_{eq} \\ &= RT \ln(\Gamma/K_{eq}) \\ &= -RT \ln(K_{eq}a/b) \end{aligned}$$

Similarly  $v_f = k_1 a$ ,  $v_r = k_2 b$ , and  $K_{eq} = k_1/k_2$ , therefore:

$$\frac{v_f}{v_r} = \frac{k_1 a}{k_2 b} = K_{eq} \frac{a}{b}$$

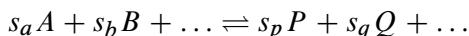
Which leads to:

$$\Delta_r G = -RT \ln(v_f/v_r)$$

This is just a reexpression of equation (1.15) where we have taken logs on both sides and multiplied by  $RT$ . It is sometimes convenient to reexpress the equation in the form:

$$\frac{v_f}{v_r} = e^{-\Delta_r G/RT}$$

Since  $\Delta_r G$  can be computed from tables of free energies of formation, it is possible to obtain  $v_f/v_r$ . Relation (1.15) is actually much more general and applies to any reaction of the form:



The disequilibrium ratio is an important quantity and reappears in later sections and chapters when we discuss enzymatic reactions. It is particularly relevant when one considers the control of cellular pathways.

## Principle of Detailed Balance

In its simplest form, the principle of detailed balance says that the forward and reverse rates must be equal at thermodynamic equilibrium. For the simple reversible reaction:



where the forward rate  $v_f$  is given by  $v_f = k_1 a$  and the reverse rate,  $v_r$  by  $v_r = k_2 b$ , detailed balance states that  $v_f = v_r$  at equilibrium, or:

$$k_1 a_{eq} = k_2 b_{eq}$$

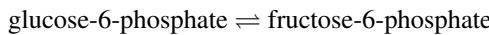
From the definition of the equilibrium constant we see that:

$$K_{eq} = \frac{b_{eq}}{a_{eq}} = \frac{k_1}{k_2}$$

### Example 1.8

The equilibrium constant for the reaction between glucose-6-phosphate and fructose-6-phosphate catalyzed by glucose-6-phosphatase isomerase (EC 5.3.1.9) is known to have a value of 0.395 at  $25^\circ C$ . The concentration of glucose-6-phosphate in liver cells is estimated to be 4.9 mM. Assuming the reaction is at equilibrium, estimate the concentration of fructose-6-phosphate.

This is a straightforward calculation. The reaction is described by:



and the equilibrium constant is therefore given by:

$$K_{eq} = \frac{\text{Fructose-6-Phosphate}}{\text{Glucose-6-Phosphate}}$$

By simple rearrangement the fructose-6-phosphate concentration is equal to:

$$\begin{aligned} \text{fructose-6-phosphate} &= K_{eq} \text{ glucose-6-phosphate} = \\ &0.395 \times 4.9 = \\ &1.94 \text{ mM} \end{aligned}$$

### Example 1.9

The previous problem can be made more difficult by stating that the *total* concentration of glucose-6-phosphate and fructose-6-phosphate is 4.9 mM. The question now is how to compute the equilibrium concentration of both species. The calculation begins by constructing an equilibrium table:

Species	G6P	F6P
Initial concentration	4.9	0
Equilibrium concentration	$4.9 - x$	$x$

$\text{G6P} \equiv \text{Glucose-6-Phosphate}$

$\text{F6P} \equiv \text{Fructose-6-Phosphate}$

From the equilibrium constant and the above table, we derive the following:

$$K_{eq} = \frac{x}{4.9 - x}$$

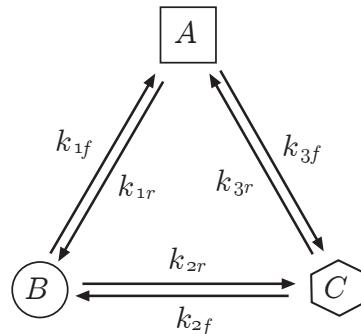
Solving for  $x$  and hence the equilibrium concentration, yields:

$$x = \frac{K_{eq} \cdot 4.9}{1 + K_{eq}}$$

Therefore the equilibrium concentration of glucose-6-phosphate is 1.387 mM and for fructose-6-phosphate, 3.514 mM. A simple check that the ratio, 1.387/3.514, equals the equilibrium constant will confirm the result.

With more complex reactions, the above method will yield polynomial equations which can have multiple solutions, usually one negative and the other positive. It should be clear however that the negative solution is physically impossible which leaves the other as the solution we seek.

Detailed balance is more useful when applied to more complex systems. Consider the system shown in Figure 1.4 which comprises of three species linked by three reversible reactions. Each reaction has a forward and reverse rate constant.



**Figure 1.4** The principle of detailed balance.

At equilibrium the following must be true:

$$\frac{k_{1f}}{k_{1r}} = \frac{a}{b}, \quad \frac{k_{2f}}{k_{2r}} = \frac{b}{c} \quad \text{and} \quad \frac{k_{3f}}{k_{3r}} = \frac{c}{a}$$

Combining the three equations and eliminating  $a$ ,  $b$ , and  $c$  yields the following relation among the rate constants:

$$k_{1f}k_{2f}k_{3f} = k_{1r}k_{2r}k_{3r}$$

That is, the product of rate constants in one direction around the loop must equal the product of rate constants in the opposite direction around the loop. A restatement of the above relation is that the product of the equilibrium constants in a loop is one:

$$\frac{a}{b} \frac{b}{c} \frac{c}{a} = K_{eq1} K_{eq2} K_{eq3} = 1 \quad (1.17)$$

This is true irrespective of the actual reaction mechanism. Detailed balance gives a constraint on the allowable rate and equilibrium constants in a reaction loop. In addition, it means that the change in free energy around the loop is zero. By analogy, one can compare detailed balance to a hike over a mountain range, where a hiker traverses one peak after another. If the route the hiker takes eventually returns them to the original starting peak, the net change in height during the entire hike is zero.

## 1.7 Reaction Order

As introduced previously, the rate of reaction for elementary reactions is given by the general equation:

$$v = k a^{s_a} b^{s_b} \dots$$

where  $s_i$  is the stoichiometric amount for the  $i^{th}$  reactant. In this formalism, the power to which a reactant is raised is termed the **reaction order**. Strictly speaking the reaction order should be determined experimentally, but often it is assumed that the reaction is elementary and then the reaction order is related directly to the stoichiometry of reaction.

Thus, the order of reaction for the simple decay reaction,  $A \rightarrow B$ , is equal to one with respect to the reactant  $A$ .

For elementary reactions the reaction orders are generally constant integers, and unless indicated otherwise, are usually related to the stoichiometry of reaction. However, reaction orders need not necessarily be integers and in many biological processes, for example enzyme catalyzed reactions, the reaction order is not a constant but is a function of the concentration of the reactant species.

For an arbitrary rate law,  $v$ , the reaction order can also be expressed using the elasticity<sup>2</sup> of the reaction:

$$\varepsilon_a^v = \frac{\partial v}{\partial a} \frac{a}{v} \quad (1.18)$$

where  $v$  is the reaction rate and  $a$  the concentration of a reactant. This definition accommodates integer, fractional and varying orders of reaction. For an elementary reaction, the

<sup>2</sup>In the older literature this expression is also referred to as the kinetic order [113, 22].

elasticities can be easily shown to equal the reaction orders, that is:

$$v = k a^{s_a} b^{s_b} \dots$$

The two elasticities with respect to  $a$  and  $b$  can be shown by differentiating the expression and scaling to be:

$$\begin{aligned}\varepsilon_a^v &= s_a \\ \varepsilon_b^v &= s_b\end{aligned}$$

---

**Example 1.10**

Consider the reaction:



where the rate of reaction is expressed as:

$$v = k_1 \text{ adp}^2 - k_2 \text{ atp amp}$$

The elasticity can be derived by differentiating the rate law with respect to each species and scaling the derivative:

$$\begin{aligned}\varepsilon_{\text{adp}}^v &= 2 \\ \varepsilon_{\text{atp}}^v &= -1 \\ \varepsilon_{\text{amp}}^v &= -1\end{aligned}$$

---

The concept of the elasticity will be examined in much more detail in Chapter 6.

---

## Further Reading

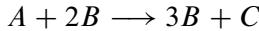
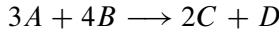
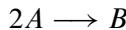
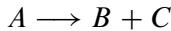
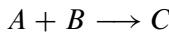
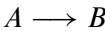
1. Sauro H M (2011) Enzyme Kinetics for Systems Biology. Ambrosius Publishing.  
ISBN: 0982477317

---

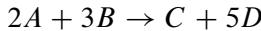
## Exercises

1. Explain the difference between the terms: Stoichiometric amount, stoichiometric coefficient, rate of change ( $dx/dt$ ), and reaction rate ( $v_i$ ).
2. What does a negative stoichiometric coefficient mean?
3. Determine the stoichiometric amount and stoichiometric coefficient for each species

in the following reactions:



4. For the following reaction:



- a) Write the stoichiometry amounts for the species,  $A$ ,  $B$ ,  $C$  and  $D$ .
- b) Write out the stoichiometric coefficients for each species in the reaction.
- c) Write the reversible mass-action rate law for the reaction.
- d) Write the rates of change for each species in terms of the reaction rate,  $v$ .

5. For the reaction  $2A \rightarrow B$ , the rate of change of  $A$  is 2 moles sec $^{-1}$ . What is the rate of reaction?

6. The number of copies of the enzyme phosphofructokinase in an *E. coli* cell is 1550. Assuming a volume of  $1.57 \times 10^{-18}$  m $^3$  for an *E. coli* cell, what is the concentration of this enzyme? Use the most appropriate units in your answer.

7. Compute the reaction order with respect to  $A$  and  $B$ , for the rate law:

$$v = k_1 a^3 b^2$$

8. If the equilibrium constant ( $K_{\text{eq}}$ ) for the reaction  $A \rightleftharpoons B$  is 6.0, calculate the equilibrium concentration of  $B$  if the equilibrium concentration of  $A$  is 10 mM.
9. A protein undergoes a reversible transformation ( $A \rightleftharpoons AP$ ) with an equilibrium constant of 25.0. Explain why the equilibrium constant unitless. If the total concentration of protein ( $A$  and  $AP$ ) is 10 mM, what are the equilibrium concentrations of  $A$  and  $AP$ ?
10. Hemoglobin contains 0.335% by mass of Fe. Four atoms of Fe are present in each molecule of hemoglobin. If the atomic mass of Fe is 55.84 g/mol, estimate the molar mass of hemoglobin.
11. The reaction  $A \rightarrow B$  has a rate of reaction of  $0.75 \text{ mmoles } s^{-1} L^{-1}$ . What is the reaction rate if the reaction takes place in  $5 \mu\text{l}$ ?

# 2

## *Enzyme Kinetics in a Nutshell*

### **2.1 Enzymes**

---

Enzymes are protein molecules that can accelerate a chemical reaction without changing the equilibrium constant of the reaction.

#### **Enzyme Kinetics**

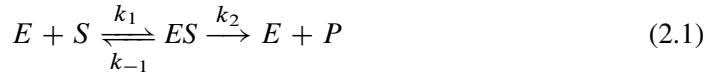
Enzyme kinetics is a branch of science that deals with the many factors that can affect the rate of an enzyme-catalysed reaction. The most important factors include the concentration of enzyme, reactants, products, and the concentration of any modifiers such as specific activators, inhibitors, pH, ionic strength, and temperature. When the action of these factors is studied, it is possible to deduce the kinetic mechanism of the reaction. That is, the order in which substrates and products bind and unbind, and the mechanism by which modifiers alter the reaction rate. This chapter will provide a very brief overview of enzyme kinetics. A much more detailed discussion can be found in the companion book “Enzyme Kinetics for Systems Biology” [109].

### **2.2 Michaelis-Menten Kinetics**

---

The standard model for enzyme action describes the binding of free enzyme to the reactant forming an **enzyme-reactant complex**. This complex undergoes a transformation, releasing product and free enzyme. The free enzyme is then available for another round of binding

to a new reactant. This mechanism can be written using the following reaction scheme:



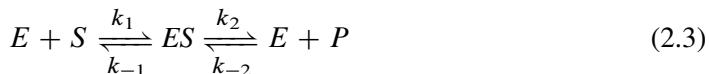
where  $k_1$ ,  $k_{-1}$  and  $k_2$  are rate constants,  $S$  is substrate,  $P$  is product,  $E$  is the free enzyme, and  $ES$  the enzyme-substrate complex. Estimating the various rate constants tends to be difficult and much more difficult for complex mechanisms. As a result, enzymologists use approximations to reduce the complexity. The first such approximation to be introduced was the rapid equilibrium assumption. In this case it is assumed that complex formation between enzyme and substrate is so fast that the binding/unbinding step is always in thermodynamic equilibrium. This allows the concentration of the enzyme-substrate complex to be derived as a function of the total enzyme, free substrate concentration and the various rate constants. The reaction rate is assumed to be proportional to the concentration of substrate complex. The resulting equation is called the Michaelis-Menten equation. Briggs and Haldane relaxed the requirement for equilibrium, and instead assumed a quasi-steady state assumption where the enzyme-substrate complex rapidly approached steady state. As with the rapid equilibrium assumption, the concentration of enzyme-substrate complex could be computed quite easily, from which the net rate could be estimated. Later on systematic methods, such as the King-Altman method [69], were derived to make it possible to derive the equations by computer. The resulting Briggs-Haldane equation for the mechanism (2.1) is given by:

$$v = \frac{V_m s}{K_m + s} \quad (2.2)$$

where  $V_m$  is the maximal velocity, and  $K_m$  the substrate concentration that yields half the maximum velocity. The Briggs-Haldane approach is often the most common way to derive enzymatic rate laws although more complex systems such as allosteric, cooperativity and gene regulation still use the rapid equilibrium approach because the approach is technically simpler. A detailed examination of the history and origins of this equation can be found in [23].

## Reversible Rate laws

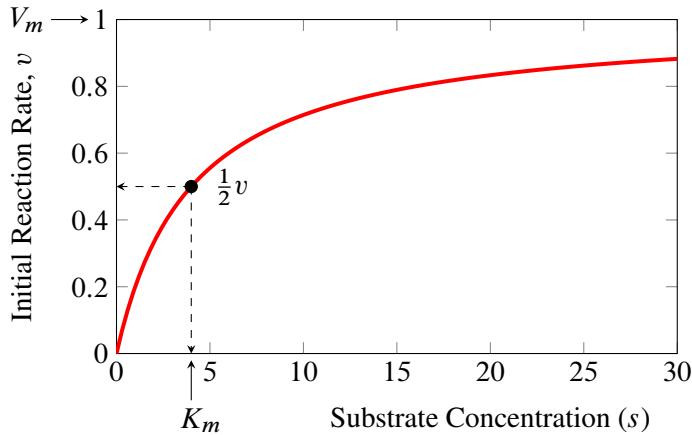
An alternative and more realistic model is the reversible form:



Assuming quasi-steady-state, the following rate law can be derived and is given by:

$$v = \frac{V_f s/K_s - V_r p/K_p}{1 + s/K_s + p/K_p} \quad (2.4)$$

where  $V_f$  and  $V_r$  are the forward and reverse maximal rates, and  $K_s$  and  $K_p$  the Michaelis constants.



**Figure 2.1** Relationship between the initial rate of reaction and substrate concentration for a simple Michaelis-Menten rate law. The reaction rate reaches a limiting value called the  $V_m$ .  $K_m$  is set to 4.0 and  $V_m$  to 1.0. The  $K_m$  value is the substrate concentration that gives half the maximal rate.

### Haldane Relationship

For the reversible enzyme kinetic law there is an important relationship:

$$K_{eq} = \frac{p_{eq}}{s_{eq}} = \frac{V_f K_p}{V_r K_s} \quad (2.5)$$

which shows that the four kinetic constants,  $V_f$ ,  $V_r$ ,  $K_p$  and  $K_s$  are not independent. Haldane relationships can be used to eliminate one of the kinetic constants by substituting the equilibrium constant in its place. This is useful because equilibrium constants tend to be more commonly known compared to kinetic constants. By incorporating the Haldane relationship we can eliminate the reverse maximal velocity ( $V_r$ ) in equation (2.4) to yield:

$$v = \frac{V_f / K_s (s - p / K_{eq})}{1 + s / K_s + p / K_p} \quad (2.6)$$

Separating out the terms makes it easier to see that the above equation can be partitioned into a number of distinct terms:

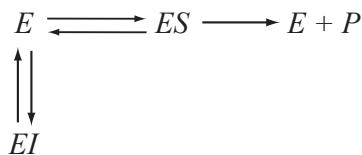
$$v = V_f \cdot (1 - \Gamma / K_{eq}) \cdot \frac{s / K_s}{1 + s / K_s + p / K_p} \quad (2.7)$$

where  $\Gamma = p/s$ . The first term,  $V_f$ , is the maximal velocity; the second term,  $(1 - \Gamma / K_{eq})$  indicates the direction of the reaction according to thermodynamic considerations. The last term refers to the fractional saturation with respect to substrate. We thus have a maximal velocity, a thermodynamic term and a saturation term.

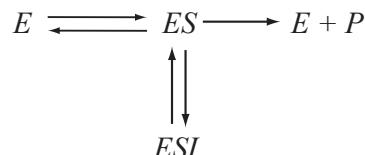
## Competitive Inhibition

There are many molecules capable of slowing down or speeding up the rate of enzyme catalyzed reactions. Such molecules are called enzyme inhibitors and activators. One common type of inhibition, called **competitive inhibition**, occurs when the inhibitor is structurally similar to the substrate or product so that it competes for the active site by forming a dead-end complex.

### a) Competitive Inhibition



### b) Uncompetitive Inhibition



**Figure 2.2** Competitive and uncompetitive inhibition.  $P$  is the product,  $E$  is the free enzyme,  $ES$  the enzyme-substrate complex, and  $ESI$  the enzyme-substrate-inhibitor complex.

The kinetic mechanism for a pure competitive inhibitor is shown in Figure 2.2(a), where  $I$  is the inhibitor and  $EI$  the enzyme inhibitor complex. If the substrate concentration is increased, it is possible for the substrate to eventually out compete the inhibitor. For this reason the inhibitor alters the enzyme's apparent  $K_m$ , but not the  $V_m$ . Assuming quasi-steady-state, the rate law for competitive inhibition is given by:

$$\begin{aligned} v &= \frac{V_m s}{s + K_m \left(1 + \frac{i}{K_i}\right)} \\ &= \frac{V_m s / K_m}{1 + s / K_m + i / K_i} \end{aligned} \tag{2.8}$$

At  $i = 0$ , the competitive inhibition equation reduces to the irreversible Michaelis-Menten equation. Note that the term  $K_m(1 + i/K_i)$  in the first equation more clearly shows the impact of the inhibitor,  $I$ , on the  $K_m$ . The inhibitor has no effect on the  $V_m$ . The reversible form of the competitive rate law can also be derived and is shown below:

$$v = \frac{\frac{V_m}{K_s} \left(s - \frac{p}{K_{eq}}\right)}{1 + \frac{s}{K_s} + \frac{p}{K_p} + \frac{i}{K_i}} \tag{2.9}$$

where  $V_m$  is the forward maximal velocity, and  $K_s$  and  $K_p$  are the substrate and product half saturation constants.

## Product Inhibition

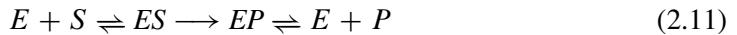
Sometimes reactions appear irreversible. That is, no discernable reverse rate is detected, and yet the forward reaction is influenced by the accumulation of product. This effect is caused by the product competing with substrate for binding to the active site and is often called **product inhibition**. Given that product inhibition is a type of competitive inhibition, we will briefly discuss it here. An important industrial example of this is the conversion of lactose to galactose by the enzyme  $\beta$ -galactosidase where galactose competes with lactose, slowing the forward rate [42].

To describe simple product inhibition with rate irreversibility, we can set the  $p/K_{eq}$  term in the reversible Michaelis-Menten rate law (2.9) to zero. This yields:

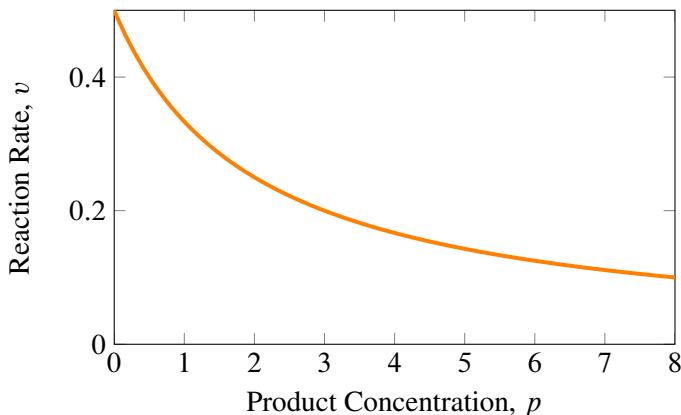
$$v = \frac{V_m s}{s + K_m \left(1 + \frac{p}{K_p}\right)} \quad (2.10)$$

It is not surprising to discover that equation (2.10) has exactly the same form as the equation for competitive inhibition (2.8). Figure 2.3 shows how the reaction rate responds to increasing product concentration at a fixed substrate concentration. As the product increases, it out competes the substrate and therefore slows down the reaction rate.

We can also derive the equation by using the following mechanism and the rapid equilibrium assumption:



where the reaction rate,  $v \propto ES$ .



**Figure 2.3** Effect of increasing the concentration of product on the reaction rate at a fixed substrate concentration for an irreversible reaction with product inhibition. The reaction rate declines monotonically to zero as the product increases.  $s = 1$ ,  $V_m = 1$ ,  $K_m = 1$ ,  $K_i = 1$ .

## 2.3 Multisubstrate Kinetics

---

It is probably fair to say that most enzyme catalyzed reactions are multisubstrate/multiproduct. Common examples include any reactions involving cofactors such as ATP, NAD, or Coenzyme A. All hydrolase, transferases and oxidoreductases catalyze multisubstrate reactions. An example of a multisubstrate reaction is the phosphorylation of glucose to form glucose-6-phosphate catalyzed by Hexokinase.



The problem when dealing with multisubstrate catalyzed enzymes is that there are a number of ways in which substrate can bind. If we assume the reaction  $A + B \rightarrow P + Q$ , then a number of situations present themselves:

**Sequential Mechanism:**  $A$  binds first, followed by  $B$ .  $P$  leaves first, followed by  $Q$ . That is, there is a strict order in which substrates bind and unbind to the active site.

**Random Mechanism:**  $A$  or  $B$  can bind first, followed by binding of the remaining substrate.  $P$  and  $Q$  unbind in any order.

**Ping-Pong:**  $A$  binds first, followed by the unbinding of  $P$ .  $B$  then binds, followed by the unbinding of  $Q$ .

Given these mechanisms, it is common to derive the rate equations using the steady state assumption, although it is possible to derive them using rapid equilibrium which can lead to simpler expressions. Segel [122] provides a large list of mechanisms and their associated rate laws.

An alternative to using specific multisubstrate rate laws is to use generalized rate laws such as the Hanekom [49, 105, 150] or Liebermesiter [76, 77] rate laws. The mechanistic basis for deriving the saturation term in the Hanekom equation is a rapid-equilibrium random-order mechanism, which according to previous studies [105], appears to fulfill most kinetic requirements. In its most general form, the equation is given by:

$$v = V_f \prod_{i=1}^n \frac{s_i}{K_{s_i}} \left( 1 - \frac{\Gamma}{K_{eq}} \right) \prod_{i=1}^n \left( \frac{1}{1 + \frac{s_i}{K_{s_i}} + \frac{p_i}{K_{p_i}}} \right) \quad (2.12)$$

Where  $V_f$  is the maximal forward rate, and  $K_{s_i}$  and  $K_{p_i}$  are the substrate and product half-saturation constants.  $\Gamma$  is the mass-action ratio and  $K_{eq}$  the equilibrium constant. To give a concrete example, consider the enzyme catalyzed reversible reaction  $A + B \rightarrow C$ . The generalized rate law is given by:

$$v = V_f \frac{a}{K_a} \frac{b}{K_b} \frac{1 - \Gamma / K_{eq}}{\left( 1 + \frac{a}{K_a} + \frac{p}{K_p} \right) \left( 1 + \frac{b}{K_b} \right)}$$

The Liebermesiter rate laws take things further by explicitly modularizing the rate law. Further details can be found in the source publications [76, 77] or the companion book [109].

More studies need to be undertaken to determine exactly what is the effect of substituting a specific kinetic mechanism with a generalized rate law. One of the key issues is the effect each kinetic mechanism has on the elasticity coefficients (1.7).

## 2.4 Cooperativity

---

Many proteins are known to be oligomeric, that is they are composed of more than one identical protein subunit where each subunit has one or more binding sites. Often the individual subunits are identical.

If the binding of a ligand (a small molecule that binds to a larger macromolecule) to one site alters the affinity at other sites on the same oligomer, we observe what is called **cooperativity**. If ligand binding increases the affinity of subsequent binding events, it is termed **positive cooperativity**, whereas if the affinity decreases, then this often results in a phenomena called **negative cooperativity**. One characteristic of positive cooperativity is that it results in a sigmoidal response instead of the usual hyperbolic response.

The simplest equation that displays sigmoid like behavior is the Hill equation:

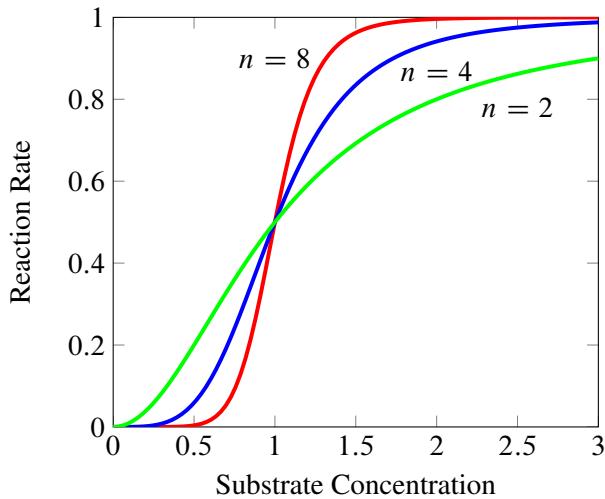
$$v = \frac{V_m s^n}{K_d + s^n} = \frac{V_m s^n}{K_H^n + s^n} \quad (2.13)$$

where  $K_H$  is the concentration of ligand that yields half the maximal rate.

The Hill equation is commonly used to model the rate of gene expression in response to an inducer. The equation originally had no mechanistic underpinning and was purely an empirical device used to fit sigmoidal data observed during oxygen binding to hemoglobin. However, one can derive the Hill equation by assuming that all ligands bind simultaneously to the protein complex, an unrealistic scenario. As such, the Hill equation offers little insight into the mechanistic understanding of cooperativity and likely misses more subtle behavior. As a result, other models with more realistic mechanisms were introduced, the most famous being the concerted model by Monod, Wyman and Changeux [85] and the sequential model by Koshland, Nemethy and Filmer [73].

One striking feature of many oligomeric proteins is the way individual monomers are physically arranged. Often one will find at least one axis of symmetry. The individual protein monomers are not arranged in a haphazard fashion. This level of symmetry may imply that the gradual change in the binding constants as ligands bind might be physically implausible. Instead, one might envision transitions to an alternative binding state that occur within the entire oligomer complex. Such a model, termed the MWC model, was proposed by Monod, Wyman and Changeux in 1965 (MWC model, [85]). The authors laid out the following criteria for the MWC model:

1. The protein is an oligomer.

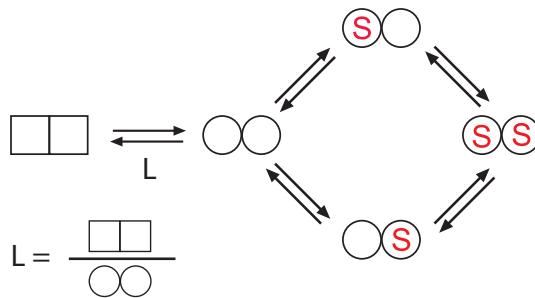


**Figure 2.4** Plots showing the response of the Hill equation (2.13) with different values for  $n$ ,  $K_d = 1$ .

2. Oligomers can exist in two states: R (relaxed) and T (tense). In each state, symmetry is preserved and all subunits must be in the same state for a given R or T state.
3. The R state has a higher ligand affinity than the T state.
4. The T state predominates in the absence of ligand  $S$ .
5. The ligand binding microscopic association constants are all identical.

Given these criteria, the MWC model assumes that an oligomeric enzyme may exist in two conformations, designated T (tensed, square) and R (relaxed, circle) with an equilibrium between the two states with equilibrium constant,  $L = T/R$ , also called the allosteric constant. If the binding constants of ligand to the two states are different, then the distribution of the R and T forms can be displaced either towards one form or the other. By this mechanism, the enzyme displays sigmoid behavior. A minimal example of this model is shown in Figure 2.5.

In the **exclusive model** (Figure 2.5), the ligand can only bind to the relaxed form (circle). The mechanism that generates sigmoidicity in this model works as follows. When ligand binds to the relaxed form, it displaces the equilibrium from the tense form to the relaxed form. In doing so, additional ligand binding sites are made available. Thus one ligand binding may generate two or more new binding sites. Eventually there are no more tense states remaining at which point the system is saturated with ligand. The overall binding curve will therefore be sigmoidal and will show positive cooperativity. Given the nature of this model, it is not possible to generate negative cooperativity. By assuming equilibrium



**Figure 2.5** A minimal MWC model, also known as the exclusive model, showing alternative microscopic states in the circle (relaxed) form.  $L$  is called the allosteric constant. The square form is called the tense state.

between the various states, it is possible to derive an aggregate equation for the dimer case of the exclusive MWC model:

$$v = V_m \frac{\frac{s}{K_R} \left(1 + \frac{s}{K_R}\right)}{\left(1 + \frac{s}{K_R}\right)^2 + L}$$

This also generalizes to  $n$  subunits as follows:

$$Y = \frac{\frac{s}{K_R} \left(1 + \frac{s}{K_R}\right)^{n-1}}{\left(1 + \frac{s}{K_R}\right)^n + L} \quad (2.14)$$

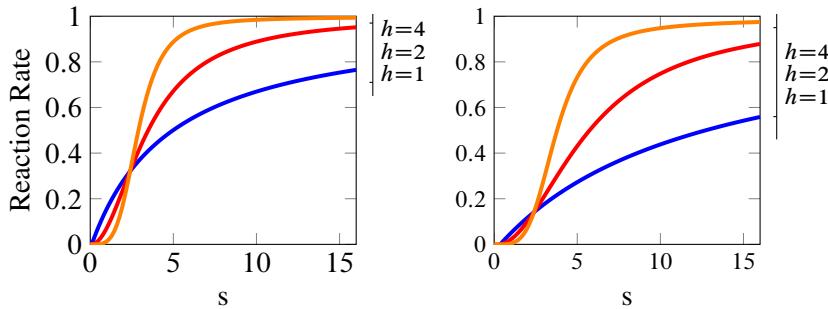
where  $Y$  is the fractional saturation by ligand. The reaction velocity is assumed to be proportional to  $Y$ . Reversible forms for the MWC model exist and can be found in the companion book [109]. The sequential model by Koshland and coworkers takes a different approach by assuming that the binding of a ligand to one subunit results in a conformational change in the subunit, and thereby alters the interactions of the subunit with its neighbors. Equations derived using the sequential model are more complex.

### Generalized Reversible Equation for Cooperativity

There is a recent trend towards using improved models of the Hill equation. For the concerted model and especially the sequential models, the rate laws can become quite complex and unmanageable. The most well known generalized model for cooperativity is the reversible Hill equation by Hofmeyr and Cornish-Bowden [56]. This is based on a much more realistic mechanistic model.

$$v = \frac{V_f s / K_s (1 - \rho) (s / K_s + p / K_p)^{h-1}}{1 + (s / K_s + p / K_p)^h} \quad (2.15)$$

Figure 2.6 illustrates the sigmoid behavior with respect to the substrate concentration. The  $K$  constants in the equation are the half saturation constants, and  $\rho$  is the disequilibrium constant for the reaction. The significance of formulating this is that the thermodynamic terms are explicitly separated from the saturation terms. The equation also reduces to familiar forms when certain restrictions are applied. For example if  $h = 1$ , the equation reduces to the non-cooperative reversible Michaelis rate law and if reversibility is removed, the equation reduces to the simple irreversible Michaelis-Menten rate law. The equation can also revert to the product inhibited and irreversible rate law by setting the  $K_{eq}$  to infinity. The reversible Hill equation also shows another interesting property. Under a certain set of parameter values, the product concentration can act as a positive regulator. The reversible Hill equation is therefore quite flexible and can be used in many situations.

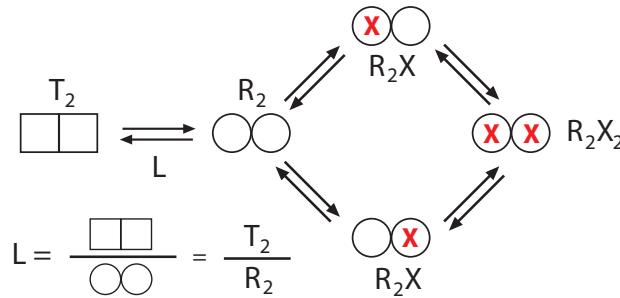


**Figure 2.6** Plot showing the response of the reaction rate for a reversible Hill model with respect to the substrate as a function of the Hill coefficient,  $h$ . The parameters were set as follows:  $V_m = 1$ ,  $\Gamma = 2$ ,  $K_{eq} = 10.95$ ,  $K_p = 0.5$ ,  $K_s = 2.75$ , Left Panel  $p = 1$ , Right Panel  $p = 4$ .

## 2.5 Allostery

An allosteric effect is where the activity of an enzyme or other protein is affected by the binding of an effector molecule at a site on the protein's surface other than the active site. The MWC model described previously can be easily modified to accommodate allosteric action.

The key to including allosteric effectors is the equilibrium between the tense (T) and relaxed (R) states (See Figure 2.7). To influence the sigmoid curve, an allosteric effector need only displace the equilibrium between the tense and relaxed forms. For example, to behave as



**Figure 2.7** Exclusive MWC model based on a dimer showing alternative microscopic states in the form of  $T$  and  $R$  states. The model is exclusive because the ligand,  $X$ , only binds to the  $R$  form.

an activator, an allosteric effector needs to preferentially bind to the  $R$  form and shift the equilibrium away from the less active  $T$  form. An allosteric inhibitor would do the opposite, bind preferentially to the  $T$  form so that the equilibrium shifts towards the less active  $T$  form. In both cases, the  $V_m$  of the enzyme is unaffected.

The net result of this is to modify the normal MWC aggregate rate law to the following if the effector is an inhibitor:

$$v = V_m \frac{\alpha (1 + \alpha)^{n-1}}{(1 + \alpha)^n + L(1 + \beta)^n} \quad (2.16)$$

where  $\alpha = s/K_s$  and  $\beta = i/K_I$ .  $K_s$  and  $K_I$  are kinetic constants related to each ligand. A MWC model that is regulated by either an inhibitor,  $i$ , or an activator,  $a$ , is represented by:

$$v = V_m \frac{\alpha (1 + \alpha)^{n-1}}{(1 + \alpha)^n + L \frac{(1 + \beta)^n}{(1 + \gamma)^n}}$$

where  $\gamma = a/K_A$  with  $a$  equal to the concentration of activator. As with the cooperative MWC model, there are also reversible forms of the allosteric MWC model but these can be fairly complex. Instead, it is possible to modify the reversible Hill rate law [56] to include allosteric ligands.

$$v = \frac{V_f \alpha \left(1 - \frac{\Gamma}{K_{eq}}\right) (\alpha + \pi)^{h-1}}{\frac{1 + \mu^h}{1 + \sigma \mu^h} + (\alpha + \pi)^h} \quad (2.17)$$

where:

$$\begin{aligned} \sigma < 1 && \text{inhibitor} \\ \sigma > 1 && \text{activator} \end{aligned}$$

Enzyme kinetics is an extensive topic in its own right, and it is not possible to cover the many topics in a brief chapter. A list of recommended texts is given below.

## **Further Reading**

---

1. Cornish-Bowden A (2012) Fundamentals of Enzyme Kinetics, 4th edition. Wiley Blackwell, ISBN: 978-3527330744
2. Palmer T (1995) Understanding Enzymes 4 SubEdition, Prentice Hall, ISBN: 978-0131344709
3. Sauro HM (2012) Enzyme Kinetics for Systems Biology. 2nd Edition, Ambrosius Publishing ISBN: 978-0982477335
4. Segel IH (1993) Enzyme Kinetics: Behavior and Analysis of Rapid Equilibrium and Steady-State Enzyme System, Wiley Classics Library, ISBN: 978-0471303091

# 3

## ***Stoichiometric Networks***

### **3.1 Stoichiometric Networks**

---

Almost all cellular events involve some kind of chemical process that includes binding, unbinding, or transformation of compounds in specific stoichiometric amounts. The binding of the yeast cell cycle proteins cdc2 and cdc13 to form a cdc2-cdc13 complex, or the isomerization of glucose-6-phosphate to fructose-6-phosphate, are two notable examples. When we put a collection of these processes together, we form a **stoichiometric network**.

One of the key characteristics of stoichiometric networks is that mass is conserved at each transformation step. For example, the transformation  $S_1 \rightarrow S_2$  means that when one molecule of  $S_1$  disappears, one molecule of  $S_2$  is formed. An example of a very simple and minimal stoichiometric network is the two step pathway shown below:



In this system mass is conserved at every stage. In more sophisticated models where electric charge is also considered, charge will be conserved as well.

### **Elementary Reactions**

Chemical reactions that involve no reaction intermediates other than a single transition state are called **elementary reactions**. Elementary reactions have been depicted in a number of ways in the literature. For example, the transformation of one species into another can be

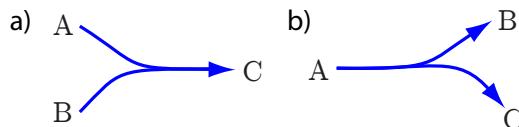
represented by a simple line with an arrow at the tip. The direction of the arrow indicates the direction of the *positive* reaction rate (Figure 3.1). If a reaction rate is  $-0.75 \text{ mol L}^{-1}$ , this means the reaction proceeds in the *opposite* direction indicated by the arrow at a rate of  $0.75 \text{ mol L}^{-1}$ . Most if not all reactions are in principle reversible, that is, the reactions can go in



**Figure 3.1** Simple Transformations. a) A single arrow indicates positive rate direction. b) Two arrows showing explicit reversibility. c) Common barb style used to indicate reversibility. d) Reversibility with dominant arrow indicating positive direction; e) Reversibility using single line with two arrows.

both directions. Unless otherwise stated, reversibility is defined by the rate law attached to the reaction. Sometimes reversibility is explicitly indicated by using multiple arrows. These come in various forms. One approach is to use two lines and add arrowheads to both the reactant and product line as shown in Figure 3.1b. Other authors add a smaller reverse arrow as shown in Figure 3.1d, or more commonly use a barbed style as shown in Figure 3.1c. In example (b), (c), and (e), it is not possible to know which direction represents the positive reaction rate unless it is assumed left to right. SBGN (Systems Biology Graphical Notation) which we will discuss in more detail shortly, uses the style (e) but does not specify the positive direction.

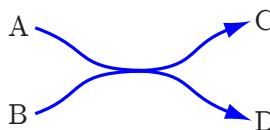
For a bimolecular reaction that depicts dissociation or association, the notation is shown in Figure 3.2 (a) and (b). This style makes it clear that there is a stoichiometric constraint



**Figure 3.2** Dissociation and Association Reactions. (a) Equal stoichiometric proportions of compounds *A* and *B* combine to form a complex, *C*. (b) Likewise, complex *A* dissociates into equal proportions of *B* and *C*.

between *A* and *B* and *B* and *C*. One molecule of *A* reacts with one molecule of *B* to form one molecule of *C*. This notation can be misused, for example, where lines departing from a branch point are joined, thereby implying a stoichiometric constraint when none actually exists.

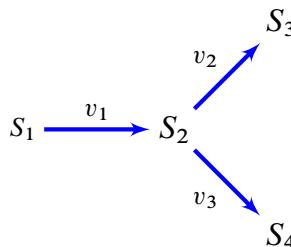
The simple association and dissociation reactions can be naturally extended to depict situations where both association and dissociation occur in the same reaction as shown in Figure 3.3.



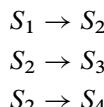
**Figure 3.3** A bimolecular interaction, coupling one process,  $A$  to  $C$ , to another,  $B$  to  $D$ . Equal proportions of  $A$  and  $B$  combine to form equal proportions of  $C$  and  $D$ .

**Example 3.1**

Write the individual reactions for the following network, taking care to indicate the correct stoichiometries.



Answer:

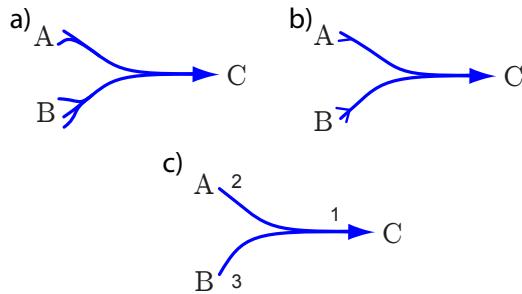


One area that is sometimes problematic is visually depicting reactions with non-unity stoichiometry. The previous examples assumed that each molecular species had a stoichiometry of one. However, what if species  $A$  in Figure 3.2 has a stoichiometry of 2, and  $B$ , a stoichiometry of 3? How should these be represented? Figure 3.4 shows three depictions that have been used by authors in the past. Sometimes simple arc extensions are used to indicate the stoichiometry, as seen in Figure 3.4a. A variation of (a) is to use small barbs at the tips of the reaction arcs [22] where the number of barbs indicate the stoichiometry as seen in Figure 3.4b. Finally, stoichiometric numbers may be placed near the tips of the arcs, as shown in Figure 3.4c.

**Example 3.2**

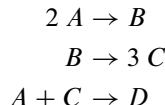
The following network is made from elementary reactions. Write the individual reactions, taking care to indicate the correct stoichiometries.





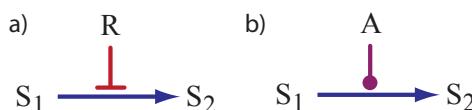
**Figure 3.4** Alternative ways for visually depicting non-unit stoichiometries. The use of numbers in (c) makes it possible to depict fractional stoichiometries.

Answer:



## Non-Elementary Reactions

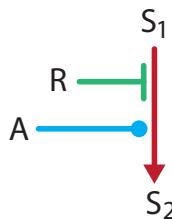
Non-elementary reactions include all reactions that have hidden reaction intermediates. The most familiar is the enzymatic reaction where the enzyme-substrate complex and free enzyme are rarely shown in network diagrams. The effect of hiding intermediates is that it is possible to show regulatory links in a diagram. For example, an enzyme may be regulated by an allosteric effector where the mechanism is quite complex. Very often this mechanism will be hidden and instead, the action of the effector will be represented by a simple regulatory line in the diagram. For example, if an enzyme that catalyzes the conversion of species  $S_1$  to  $S_2$  is inhibited by a repressor molecule  $R$ , or activated by  $A$ , then we can depict this situation as shown in Figure 3.5.



**Figure 3.5** Depicting regulation: a) Repression; b) Activation.

The blunt end representing inhibition is fairly well established in the literature, while the activation symbol is more variable. Here we will employ a filled circle at the end point to

indicate activation. If a non-elementary reaction is regulated by multiple inputs, we would use a depiction similar to what is shown in Figure 3.6.



**Figure 3.6** Multiple regulators on one reaction.

In hiding detailed mechanisms we also invoke certain assumptions when converting the diagrams to a mathematical model. In the case of a simple enzyme mechanism, we will often assume the rapid-equilibrium or steady state assumption for the formation of the enzyme-substrate complex. Sometimes these assumptions are reasonable, other times they are not. For a more comprehensive discussion of these issues see the companion book ‘Enzyme Kinetics for Systems Biology’ [109].

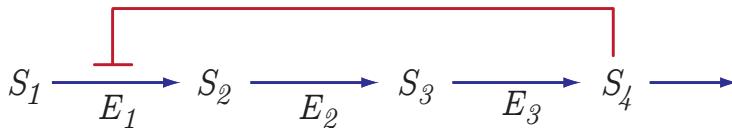
Figure 3.7 shows a common approach to representing a single genetic unit where  $I$  represents a repressor and  $P$  the expressed protein. This single genetic unit is certainly non-elementary as it hides a considerable amount of detail. We can treat the unit as if it were a reaction step whose rate of reaction is the rate of protein expression. It is important to remember that



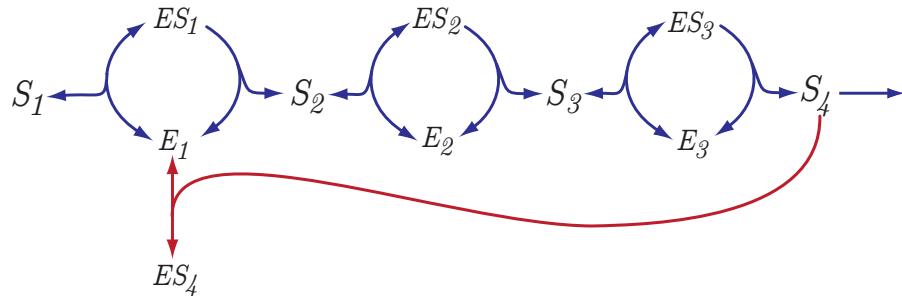
**Figure 3.7** Representing a single gene.  $I$  represents the inducer and  $P$  the expressed protein.

whenever one sees a regulatory link in a reaction step, it always means that the reaction is non-elementary and hides other mechanistic details. The use of non-elementary reactions is a high level representation because unwrapping every non-elementary reaction into its full set of elementary reactions would make the network overly complex to view, understand and parameterize. Figure 3.8 illustrates an example of a simple pathway drawn using non-elementary reactions together with a feedback inhibition step and the equivalent exploded view of the same system. The exploded view is clearly more complex. The mechanism chosen for the inhibition is the simplest possible, and therefore the exploded view could potentially be even more complex. There will be many instances where we will not know how an effector acts mechanistically, therefore obtaining more details on such a reaction is not an option.

A) Network made from Non-elementary Steps.



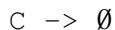
B) Equivalent Network made from Elementary Steps.



**Figure 3.8** Equivalent networks made from non-elementary and elementary components.

## Text Representation

Although representing biochemical networks using pictures is very common, it is also possible to represent networks using a text notation. Text representations are particularly easy for computers to read. For example, a linear chain of four reactions is shown in Figure 3.9. If a species is converted to a waste product such as degradation fragments, then the symbol  $\emptyset$  is typically used to represent the empty species set. For example:

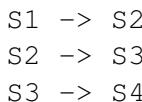


Three software tools that support text based input are Jarnac [108], Antimony [125], and PySCeS [92]. The Python based application called Tellurium<sup>1</sup> integrates Antimony and the simulator libRoadRunner<sup>2</sup> and will be used to illustrate all simulations in this book. There are also rule-based text notations but these are beyond the scope of this book and are supported by tools such as PySB and Bionetgen [9, 36]. See Maus et al. for a review [82].

In this book, models will be expressed in the Antimony syntax [125]. For example, to represent the above model, we would write the script in Antimony as shown in Listing 3.1.

<sup>1</sup>[tellurium.analogmachine.org](http://tellurium.analogmachine.org)

<sup>2</sup>[libroadrunner.org](http://libroadrunner.org)



**Figure 3.9** Simple textual representation of a linear chain of three reactions and four molecular species.

```
// Example Antimony script
$A + B -> 2 C; k1*A*B;
C -> ; k2*C;
```

**Listing 3.1** Example Antimony model.

The \$ sign in front of species A means that the species concentration is fixed. This means that when compiled by a simulator such as libRoadRunner, no differential equations for this species will be generated<sup>3</sup>. Also note that Antimony permits empty reactants or products in a reaction. In this case, the second reaction that consumes C does not specify what C is converted to, only its rate,  $k2*C$ . The implication here is that the products of reaction from C emerge into a large volume such that their concentrations remain approximately unchanged during the process. It also implies that the products do not affect the reaction, also evident in the rate law. The line starting with // is a comment line and is not part of the model.

We can also use Antimony to initialize concentrations and parameters, Listing 3.2.

```
1 // Example Antimony script with value initialization
2 $A + B -> C; k1*A*B;
3 C -> ; k2*C;
4
5 // Initialize values
6 k1 = 0.34; k2 = 4.5;
7 A = 10; B = 0; C = 0;
```

**Listing 3.2** Example Antimony model with initialization.

Antimony offers a host of other features including models composed of other models, and the ability to specify discrete events directly in the model, for example:

```
// Example Antimony script
A + B -> C; k1*A*B;
C -> ; k2*C;

// When time reaches 5 time units, halve the k2 rate constant
```

<sup>3</sup>We will return to the notion by ‘fixed species’ in the next chapter.

```

at (time >= 5) : k2 = k2/2;

k1 = 0.34; k2 = 4.5;
A = 10; B = 0; C = 0;

```

**Listing 3.3** Example Antimony script with events.

See Appendix D for more details on the software.

## 3.2 Standard Visualization Notation

Cellular networks have been depicted on wall charts for many decades using a variety of informal notations which we have briefly reviewed. With the increased interest in protein and gene regulatory networks, the variety of notations has proliferated. As a result, there have been some efforts, most notably the Systems Biology Graphical Notation (SBGN), to define a standard set of node and edge symbols to represent stoichiometric networks. Another visual notation is employed by Biotapestry [79] which provides a concise and easy to read notation for representing gene regulatory networks.

SBGN can represent stoichiometric networks using a notation called SBGN process description. For example, Figure 3.10 illustrates the SBGN approach to representing an enzyme catalyzed reaction. Round shaped nodes or a stadium shape (pill shaped) represent small molecules such as DHAP, ATP and F6P. Rounded rectangles are used to represent macromolecules, in this case enzymes TPase (triose-phosphate Isomerase) and PFK (phosphofructokinase). In the second reaction (Figure 3.10) ATP negatively regulates the reaction.

Full details of this visual specification can be found at the SBGN web site [www.sbgn.org](http://www.sbgn.org); Figure 3.11 summarizes the main symbols.

## 3.3 Mass-Balance Equations

Consider a simple network made up of two reactions,  $v_1$  and  $v_2$ , with a common species,  $S$ .  $v_1$  and  $v_2$  are the rates of reaction such that  $v_1$  is the rate at which  $S$  is produced and in the second reaction,  $v_2$  is the rate at which  $S$  is consumed (Figure 3.12).

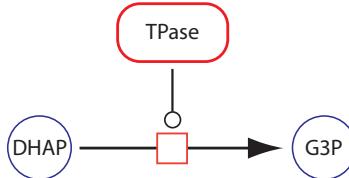
According to the law of conservation of mass, any observed change in the amount of species  $S$  must be due to the difference between the inward rate,  $v_1$ , and outward rate,  $v_2$ . That is, the change in  $S$  is the difference in the two rates, leading to the differential equation:

$$\frac{ds}{dt} = v_1 - v_2 \quad (3.1)$$

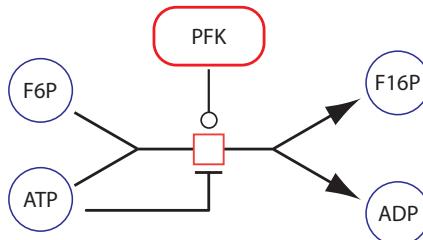
This equation is called a **mass-balance equation**. We can reexpress equation (3.1) as:

$$\frac{dS}{dt} \frac{1}{V} = v_1 - v_2$$

a) Simple uni-uni reaction.



b) More complex reaction with regulation.



**Figure 3.10** SBGN notation for enzyme catalyzed reactions.

where  $S$  is used to designate the amount in moles and  $V$  the volume. Alternatively, we note that:

$$\frac{dS}{dt} = V(v_1 - v_2)$$

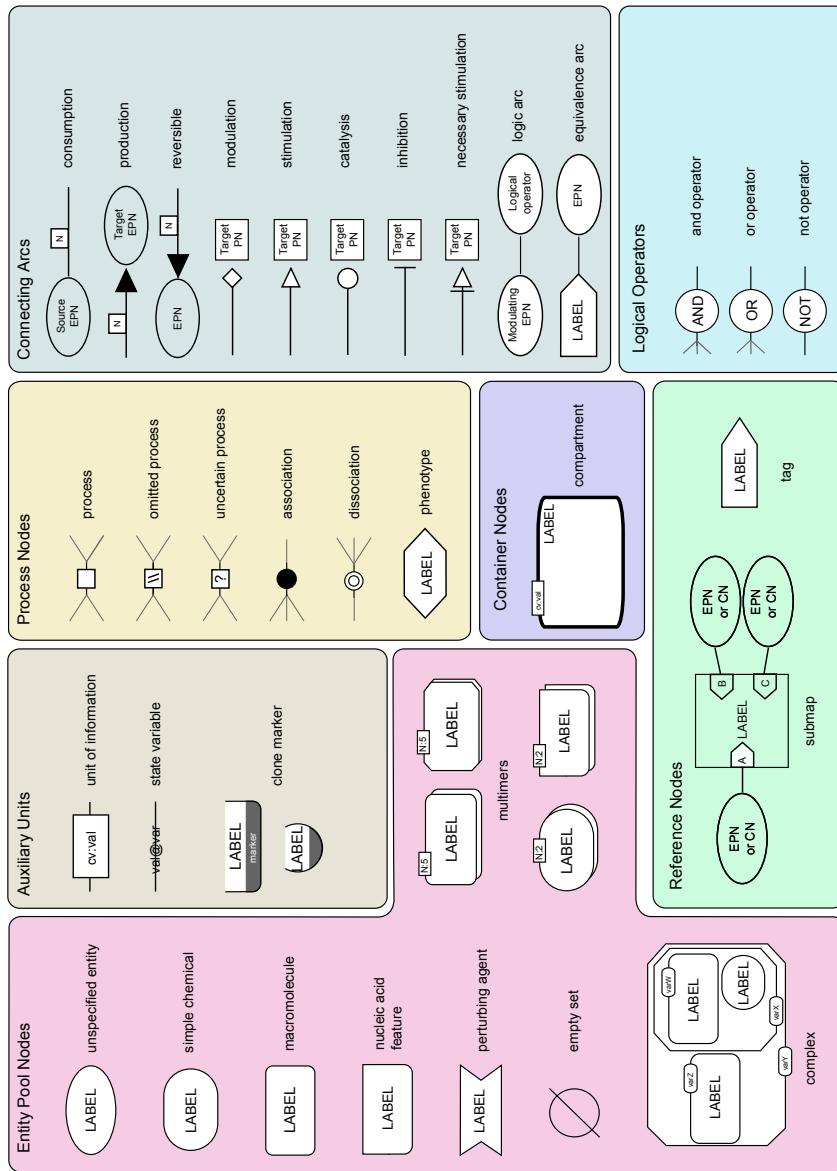
This assumes the reaction rates are expressed in  $\text{mol L}^{-1} \text{ t}^{-1}$ . Biochemical models will sometimes assume a constant unit volume so that numerically:

$$\frac{ds}{dt} = \frac{dS}{dt}$$

Although we will express the rate of change in terms of concentration, it is implied that we are dealing with a constant unit volume so that the change in concentration is the same as the change in amount. It is important to note that it is amounts that are mass conserved, not concentration. For example, if movement is from one compartment to another compartment with a different volume, it is necessary to factor in the volume difference and explicitly express the rate of change in amounts.

Unless otherwise stated, the following assumptions should be made about models in this book:

1. *Well-Stirred Reactor.* Many biochemical models assume that the volume in which reactions take place is well-stirred. This means there are no spatial inhomogeneities. For small cells such as *E. coli*, this is a reasonable assumption. The diffusion rate of molecules



**Figure 3.11** SBGN notation Reference Card, reused but modified with color shading from [www.sbgn.org](http://www.sbgn.org).



**Figure 3.12** Simple two step pathway.

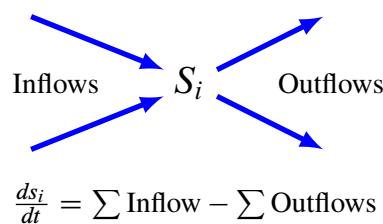
in the cytoplasm is so fast that a given molecule will, on average, sample every location in the *E. coli* cell in one second. In larger eukaryotic cells spatial homogeneity may occur, but it is known that there are mechanisms to restrict diffusion of important molecules from a given location. As such, the assumption of a well-stirred reactor may still apply. Ultimately, the validity of the assumption rests with whether the model generates useful and verifiable predictions.

2. *Large number of molecules.* The number of molecules in cells varied from a few copies for the LacI repressor protein to many millions in the case of ions. When there are large numbers of molecules or ions, concentrations can be approximated using a value that continuously changes. In such cases we can use differential equations to model the rates of change. When dealing with small numbers of molecules however, concentration as a continuous variable may no longer make sense. For example, given that 1 nM roughly equates to one molecule per *E. coli* cell, it doesn't make sense to quote a figure of 1.5 nM since that implies 1.5 molecules per cell. When dealing with small numbers of molecules, it is not possible to have a continuous range of concentrations. Under these circumstances a discrete probabilistic approach is best. Control of systems with small numbers of molecules is beyond the scope of this book and is still an area of active research [68].

3. *Unit Volume.* Unless otherwise indicated, we will assume *fixed unit volumes*.

## Model Networks

For more complex systems such as the one shown in Figure 3.13 where there are multiple inflows and outflows, the mass-balance equation is given by:



**Figure 3.13** Mass Balance: The rate of change in species  $S_i$  is equal to the difference between the sum of the inflows and the sum of the outflows.

$$\frac{ds_i}{dt} = \sum \text{Inflows} - \sum \text{Outflows} \quad (3.2)$$

For an even more general representation, we can reexpress the mass-balance equations by taking into account the stoichiometric coefficients. The rate at which a given reaction,  $v_j$ , contributes to change in a species  $S_i$ , is given by the stoichiometric coefficient of the species,  $S_i$  with respect to the reaction,  $c_{ij}$  multiplied by the reaction rate,  $v_j$ . That is, a reaction  $j$  contributes  $c_{ij} v_j$  rate of change in species  $S_i$ . For example, with the reaction  $A \rightarrow B$  which has a reaction rate  $v$ , and  $c_a$  is  $-1$ , we can say that the reaction contributes  $-1v$  to the rate of change in  $A$ . For a species  $S_i$  with multiple reactions producing and consuming  $S_i$ , the mass-balance equation (assuming constant unit volume) is given by:

$$\frac{ds_i}{dt} = \sum_j c_{ij} v_j \quad (3.3)$$

where  $c_{ij}$  is the stoichiometric coefficient for species  $i$  with respect to reaction,  $j$ . For reactions that consume a species, the stoichiometric coefficient is often *negative*; otherwise the stoichiometric coefficient is *positive*. In considering the simple example in Figure 3.12, the stoichiometric coefficient for  $S$  with respect to  $v_1$  is  $+1$ , and for  $v_2$  is  $-1$ . That is:

$$\frac{ds}{dt} = c_{s1} v_1 + c_{s2} v_2$$

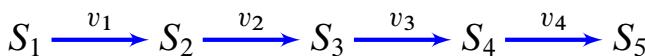
or

$$\frac{ds}{dt} = (+1)v_1 + (-1)v_2 = v_1 - v_2$$

How we describe the construction of the mass-balance equation may seem overly formal, however the formality allows us to write software that can automatically convert network diagrams into mass-balance differential equations.

### **Example 3.3**

Consider a linear chain of reactants from  $S_1$  to  $S_5$  shown in Figure 3.14. Write the mass-balance equations for this simple system.



**Figure 3.14** Simple straight chain pathway.

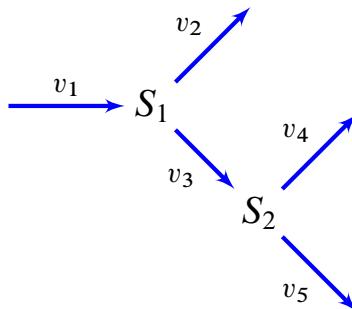
$$\begin{aligned}
 \frac{ds_1}{dt} &= -v_1 & \frac{ds_2}{dt} &= v_1 - v_2 \\
 \frac{ds_3}{dt} &= v_2 - v_3 & \frac{ds_4}{dt} &= v_3 - v_4 \\
 \frac{ds_5}{dt} &= v_4
 \end{aligned} \tag{3.4}$$

Each species in the network is assigned a mass-balance equation which accounts for the flows into and out of the species pool.

---

#### **Example 3.4**

Write the mass-balance equation for the branched system shown in Figure 3.15:



**Figure 3.15** Multi-branched pathway.

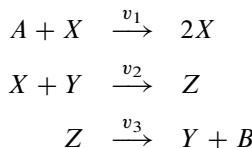
The mass-balance equations are given by:

$$\begin{aligned}
 \frac{ds_1}{dt} &= v_1 - v_2 - v_3 \\
 \frac{ds_2}{dt} &= v_3 - v_4 - v_5
 \end{aligned}$$


---

#### **Example 3.5**

Write the mass-balance equation for the more complex pathway:



This example is more subtle because we need to take into account the stoichiometry change between the reactant and product side in the first reaction ( $v_1$ ). In reaction  $v_1$ , the stoichiometric coefficient

for  $X$  is  $+1$  because two  $X$  molecules are made for every one consumed. Taking this into account, the rate of change of species  $X$  can be written as:

$$\frac{dx}{dt} = -v_1 + 2v_1 - v_2$$

or more simply as  $v_1 - v_2$ . The full set of mass-balance equations can therefore be written as:

$$\begin{aligned}\frac{da}{dt} &= -v_1 & \frac{dx}{dt} &= v_1 - v_2 \\ \frac{dy}{dt} &= v_3 - v_2 & \frac{dz}{dt} &= v_2 - v_3 \\ \frac{db}{dt} &= v_3\end{aligned}$$

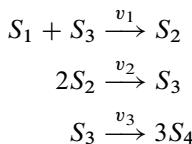
The last example (3.5) illustrates a very important aspect of converting a network diagram into a set of differential equations. The process is potentially **lossy**. That is, it is *not always possible* to fully recover the original network diagram from the set of derived differential equations. This is because in one or more of the reactions, the stoichiometries may cancel. In example (3.5) the reaction  $A + X \longrightarrow 2X$  is not recoverable from the final set of differential equations. Instead, if we reverse engineered the differential equations, the first reaction would be:



which is not like the original. This is not a common occurrence although in protein signaling pathways it is more common than other kinds of networks. What it means however is that sharing models by exchanging differential equations is not recommended. This is one reason why standard exchange formats such as SBML [59] store models explicitly as a set of reactions, not as a set of differential equations. Many models are exchanged using Matlab which means that much of the biological information, particularly information on the underlining network, is lost. *Exchanging models* via computer languages such as Matlab is therefore not recommended.

### **Example 3.6**

Write the mass-balance equation for pathway:



In this example we have non-unity stoichiometries in the second and third reaction steps. The

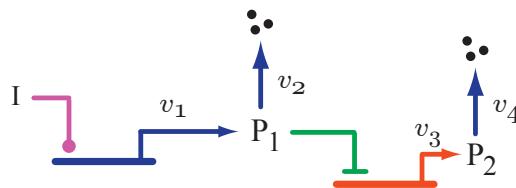
mass-balance equations are given by:

$$\begin{aligned}\frac{ds_1}{dt} &= -v_1 & \frac{ds_2}{dt} &= v_1 - 2v_2 \\ \frac{ds_3}{dt} &= v_2 - v_3 - v_1 & \frac{ds_4}{dt} &= 3v_3\end{aligned}$$


---

### Example 3.7

Write the mass-balance equations for  $P_1$  and  $P_2$  for the following gene regulatory network:



The key to this problem is that the network diagram suggests that the regulation from  $P_1$  to  $v_3$  results in no consumption of  $P_1$ .  $P_1$  acts only as a regulator. That being the case, the two mass-balance equations are:

$$\frac{dp_1}{dt} = v_1 - v_2$$

$$\frac{dp_2}{dt} = v_3 - v_4$$


---

From the previous examples we see that it is fairly straightforward to derive the mass-balance equations from a visual inspection of the network. Many software tools exist to assist in this effort by converting network diagrams, either represented visually on a computer screen (for example, pathwayDesigner<sup>4</sup>), or by processing a text file that lists the reactions in the network (for example via Antimony) into a set of differential equations (See Appendix D).

## 3.4 Stoichiometry Matrix

When describing multiple reactions in a network, it is convenient to represent the stoichiometries in a compact form called the **stoichiometry matrix**. Traditionally the matrix is denoted by **N**, where the symbol N refers to ‘number’<sup>5</sup>. The stoichiometry matrix is a  $m$

<sup>4</sup>pathwaydesigner.org

<sup>5</sup>Some recent flux balance literature uses the symbol S; the traditional symbol N will be used here to avoid confusion with S for species.

row by  $n$  column matrix, where  $m$  is the number of species and  $n$  the number of reactions:

$$\mathbf{N} = m \times n \text{ matrix}$$

The columns of the stoichiometry matrix correspond to the individual chemical reactions in the network. The rows correspond to the molecular species, with one row per species. Thus, the intersection of a row and column in the matrix indicates whether a certain species takes part in a particular reaction or not. The sign of the element determines whether there is a net loss or gain of substance, and the magnitude describes the relative quantity of substance taking part in the reaction.

The elements of the stoichiometry matrix *do not* concern themselves with the rate of reaction. This latter point is particularly important because various stoichiometric analyses can be carried out purely on the stoichiometry without *any* reference to reaction rate laws.

The stoichiometric matrix is not concerned with describing reaction rates. Reaction rates are given by rate laws in a separate vector (See section 3.9).

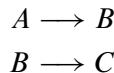
In general, the stoichiometry matrix has the form:

$$\mathbf{N} = \begin{matrix} & \xleftarrow{\quad} & v_j & \xrightarrow{\quad} \\ S_i & \uparrow & \left[ \begin{matrix} c_{ij} & \dots & \dots \\ \vdots & & \end{matrix} \right] \\ & \downarrow & \end{matrix}$$

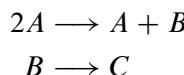
where  $c_{ij}$  is the stoichiometry coefficient for the  $i^{\text{th}}$  species and  $j^{\text{th}}$  reaction. As mentioned before, the stoichiometry matrix is generally a lossy representation. That is, it is not always possible to revert back to the original biochemical network from which the matrix was derived. For example, consider the simple stoichiometry matrix:

$$\mathbf{N} = \begin{bmatrix} -1 & 0 \\ 1 & -1 \\ 0 & 1 \end{bmatrix}$$

The most obvious network that this matrix could have been derived from is:



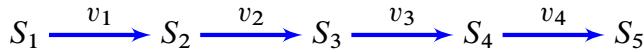
But an equally plausible network is:



It is not possible from the stoichiometry matrix alone to determine the original network.

**Example 3.8**

Write the stoichiometry matrix for the simple chain of reactions which has five molecular species and four reactions as shown below. The four reactions are labeled,  $v_1$  to  $v_4$ .



The stoichiometry matrix for this simple system is given by:

$$\mathbf{N} = \begin{bmatrix} v_1 & v_2 & v_3 & v_4 \\ -1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{array}{l} S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \end{array}$$

The rows and columns of the matrix have been labeled for convenience. Normally labels are absent.

**Example 3.9**

Write the stoichiometry matrix for the multibranched pathway shown in Figure 11.5.

$$\mathbf{N} = \begin{bmatrix} v_1 & v_2 & v_3 & v_4 & v_5 \\ 1 & -1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 & -1 \end{bmatrix} \begin{array}{l} S_1 \\ S_2 \end{array}$$

## 3.5 Reversibility

Up to this point we have not discussed whether a given reaction is reversible or not. When dealing with kinetic models, reversibility often manifests itself as a negative reaction rate in the rate law. For example, the rate law for the simple mass-action reversible reaction  $A \rightleftharpoons B$  is given by:

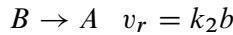
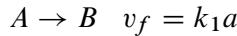
$$v = k_1 a - k_2 b$$

When this reaction goes in the reverse (right to left) direction, the reaction rate,  $v$ , will be negative. This may not be apparent from the stoichiometry matrix, which in this case is:

$$\mathbf{N} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Information on reversibility is therefore traditionally found in the rate law. In this example the rate law could equally have been  $k_1 A$ , suggesting an irreversible reaction. Depending on

the modeling problem, reversibility can be made more explicit in the stoichiometry matrix by specifying a *separate* reaction path for the reverse reaction. For example, in the previous example we might instead represent the system by two separate rate laws:



The stoichiometry matrix now becomes:

$$\mathbf{N} = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$$

Splitting a reaction into separate forward and reverse steps might not always be possible however. For example, an enzyme catalyzed reversible reaction such as  $A \rightleftharpoons B$  *cannot* be represented using:

$$\frac{db}{dt} = v_f - v_r$$

where  $v_f$  is the forward rate and  $v_r$  the reverse rate. At first glance we might choose to model the forward and reverse rates using irreversible Michaelis-Menten rate laws (2.2). However, the forward and reverse reactions are not independent. They are connected by the shared free enzyme pool so that when the forward rate rises, the reverse rate falls due to competition for free enzyme. If the modeler insists on separating the forward from the reverse rate, then the full enzyme mechanism in terms of elementary steps must be used (See the companion text book Enzyme Kinetics for Systems Biology for more details [109]). Alternatively, and more commonly, a reversible enzyme catalyzed reaction is expressed using the reversible Michaelis-Menten equation (2.4):

$$v = \frac{V_f / K_s (s - p / K_{eq})}{1 + s / K_s + p / K_p}$$

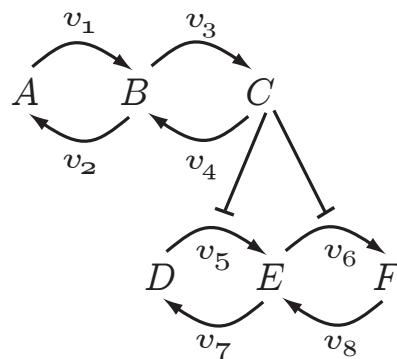
where  $S$  and  $P$  are the substrate and product concentrations, respectively.  $V_f$  is the maximal forward rate. Some modelers will choose to express all reactions using elementary reactions but this poses its own problems, particularly when trying to set values for the many elementary rate constants that result. Ultimately, the decision has to be made on a case by case basis and will depend on the models purpose.

Before leaving the topic of reversibility, it is worth repeating again the discussion we had in the last chapter on **product inhibition**. Product inhibition occurs when the product binds to an enzyme without resulting in any reverse reaction rate. However, binding of product competes with substrate which in turn represses the forward rate. Reactions that are often considered irreversible can therefore still be affected by product.

To illustrate how we apply the stoichiometry matrix to different kinds of networks, let's look at a simple signaling network and two simple gene regulatory networks.

## 3.6 Signaling Networks

Figure 3.16 illustrates a simple protein signaling network comprised of two double phosphorylation cycles coupled through activation by protein  $C$  on the lower double cycle ( $D, E$  and  $F$ ). In this model all species are proteins and we assume that protein  $A$  and  $D$  are unphosphorylated,  $B$  and  $E$  singly phosphorylated, and  $C$  and  $F$  doubly phosphorylated.  $C$  acts as a kinase and phosphorylates  $D$  and  $E$ . The reverse reactions,  $v_2, v_4, v_7$  and  $v_8$  are assumed to be catalyzed by phosphatases.



**Figure 3.16** Simple signaling network. Protein  $C$  activates the activity of reactions  $v_5$  and  $v_6$ .

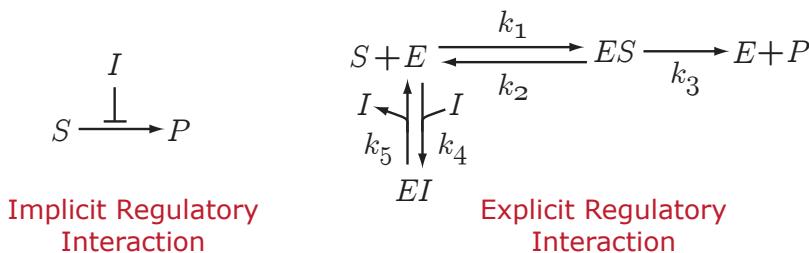
There is no specified stoichiometric mechanism for the activation on  $v_5$  and  $v_6$ . Therefore, the stoichiometric matrix will contain no information about this. The stoichiometric matrix for this system is:

$$\mathbf{N} = \begin{matrix} & v_1 & v_2 & v_3 & v_4 & v_5 & v_6 & v_7 & v_8 \\ A & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ B & 1 & -1 & -1 & 1 & 0 & 0 & 0 & 0 \\ C & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ D & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 \\ E & 0 & 0 & 0 & 0 & 1 & -1 & -1 & 1 \\ F & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{matrix} \quad (3.5)$$

The matrix is composed of two separate blocks corresponding to the two cycle layers. It is important to emphasize again that whenever there are *regulatory* interactions in a pathway diagram, these *do not* appear in the stoichiometry matrix. Instead, such information will reside in the rate laws that describe the regulation. If however the mechanism for the regulation is made explicit, then details of the regulation will appear in the stoichiometry matrix. Figure 3.17 shows a simple example of an inhibitor,  $I$ , regulating a reaction,  $S$

to  $P$ . The left displays an implicit regulatory interaction. All we see is a blunt ended arrow indicating inhibition. In this case details of the regulation will be found in the rate law governing the conversion of  $S$  to  $P$ . On the right is an explicit mechanism, a simple competitive inhibition. In this case details of the inhibition mechanism will find its way into the stoichiometry matrix, although from an inspection of the matrix, the type of regulation may not be obvious.

Figure 3.18 shows a comparison of the implicit and explicit models in terms of the stoichiometry matrix. In each case the rate laws also change. In the implicit form, the rate law will be a Michaelis-Menten competitive inhibition model whereas in the explicit model, the rates laws (now multiplied in number) will be simple mass-action rate laws. The choice of what to use, an implicit or explicit model, will depend entirely on the type of question that the model is attempting to answer. *There is no right or wrong way to do this*, the details of a model will depend on the type of question being asked.



**Figure 3.17** Example of implicit and explicit depiction of a regulatory interaction. The left-hand mechanism involving inhibitor,  $I$ , will not appear in the stoichiometry matrix whereas in the explicit mechanism, it will.

$$\mathbf{N} = \begin{matrix} S & \left[ \begin{matrix} v_1 \\ -1 \\ 1 \\ 0 \end{matrix} \right] \\ P & \\ I & \end{matrix} \quad \text{Implicit}$$

$$\mathbf{N} = \begin{matrix} S & \left[ \begin{matrix} v_1 & v_2 & v_3 & v_4 & v_5 \\ -1 & 1 & 0 & -1 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 \\ -1 & 1 & 1 & -1 & 1 \\ 1 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 \end{matrix} \right] \\ P & \\ I & \\ E & \\ ES & \\ EI & \end{matrix} \quad \text{Explicit}$$

**Figure 3.18** Stoichiometry matrices corresponding to the two models in Figure 3.17.

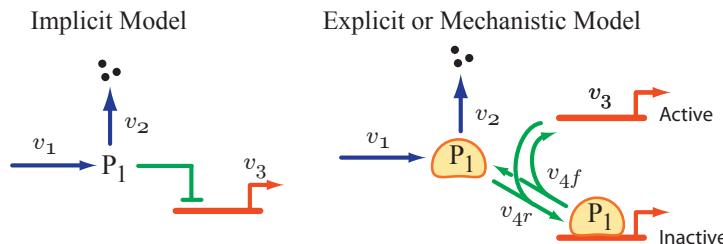
## 3.7 Gene Regulatory Networks

Consider a transcription factor  $P_1$  that represses a gene with expression rate  $v_3$  shown in Figure 3.19, left panel. In this model we have production of  $P_1$  from reaction  $v_1$ , and degradation of  $P_1$  via  $v_2$ . The construction of the stoichiometry matrix will depend on how we represent the regulated step,  $v_3$ . If regulation is implied, meaning there is no explicit kinetic mechanism, then the regulation will not appear in the stoichiometry matrix. For the network on the left in Figure 3.19, the stoichiometry matrix is:

$$\mathbf{N} = P_1 \begin{bmatrix} v_1 & v_2 \\ 1 & -1 \end{bmatrix} \quad (3.6)$$

The stoichiometry matrix has only one row indicating that there is only one species in the model,  $P_1$ , and there is no hint in the stoichiometry matrix of any regulation. In this model  $P_1$  is not explicitly sequestered by the operator site upstream of the gene. We make the significant assumption that when  $P_1$  regulates, its own state is not affected in any way.

Consider now that the interaction between  $P_1$  and  $v_3$  is made mechanistically explicit. The right-hand network in Figure 3.19 shows one possible way in which to represent the interaction of the transcription factor,  $P_1$  with gene  $v_3$ . In the explicit model the transcription factor  $P_1$  is assumed to bind to a repressor site preventing gene expression. In the explicit model



**Figure 3.19** Two simple gene regulatory networks involving gene repression. On the left side is the implicit model where  $P_1$  represses  $v_3$ , on the right side is the explicit model showing a more detailed mechanism for the regulation.

there are two new species, designated active gene and inactive gene. The stoichiometry matrix will therefore include two additional rows corresponding to these two new species. The stoichiometry matrix for the explicit model is shown here:

$$\mathbf{N} = \begin{array}{c} P_1 \\ P_1(\text{Active}) \\ P_1(\text{InActive}) \end{array} \begin{bmatrix} v_1 & v_2 & v_{4r} & v_{4f} \\ 1 & -1 & -1 & 1 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 1 & -1 \end{bmatrix} \quad (3.7)$$

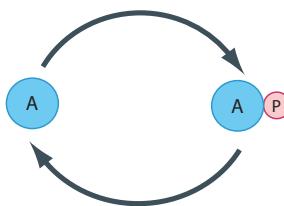
In this case  $P_1$  is actively sequestered onto the operator site and therefore appears in the stoichiometry matrix. Processes such as consumption, production, or sequestration by some binding mechanism will appear as columns in the stoichiometry matrix.

In conclusion, regulation does not appear explicitly in a stoichiometry matrix unless the regulation is represented as an explicit mechanistic scheme. The choice of implicit or explicit representations depends on the question being asked and the availability of suitable data.

### 3.8 Moiety Conserved Cycles

Many cell processes operate on different time scales. For example, metabolic processes tend to operate on a faster scale than protein synthesis and degradation. Such time scale differences have a number of implications to model builders, software designers, and model behavior. In this chapter we will briefly examine some of these aspects in relation to species conservation laws.

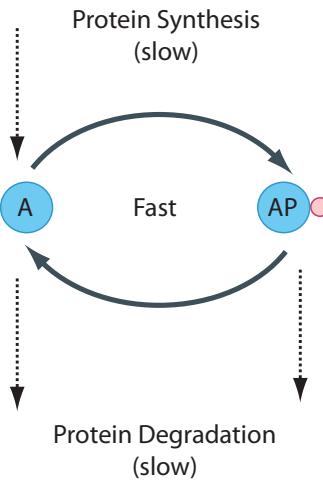
To introduce this topic, consider a simple protein phosphorylation cycle such as the one shown in Figure 3.20. This shows a protein undergoing phosphorylation (upper limb) and dephosphorylation (lower limb) via a kinase and phosphatase, respectively.



**Figure 3.20** Phosphorylation and dephosphorylation cycle forming a moiety conservation cycle between unphosphorylated (left species,  $A$ ) and phosphorylated protein (right species,  $AP$ ).

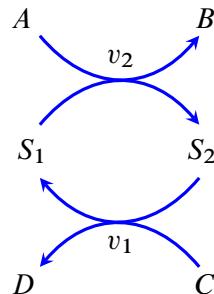
The depiction in Figure 3.20 is a simplification. The ATP used during phosphorylation and the release of free phosphate during the dephosphorylation event are not shown. In addition, synthesis and degradation of protein is also absent. In many cases we can leave these aspects out of the picture. ATP for instance is held at a relatively constant level by strong homeostatic forces from metabolism so that within the context of the cycle, changes in ATP isn't something we must worry about. More interesting is that within the time scale of phosphorylation and dephosphorylation, we can assume that the rate of protein synthesis and degradation is negligible (Figure 3.21). This assumption is more significant and leads to the emergence of a new property called **moiety conservation** [103].

In chemistry a **moiety** is described as a subgroup of a larger molecule. In this case the moiety is a protein. During the interconversion between the phosphorylated and unphosphorylated



**Figure 3.21** Phosphorylation and dephosphorylation cycle that also includes the slower process of protein synthesis and degradation. We assume that the phosphorylated and unphosphorylated protein can be degraded, but only the unphosphorylated protein is synthesized.

states, the amount of moiety (protein) remains constant. More abstractly we can draw a cycle in the following way (Figure 3.22), where  $S_1$  and  $S_2$  are the cycle species:



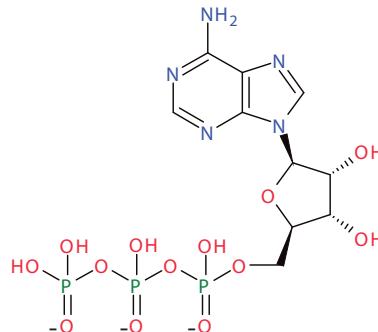
**Figure 3.22** Simple conserved cycle where  $s_1 + s_2 = \text{constant}$ .

The two species  $S_1$  and  $S_2$  are conserved because the total  $s_1 + s_2$  remains constant over time (at least over a time scale shorter than other processes that may be involved). Such cycles are collectively called **moiety conserved cycles**.

**Moiety:** A subgroup of a larger molecule.

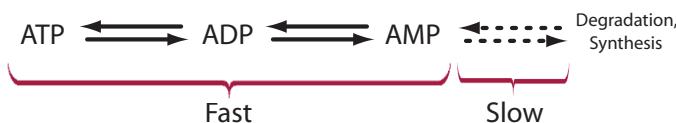
**Conserved Moiety:** A subgroup whose interconversion through a sequence of reactions leaves it unchanged.

Protein signaling pathways abound with conserved cycles such as these, although many are more complex and may involve multiple phosphorylation reactions. In addition to protein networks, other pathways also possess conservation cycles. One of the earliest conservation cycles to be recognized was the adenosine triphosphate (ATP) cycle. ATP is a chain of three phosphate residues linked to a nucleoside adenosine group as shown in Figure 3.23.



**Figure 3.23** Adenosine Triphosphate: Three phosphate groups plus an adenosine subgroup.

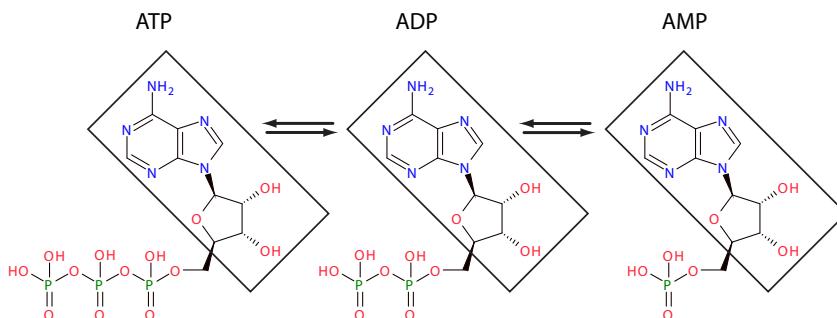
The linkage between the phosphate groups involves an unstable phosphoric acid anhydride bond. These bonds can be cleaved by hydrolysis one at a time leading to the formation of adenosine diphosphate (ADP) and adenosine monophosphate (AMP), respectively. The hydrolysis provides much of the free energy to drive endergonic processes in the cell. Given the insatiable need for energy, there is a continual and rapid interconversion between ATP, ADP and AMP as energy is released or captured. One constant during these interconversions is the amount of adenosine group (Figure 3.25). Adenosine is a conserved moiety. Over longer time scales there is also the slower process of AMP degradation and biosynthesis via the purine nucleotide pathway; but for many models, we assume that this process is very slow compared to ATP turnover by energy metabolism.



**Figure 3.24** The interconversion of ATP, ADP and AMP is generally considered fast in comparison to the slow process of synthesis and degradation of AMP.

There are many other examples of conserved moieties such enzyme/enzyme-substrate com-

plexes, NAD/NADH, phosphate and coenzyme A. In all these cases the basic assumption is that the interconversions of the subgroups is rapid compared to their net synthesis and degradation. We should emphasize that in reality, conserved moieties do not exist since all molecular subgroups will at some point be subject to synthesis and degradation. However, over sufficiently short time scales, the sum total of these groups can be considered constant.



**Figure 3.25** The adenosine moiety, indicated by the boxed molecular group, is conserved during the interconversion of ATP, ADP and AMP.

## 3.9 The System Equation

Equation (3.3), which describes the mass-balance equation, can be reexpressed in terms of the stoichiometry matrix to form the **system equation**:

$$\frac{ds}{dt} = \mathbf{N}v \quad (3.8)$$

where  $\mathbf{N}$  is the  $m \times n$  stoichiometry matrix and  $v$  is the  $n$  dimensional rate vector, whose  $i$ th component gives the rate of reaction  $i$  as a function of the species concentrations.  $s$  is the  $m$  vector of species. This is a key equation for describing a network of processes inside a cell. Of particular significance is that the equation explicitly separates the network, in the form of  $\mathbf{N}$ , from the process rates,  $v$ .

Looking again at the simple chain of reactions in Figure 3.14, the system equation can be written as:

$$\frac{ds}{dt} = \mathbf{N}v = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix} \quad (3.9)$$

If the stoichiometry matrix is multiplied into the rate vector, the mass-balance equations shown earlier (3.4) are recovered. To illustrate what the system equation might look like for a simple system, consider the following model expressed in Antimony format:

```
A -> B; k1*A - k2*B;
B -> C; k3*B - k4*C;
```

The system equation for this model is:

$$\frac{ds}{dt} = \mathbf{N}v = \begin{bmatrix} -1 & 0 \\ 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} k_1A - k_2B \\ k_3B - k_4C \end{bmatrix} \quad (3.10)$$

## 3.10 Tellurium

The modeling platform Tellurium [108] provides facilities to extract the stoichiometry matrix from a model. The command for generating the stoichiometry matrix is `getFullStoichiometryMatrix()`. The short-hand version for this command is `sm`. The script and results of a run are given below:

```
import tellurium as te

r = te.loada '''
J1: A -> B; k1*A - k2*B;
J2: B -> C; k3*B - k4*C;

k1 = 0.1; k2 = 0.02;
k3 = 0.3; k4 = 0.04;
A = 10; B = 0; C = 0;
''')

print r.getFullStoichiometryMatrix()
```

If this script is run, the output is:

```
J1, J2
A [[ -1,  0],
B [ 1, -1],
C [ 0,  1]]
```

If you need the explicit names for the rows and columns, use the call for the column labels:

```
rr.getReactionIds()
```

and for the row labels:

```
rr.getFloatingSpeciesIds()
```

Using the supported shortcuts (See appendix), the script becomes:

```
import tellurium as te

r = te.loada '''
J1: A -> B; k1*A - k2*B;
J2: B -> C; k3*B - k4*C;

k1 = 0.1; k2 = 0.02;
k3 = 0.3; k4 = 0.04;
A = 10; B = 0; C = 0;
'''

print r.sm()
print r.rs()
print r.fs()
```

---

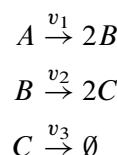
## Further Reading

1. Palsson BO (2006) Systems Biology Systems Biology: Properties of Reconstructed Networks. Cambridge University Press, ISBN: 978-0521859035
2. Sauro HM (2012) Enzyme Kinetics for Systems Biology. 2nd Edition, Ambrosius Publishing ISBN: 978-0982477335
3. Stephanopoulos G, Aristidou A, and Nielsen J (1998) Metabolic engineering: principles and methodologies. Academic Press, ISBN: 978-0126662603

---

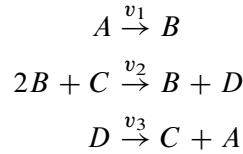
## Exercises

1. Derive a set of differential equations for the following model in terms of the rate of reaction,  $v_1$ ,  $v_2$ , and  $v_3$ :



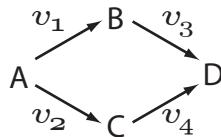
2. Derive the set of differential equations for the following model in terms of the rate of

reaction,  $v_1$ ,  $v_2$  and  $v_3$ :

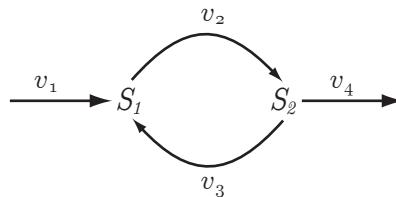


3. Write the stoichiometry matrix for the networks in question 3 and 4.
4. Enter the previous models, 3 and 4, into Tellurium and confirm that the stoichiometry matrices are the same as those derived manually in the previous question.
5. Derive the stoichiometry matrix for each of the following networks. In addition, write the mass-balance equations in each case.

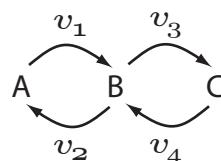
(a)



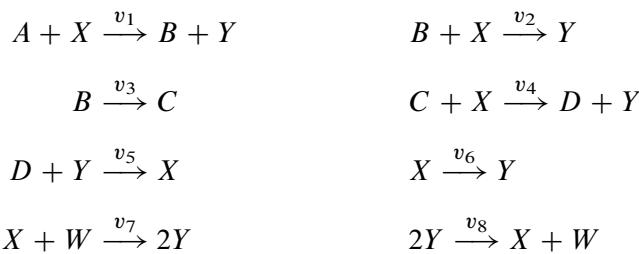
(b)



(c)



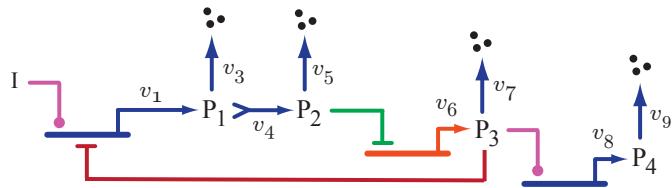
(d)



6. For the irreversible enzyme catalyzed reaction,  $A \rightarrow B$ :
- Write the stoichiometry matrix.
  - Write the stoichiometry matrix in terms of the elementary reactions that make up the enzyme mechanism.
7. A gene  $G_1$  expresses a protein  $p_1$  at a rate  $v_1$ .  $p_1$  forms a tetramer (4 subunits), called  $p_1^4$  at a rate  $v_2$ . The tetramer negatively regulates a gene  $G_2$ .  $p_1$  degrades at a rate  $v_3$ .  $G_2$  expresses a protein,  $p_2$  at a rate  $v_9$ .  $p_2$  is cleaved by an enzyme at a rate  $v_4$  to form two protein domains,  $p_2^1$  and  $p_2^2$ .  $p_2^1$  degrades at a rate  $v_5$ . Gene  $G_3$  expresses a protein,  $p_3$  at a rate  $v_6$ .  $p_3$  binds to  $p_2^2$  forming an active complex,  $p_4$  at a rate  $v_{10}$ , which can bind to gene  $G_1$  and activate  $G_1$ .  $p_4$  degrades at a rate  $v_7$ . Finally,  $p_2^1$  can form a dead-end complex,  $p_5$ , with  $p_4$  at a rate  $v_8$ .
- Draw the network represented in the description given above.
  - Write the differential equation for each protein species in the network in terms of  $v_1, v_2, \dots$
  - Write the stoichiometric matrix for the network.
8. Write the differential equations for the system depicted in equation (3.9).
9. Given the following stoichiometry matrix, write out the corresponding network diagram. Why might this process not fully recover the original network from which the stoichiometry matrix was derived?

$$\begin{array}{c}
 \begin{matrix}
 & v_1 & v_2 & v_3 & v_4 & v_5 \\
 A & -1 & 0 & -1 & 0 & 0 \\
 B & 1 & -1 & 0 & 0 & 3 \\
 C & 0 & 2 & -1 & 0 & 0 \\
 D & 0 & 0 & 1 & -1 & 0 \\
 E & 0 & 0 & 0 & 1 & -1 \\
 F & 0 & 0 & 0 & 0 & 1 \\
 G & 0 & 0 & 0 & -1 & 0
 \end{matrix} \\
 (3.11)
 \end{array}$$

10. Derive the mass-balance equations for the following gene regulatory network:



11. Why is it better to store a model as a list of reactions rather than a set of differential equations?

# 4

## *How Systems Behave*

---

### 4.1 System Behavior

---

As we proceed through the book we will encounter many different kinds of behavior. As such, it is worth describing the states that are fundamental to all systems. These states fall into three groups: **(Thermodynamic) equilibrium, steady state, and transients**. In the literature the terms equilibrium and steady state are often used to mean the same thing, but here they will be used to describe two very different states.

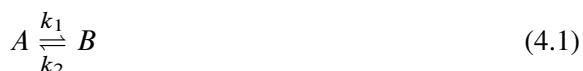
The simplest and arguably the least interesting is equilibrium, or more precisely thermodynamic equilibrium.

---

### 4.2 Equilibrium

---

Thermodynamic equilibrium, or simply equilibrium, refers to the state of a system when all forces are balanced. In chemistry, thermodynamic equilibrium is when all forward and reverse rates are equal. This also means that the concentration of chemical species are unchanging and all net flows are zero. Equilibrium is easily achieved in an isolated system. For example, consider the simple chemical isomerization:



Let the net forward rate of the reaction,  $v$ , be equal to  $v = k_1a - k_2b$ . The rates of change

of  $A$  and  $B$  are given by:

$$\frac{da}{dt} = -v \quad \frac{db}{dt} = v$$

At equilibrium  $da/dt$  and  $db/dt$  equal zero, that is  $k_1a = k_2b$ , or  $v = 0$ . The analytical solution to the chemical isomerization can be derived as follows. Given that the system is isolated, we know that the total mass in the system,  $a + b$  is constant. This constant is given by the sum of the initial concentrations of  $A$  and  $B$  which we will define as  $a_o + b_o$ . Note that  $a_o + b_o = a(t) + b(t)$  is always true. We assume that the volume is constant and set to unit volume, which allows us to state that the sum of the concentrations is conserved. The differential equation for  $A$  is given by:

$$\frac{da}{dt} = k_2b - k_1a \quad (4.2)$$

Before solving this equation, let us replace  $b$  by the term  $a_o + b_o - a$ . This yields:

$$\frac{da}{dt} = k_2a_o + k_2b_o - k_2a - k_1a = k_2(a_o + b_o) - a(k_1 + k_2) \quad (4.3)$$

The easiest way to solve this equation is to use Mathematica or Maxima. The Mathematica command is `DSolve[{ a'[t] == k2 (ao + bo) - a[t] (k1 + k2), a[0] == ao}, a[t], t]`, where  $a[0] == ao$  sets the initial condition for the concentration of  $A$  to be  $a_o$ . By implication, the initial condition for  $b_o$  is  $(a_o + b_o) - a_o = b_o$ . It is also not difficult to derive the solution by hand if we use the standard solution for a linear differential equation:  $dx/dt = ax + b$  which is  $x = Ce^{at} - b/a$ . We need to specify some initial conditions in order to determine the constant  $C$  and here we assume that at  $t = 0$ ,  $a = a_o$ . Using either method results in the following solution:<sup>1</sup>

$$a(t) = a_o e^{-(k_1+k_2)t} + \frac{(a_o + b_o)k_2}{k_1 + k_2} \left( 1 - e^{-(k_1+k_2)t} \right) \quad (4.4)$$

It easier to see the structure if we rearrange the solution to:

$$a(t) = \frac{(a_o + b_o)k_2}{k_1 + k_2} + e^{-(k_1+k_2)t} \left( a_o - \frac{(a_o + b_o)k_2}{k_1 + k_2} \right)$$

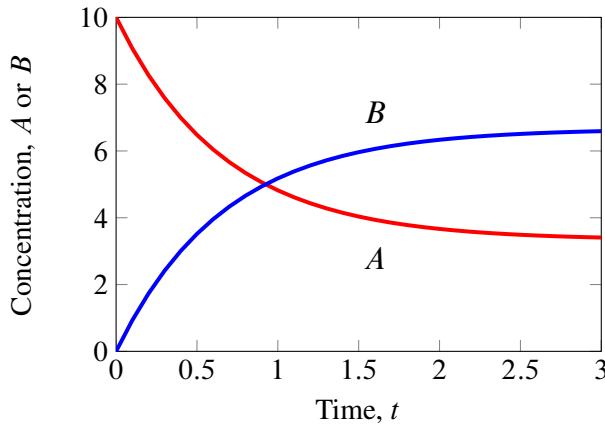
The first term in the equation is a constant and equals the equilibrium concentration of  $A$ . The term is also a function of the total mass in the system ( $a_o + b_o$ ), which means that the equilibrium solution is independent of the starting concentrations so long as the total remains the same. The second term is time dependent and describes the evolution of the system when the initial concentrations of  $A$  and  $B$  are not set to the equilibrium concentrations.

At equilibrium the reaction rate can be computed by substituting the equilibrium concentration of  $A$  and  $B$  into the reaction rate,  $v = k_2b - k_1a$ . We note that the equilibrium concentration of  $A$  is given by:

$$a_{eq} = \frac{(a_o + b_o)k_2}{k_1 + k_2}$$

---

<sup>1</sup>For the curious the differential equations are solved using Laplace Transforms in the Chapter Appendix.



**Figure 4.1** Time course for equilibration of the reversible reaction in model (4.1) where  $k_1 = 1$ ,  $k_2 = 0.5$ ,  $a_o = 10$ ,  $b_o = 0$ . The ratio of the equilibrium concentration is given by  $k_1/k_2$ . Tellurium listing of model: (4.1)

and for  $B$  by subtracting  $a_{eq}$  from  $a_o + b_o$ . When the  $a_{eq}$  and  $b_{eq}$  relations are substituted into  $v$ , the result is:

$$v = 0$$

From this somewhat long winded analysis, it has been determined for the isolated reversible system, at infinite time, the concentrations of  $A$  and  $B$  reach some constant values and that the net rate,  $v$  is zero. The system is therefore at thermodynamic equilibrium.

In biochemical models it is often assumed that when the forward and reverse rates for a particular reaction are very fast compared to the surrounding reactions, the reaction is in **quasi-equilibrium**. That is, although the entire system may be out of equilibrium, there may be parts of the system that can be approximated as though they were in equilibrium. This is often done to simplify the modeling process. Living organisms are not themselves at thermodynamic equilibrium; if they were, then they would technically be dead. Living systems are open so that there is a continual flow of mass and energy across the system's boundaries.

## 4.3 Boundary and Floating Species

In not isolated systems there are boundaries between the system and the surrounding environment. The surrounding environment provides the sources and sinks for any chemical species the system depends on. These sources are often called **boundary species** and ensure that the system does not tend to thermodynamic equilibrium. The source species supply mass to the system and the sink species provide an exit point of mass from the system. In a real system, for example a yeast culture, glucose will often be the source species and the waste or sink species ethanol. In a computer model these species that have fixed concentrations.

In a typical experiment, boundary species are clamped to some fixed values. The clamping mechanism can simply be a large external reservoir so that any exchange of mass between the system and the external reservoir has a negligible effect on the external concentration. Alternatively, there may be active mechanisms maintaining an external concentration such as a chemostat. External concentrations may also change slowly in time compared to the timescale of the experiment so that over the study period, the external concentrations change very little. A typical example is the study of a metabolic response over a timescale that is shorter than change in gene expression. This permits an experimenter to study a metabolic pathway without considering the effect of changes in gene expression.

The same tactic is used when constructing a computational model. In the case of a model it is much simpler to achieve as it only requires the software to keep the boundary species fixed during a simulation.

Those species that changed during a simulation or experiment are called **floating species**.

## 4.4 Steady State

---

The **steady state**, also called the stationary state, is where the rates of change of all species,  $da/dt$  are zero, **but** at the same time the net rates are non-zero, that is  $v_i \neq 0$ . This situation can only occur in an open system; that is a system that can exchange matter with the surroundings. To convert the simple reversible model described in the last section into an open system we need only add a source reaction and a sink reaction as shown in the following scheme:



In this case simple mass-action kinetics is assumed for all reactions. It is also assumed that the source reaction, with rate  $v_o$ , is irreversible and originates from a boundary species,  $X_o$ , which itself is fixed. In addition, it is assumed that the sink reaction, with rate constant,  $k_3$  is also irreversible. For the purpose of making it easier to derive the time course solution, the reverse rate constant,  $k_2$ , will be assumed to equal zero. We will also set the initial conditions for  $A$  and  $B$  to equal zero. The mathematical solution for the system can be obtained using Mathematica:

$$a(t) = v_o \frac{1 - e^{-k_1 t}}{k_1}$$
(4.6)

$$b(t) = v_o \frac{k_1 (1 - e^{-k_3 t}) + k_3 (e^{-k_1 t} - 1)}{k_3 (k_1 - k_3)}$$

As  $t$  tends to infinity,  $a(t)$  tends to  $v_o/k_1$  and  $b(t)$  tends to  $v_o/k_3$ . In addition, the reaction rate through each of the three reaction steps is  $v_o$ . This can be confirmed by substituting

the solutions for  $A$  and  $B$  into the reaction rate laws. Given that  $v_o$  is greater than zero and that  $A$  and  $B$  reach constant values given sufficient time, we conclude that this system eventually settles to a steady state rather than thermodynamic equilibrium. The system displays a continuous flow of mass from the sink to the source. This can only continue undisturbed so long as the source material,  $X_o$ , never runs out and the sink is continuously emptied. Figure 4.2 shows a simulation of this system.

Usually however, we cannot solve the equations and so must revert to computer simulation or using specialist software (such as Tellurium) to compute the steady state. The script below shows a Tellurium model where we ask the software to solve for the steady state using the command `r.steadyState()`.

At steady state, the rate of mass transfer across a reaction is often called the **flux**, or  $J$ .

At steady state the net reaction rate is also called the **pathway flux**, often symbolized with the letter  $J$ .

We can sometimes also calculate the steady state mathematically. In the last example we used the simplified model:



The differential equations for this system are:

$$\frac{da}{dt} = v_o - k_1 a$$

$$\frac{db}{dt} = k_1 a - k_3 b$$

If we set the rates of change to zero:

$$0 = v_o - k_1 a$$

$$0 = k_1 a - k_3 b$$

We have equations in two unknowns,  $a$  and  $b$ . We can solve for  $A$  and  $B$  to obtain:

$$a = v_o / k_1$$

$$b = v_o / k_3$$

More often we cannot solve the equations and so must revert to computer simulation or using specialist software (such as Tellurium) to compute the steady state. The script below shows a Tellurium model where we ask the software to solve for the steady state using the command `r.steadyState()`.

### Thermodynamic Equilibrium and Steady State

Thermodynamic equilibrium (or equilibrium for short) and the steady state are distinct states of a chemical system. If we consider a system where every part is in equilibrium, then we can be sure of two things: (i) the species concentrations are unchanging, and (ii) most importantly, there are no net flows of mass or energy within the system or between the system and the environment. A system that is in equilibrium must have the following properties:

$$\frac{d\mathbf{s}}{dt} = 0$$

$$\text{for all } i: v_i = 0$$

where  $v_i$  is the net reaction rate for the  $i^{\text{th}}$  reaction step.  $\mathbf{s}$  is bold to indicate that it is a vector. When a biological system is at equilibrium, we say it is **dead**. Thermodynamically, we can also say that entropy production is at zero and has reached its **maximum value**.

The steady state has some similarities with the equilibrium state, species concentrations are still unchanging, however **there are net flows** of energy and mass within the system and with the environment. Systems at steady state must therefore be open and will necessarily continuously dissipate any gradients between the system and the external environment. This means that one or more  $v_i$ 's must be non-zero.

The steady state is defined when all  $dS_i/dt$  are equal to zero while one or more reaction rates are non-zero:

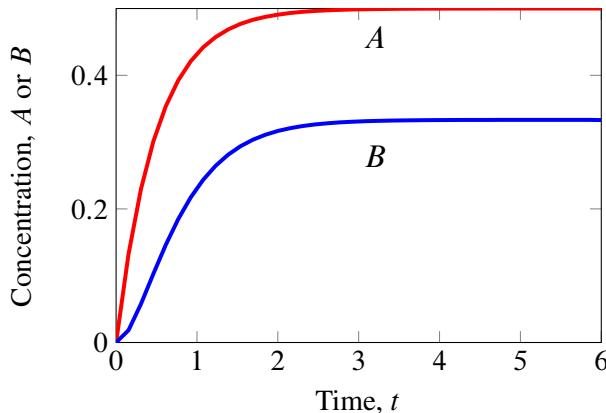
$$\frac{d\mathbf{s}}{dt} = 0$$

$$v_i \neq 0$$

Thermodynamically, we can also say that the entropy production of the system at steady state is lower than the entropy production in the environment. In some of the literature the terms equilibrium and steady state are used interchangeably resulting in possible confusion. In this book, the word equilibrium will be used to refer to a system at thermodynamic equilibrium, not at steady state.

```
import tellurium as te
r = te.loada '''
$X0 -> A;    vo;
A -> B;    k1*A;
B -> $X1;  k3*B;

# Set up the model initial conditions
```



**Figure 4.2** Time course for an open system reaching steady state in model (4.7) where  $v_o = 1, k_1 = 2, k_2 = 0, k_3 = 3, a_o = 0, b_o = 0$ .  $x_o$  is assumed to be fixed. Tellurium model script: (4.2).

```

Xo = 1;      X1 = 0;
k1 = 0.2;   k3 = 0.3;
vo = 0.5;
''')

# Evaluate the steady state
r.steadyState();

print "Steady State values:", r.A, r.B;

# Output follows:
Steady State values:  2.5  1.66667

```

## 4.5 Transients

The final simple behavior that a system can show is a transient. A transient is usually the change that occurs in the species concentrations as the system moves from one state, often a steady state, to another. Equation (4.6) shows the solution to a simple system that describes the transient behavior of species  $A$  and  $B$ . Figure 4.2 illustrates the transient from an initial condition, in this case from a non-steady state condition to a steady state. A periodic (such as an oscillation) or a chaotic system may be considered a transient, one that is unable to settle to a fixed steady state. In the case of a system showing periodic behavior, the transient repeats itself indefinitely at regular intervals called the period. In a chaotic system, the transient never repeats the exact same trajectory but will continue indefinitely.

## 4.6 Setting up a Model in Software

---

There are many software tools both commercial and free (including open source) that one can use to build models of cellular networks. In this book we will use Tellurium, a Python based software tool written by the author and coworkers. Tellurium is a script based tool where one enters a model as a text file, the model is then compiled, run, and the results displayed. For those who wish to use other tools, such as PySCeS<sup>2</sup>, COPASI<sup>3</sup>, CellDesigner<sup>4</sup>, iBioSim<sup>5</sup>, VCell<sup>6</sup> or even Matlab<sup>7</sup>, it is easy to convert Tellurium files into standard Systems Biology Markup Language (SBML) or Matlab scripts and then load the models into the simulation tool of choice.

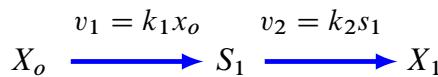
## 4.7 Effect of Different Kinds of Perturbations

---

When we talk about model dynamics we mean how species levels and reaction rates change in time as the model evolves. There are a number of ways to elicit a dynamic response in a model, the two we will consider here are perturbations to species and perturbations to model parameters.

### Effect of Perturbing Floating Species

Let us consider a two step pathway of the following form:



We assume that  $X_o$  and  $X_1$  are fixed. If the initial concentration of  $S_1$  is zero then we can run a simulation and allow the system to come to steady state. This is illustrated in Figure 4.3.

Once at steady state, we can consider applying perturbations to see what happens. For example, Figure 4.4 illustrates the effect of adding 0.35 units of  $S_1$  at  $t = 20$  and watching the system evolve. In practice, this could be accomplished by injecting 0.35 units of  $S_1$  into the volume where the pathway operates. What we observe is that the concentration of  $S_1$  initially jumps by the amount 0.35, then relaxes back to its steady state concentration before the perturbation was made (Figure 4.4). When we apply perturbations to species concentrations and the change relaxes back to the original state, we call the system **stable**. We will return to the topic of stability in a later chapter.

<sup>2</sup><http://pysces.sourceforge.net/>

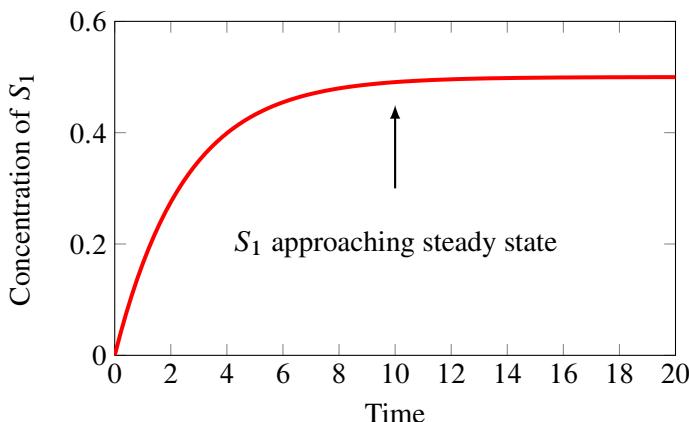
<sup>3</sup><http://www.copasi.org>

<sup>4</sup><http://celldesigner.org/>

<sup>5</sup><http://www.async.ece.utah.edu/ibiosim>

<sup>6</sup>[vcell.org](http://vcell.org)

<sup>7</sup><http://www.mathworks.com>



**Figure 4.3**  $S_1$  approaching steady state. Tellurium model script: (4.3).

Figure 4.4 illustrates perturbing one of the floating molecular species by physically adding a specific amount of the substance to the pathway. In many cases the system will recover from such perturbations as we see in Figure 4.4. We are not limited to single perturbations; Figure 4.5 shows multiple perturbations, both positive and negative. Not all systems show recovery like this, those that do not are called **unstable**. That is when we perturb a species concentration, instead of the perturbation relaxing back, it begins to diverge.

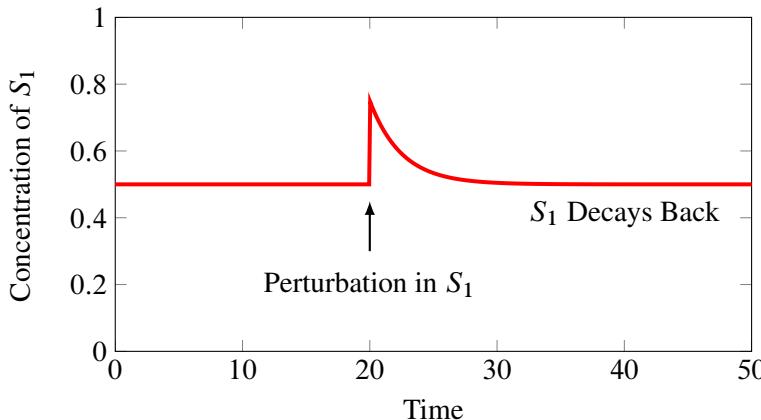
There is one caveat to this definition. For those models that include conserved moieties, it is important that when perturbing a species that is part of a conserved moiety, that one of the other conserved spools be decreased by the equivalent amount. If this is not done, then the perturbation is actually a change to a total amount of conserved species which will result in a shift to the steady state solution. That is the system won't appear to relax back to its original state.

### Stability:

A biochemical pathway is dynamically **stable** at steady state if small perturbations in the floating species concentrations relax back to the original state. For species that are part of a conserved moiety cycle, it is important to perturb one of the species in one direction and another, by exactly the same amount, in the opposite direction.

Conversely if a small perturbation at steady state to one of the floating species causes the system to diverge, this is called an **unstable** steady state.

System stability will be revisited in Chapter 13 where a much more detailed discussion will be presented. For now it can be safely assumed that all the steady states we will consider in subsequent chapters are stable.



**Figure 4.4** Stability of a simple biochemical pathway at steady state. The steady state concentration of the species  $S_1$  is 0.5. A perturbation is made to  $S_1$  by adding an additional 0.35 units of  $S_1$  at time = 20. The system is considered stable because the perturbation relaxes back to the original steady state. Tellurium model script (4.4).

### Effect of Perturbing Model Parameters

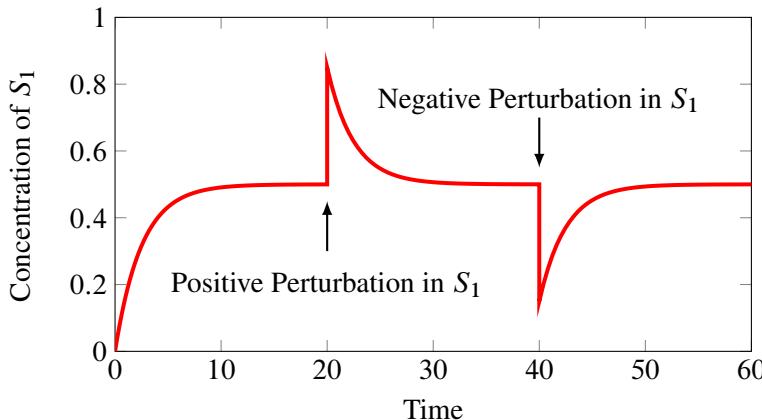
In addition to perturbing floating species, we can also perturb the model parameters. Such parameters include kinetic constants and boundary species. When changing a parameter we can do it in two ways, we can make a permanent change, or we can make a change and at some time later return the parameter to its original value. Assuming that the steady state is stable, a temporary change will result in the steady state changing then recovering to the original state when the parameter is changed back. Figure 4.6 shows the effect of perturbing the rate constant,  $k_1$  and then restoring the parameter to its original value at some time later.

In some applications other types of perturbations are made. For example, in studying the infusion of a drug where the concentration of the drug is a model parameter, one might use a slow linear increase in the drug concentration. Such a perturbation is called ramp. More sophisticated analyzes might require a sinusoidal change in a parameter, an impulse, a pulse or an exponential change. The main point to remember is that parameter changes will usually result in changes to the steady state concentrations and fluxes.

For completeness, Figure 4.7 shows what happens when we perturb both a parameter and a species concentration. As expected the species concentration does not recover to the original steady state.

## 4.8 Sensitivity Analysis

Sensitivity analysis at steady state looks at how particular model variables are influenced by model parameters. There are at least two main reasons why it is interesting to examine



**Figure 4.5** Multiple Perturbations. The steady state concentration of the species  $S_1$  is 0.5 and a perturbation is made to  $S_1$  by adding an additional 0.35 units of  $S_1$  at time = 20 and removing 0.35 units at time = 40. In both cases the system relaxes back. Tellurium model script: (4.5).

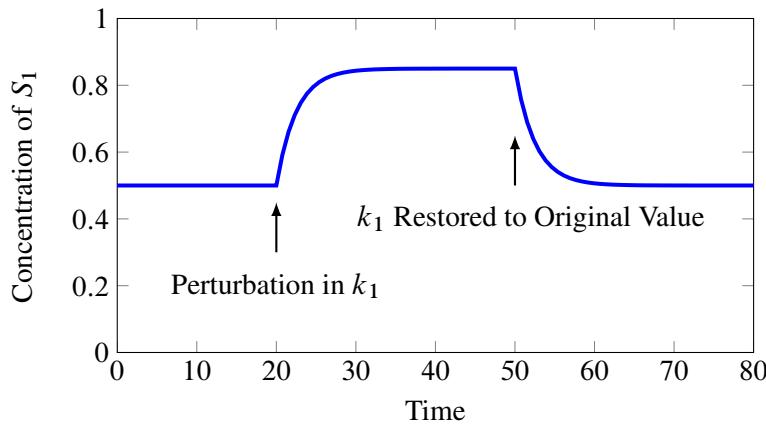
sensitivities. The first is a practical one. Many kinetic parameters we use in building biochemical models can have a significant degree of uncertainty about them. By determining how much each parameter has an influence on the model's state, we can decide whether we should improve our confidence in the particular parameter. A parameter that has considerable influence but at the same time has significant uncertainty, is a parameter that should be determined more carefully by additional experimentation. On the other hand, a parameter that has little influence but has significant uncertainty associated with it, is relatively unimportant. A sensitivity analysis can therefore be used to highlight parameters that need better accuracy.

The second reason for measuring sensitivities is to provide insight. The degree to which a parameter can influence a variable tells us something about how the network is responding to perturbations. Such a study can be used to answer questions about robustness and adaptation. We will delay further discussion of this important topic to the second half of the book where we will describe it much more detail.

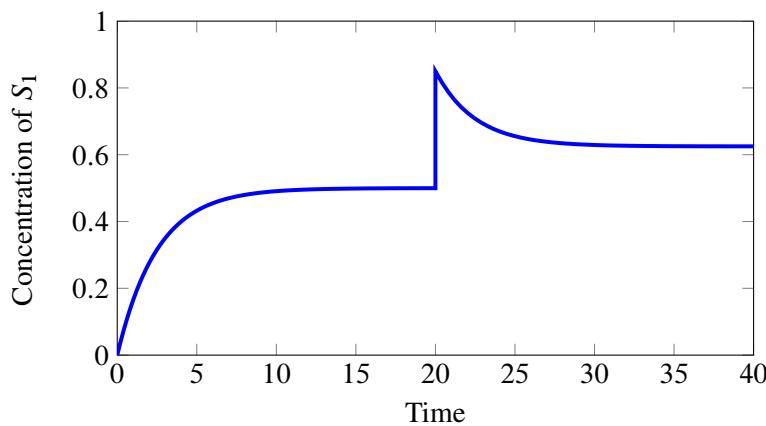
How are sensitivities represented? Traditionally there are two way, one based on absolute sensitivities and the second based on relative sensitivities. Absolute sensitivities are simply given by the ratio of the absolute change in the variable,  $V$ , to the absolute change in the parameter,  $p$ . That is:

$$S = \frac{\Delta V}{\Delta p}$$

where  $S$  is the absolute sensitivity. This equation shows finite changes to the parameter and variable. Unfortunately because most systems are nonlinear, the value for the sensitivity will be a function of the size of the finite change. To make the sensitivity independent of



**Figure 4.6** Effect of Perturbing Model Parameters. Tellurium model script: (4.6).



**Figure 4.7** Effect of Perturbing Model Parameters and Species Concentration.

the size of the change, the sensitivity is usually defined in terms of infinitesimal changes:

$$S = \frac{dV}{dp}$$

Although absolute sensitivities are straightforward, they have one drawback, the value can be influenced by the units used to measure the variable and parameter. In making experimental measurements we won't often be able to measure the quantity using the most natural units, instead we may have measurements in terms of fluorescence, colony counts, staining on a gel and so on. It is most likely that the variable and parameter units will be quite different and each laboratory may have its own preferences for the unit it uses. Absolute sensitivities can therefore be quite difficult to compare.

To get around the problem of units, many people will use relative sensitivities. These are scaled absolute sensitivities:

$$S = \frac{dV}{dp} \frac{p}{V}$$

The sensitivity is defined in terms of infinitesimal changes for the same reason cited before. Relative sensitivities are immune from the units we use to measure quantities but also relative sensitivities correspond more closely to how many measurements are made, often in terms of relative or fold changes. In practice steady state relative sensitivities should be measured by taking a measurement at the operating steady state, making a perturbation (preferable a small one), waiting for the system to reach a new steady state and then measuring the system again. It is important to be aware that the steady state sensitivities measure how a perturbation in a parameter moves the system from one steady state to another.

#### ***Example 4.1***

---

Estimate the unscaled and scaled steady state sensitivity of  $S_1$  with respect to the boundary species,  $X_o$  in the following model:

```
import tellurium as te
r = te.loada ('''
$Xo -> S1; k1*Xo;
S1 -> $X1; k2*S1;

k1 = 0.2;
k2 = 0.4;
Xo = 1;
S1 = 0.0;
''')
```

Run a time course simulation of the model until the concentration of  $S_1$  reaches a steady value. Use the following script to find the steady state:

```
m = r.simulate (0, 25, 10)
print "Steady State value = ", m[9,1]
```

From this we obtain the steady state value of  $S_1$  to be  $0.49997 = 0.5$ . Call this value  $S_1(\text{ref})$ . Make a perturbation in  $X_o$  of 10 percent and rerun the model to compute the new steady state. From this we obtained a new steady state concentration of  $S_1$  to be  $0.54997 = 0.55$ . Call this value  $S_1(\text{pert})$ . The unscaled sensitivity can be computed approximately using:

$$\frac{S_1(\text{pert}) - S_1(\text{ref})}{X_o(\text{pert}) - X_o(\text{ref})} = \frac{0.55 - 0.5}{1.1 - 1.0} = 0.5$$

The scaled sensitivity can be computed using:

$$\frac{S_1(\text{pert}) - S_1(\text{ref})}{X_o(\text{pert}) - X_o(\text{ref})} \frac{X_o(\text{ref})}{S_1(\text{ref})} = 0.5 \frac{1.0}{5} = 0.1$$


---

## 4.9 Robustness and Homeostasis

---

Biological organisms are continually subjected to perturbations. These perturbations can originate from external influences such as changes in temperature, light or the availability of nutrients. Perturbations can also arise internally due to the stochastic nature of molecular events or by natural genetic variation. One of the most remarkable and characteristic properties of living systems is their ability to resist such perturbations and maintain very steady internal conditions. For example, the human body can maintain a constant core temperature of  $36.8^{\circ}\text{C} \pm 0.7$  even though external temperatures may vary widely. The ability of a biological system to maintain a steady internal environment is called **homeostasis**, a phrase introduced by Claude Bernard almost 150 years ago. Modern authors may also refer to this behavior as **robustness**, although this word is used in many other contexts.

There are a number of mechanisms that are used in biology to maintain homeostasis. Perhaps the most common is negative feedback. This is where the difference between the desired output, and the actual output is used to modulate the process that determines the output. For example, if the output is lower than the desired output then the process will increase the output. Such systems are found at multiple levels in a living organism, including subcellular processes such as metabolism and multicellular processes that control, for example, the level of glucose in the blood stream. One way to measure the degree of robustness or homoeostasis in a system is to use sensitivity analysis. We will investigate the use of negative feedback to maintain concentrations within a narrow range in a later chapter.

## Further Reading

---

1. Klipp E, Herwig R, Kowald A, Wierling C and Lehrach H (2005) Systems Biology in Practice, Wiley-VCH Verlag.
2. Steuer R, Junker BH (2008) Computational Models of Metabolism: Stability and Regulation in Metabolic Networks, Advances in Chemical Physics, Volume 142, (ed S. A. Rice), John Wiley & Sons, Inc.

---

## Exercises

---

1. Describe the difference between thermodynamic equilibrium and a steady state.
2. Write out the differential equations for the system  $A \rightarrow B \rightarrow C$  where the reaction rates are given by:

$$\begin{aligned}v_1 &= k_1a - k_2b \\v_2 &= k_3b - k_4c\end{aligned}$$

Find the concentrations of  $A$ ,  $B$  and  $C$  when the rates of change are zero  $da/dt = 0$ ,  $db/dt = 0$ ,  $dc/dt = 0$ . Show that this system is at thermodynamic equilibrium when the rates of change are zero.

3. What do we mean by the phrase quasi-equilibrium?
4. Find the mathematical expression for species  $A$  and  $B$  that describes the steady state for the network:



Assume that  $X_o$  is fixed and that all reactions are governed by simple mass-action kinetics.

5. Explain what is meant by a stable and unstable steady state.
6. The steady state of a given pathway is stable. Explain the effect in general terms on the steady state if:
  - a) A bolus of floating species is injected into the pathway.
  - b) A permanent change to a kinetic constant.
7. Why are scaled sensitivities sometimes of more advantage than unscaled sensitivities?

---

## Appendix

---

### Solution using Laplace Transforms

Solution to equation (4.3) using Laplace transforms. In the following differential equation,  $a$  has been substituted with  $x$  to make the notation more conventional:

$$\frac{dx}{dt} = k_2T - x(k_1 + k_2)$$

Assume the initial conditions for  $x = x_o$  and that the total mass at  $t = 0$  is  $T$ . Take the Laplace Transforms on both sides:

$$sX(s) - x_o = \frac{k_2 T}{s} - X(s)(k_1 + k_2)$$

Rearrange and solve for  $X(s)$ :

$$X(s) = \frac{k_2 T + s x_o}{s(s + (k_1 + k_2))}$$

Separate the right-hand term into two simpler terms:

$$X(s) = \frac{x_o}{s + (k_1 + k_2)} + \frac{k_2 T}{s(s + (k_1 + k_2))}$$

We now take the inverse Laplace transform. Both terms on the right can be found in standard Laplace transform tables such that:

$$x(t) = x_o e^{-(k_1 + k_2)t} + \frac{k_2 T}{k_1 + k_2} \left(1 - e^{-(k_1 + k_2)t}\right)$$

■

## Tellurium Scripts

Tellurium modeling scripts using Python. See Appendix D for more details of Tellurium.

```
import tellurium as te
r = te.loada ('''
    A -> B;    k1*A;
    B -> A;    k2*B;
end;

p.A = 10; p.k1 = 1;
p.B = 0;  p.k2 = 0.5;

m = p.sim.eval (0, 6, 100);
graph (m);
```

**Listing 4.1** Script for Figure 4.1.

```
import tellurium as te
r = te.loada ('''
    $Xo -> S1;    vo;
    S1 -> S2;    k1*S1 - k2*S2;
    S2 -> $X1;    k3*S2;
```

```

vo = 1;
k1 = 2; k2 = 0;
k3 = 3;
''')

m = r.simulate (0, 6, 100);
r.plot()

```

**Listing 4.2** Script for Figure 4.2.

```

import tellurium as te
r = te.loadada ('''
    $Xo -> S1; k1*Xo;
    S1 -> $X1; k2*S1;

    k1 = 0.2;
    k2 = 0.4;
    Xo = 1;
    S1 = 0.0;
''')

m = r.simulate (0, 20, 100, ['time', 'S1']);
r.plot()

```

**Listing 4.3** Script for Figure 4.3.

```

import tellurium as te
import numpy as np
r = te.loadada ('''
    $Xo -> S1; k1*Xo;
    S1 -> $X1; k2*S1;

    k1 = 0.2;
    k2 = 0.4;
    Xo = 1;
    S1 = 0.5;
''')

# Simulate the first part up to 20 time units
m1 = r.simulate (0, 20, 100, ['time', 'S1']);

# Perturb the concentration of S1 by 0.35 units
r.S1 = r.S1 + 0.35;

# Continue simulating from last end point

```

```
m2 = r.simulate (20, 50, 100, ['time', 'S1']);

# Merge and plot the two halves of the simulation
m3 = np.vstack ((m1, m2))
r.plot(m3)
```

**Listing 4.4** Script for Figure 4.4.

```
import tellurium as te
import numpy as np
r = te.loada ('''
    $Xo -> S1; k1*Xo;
    S1 -> $X1; k2*S1;

    k1 = 0.2; k2 = 0.4;
    Xo = 1;   S1 = 0.0;
''')

# Simulate the first part up to 20 time units
m1 = r.simulate (0, 20, 100, ['time', 'S1']);

# Perturb the concentration of S1 by 0.35 units
r.S1 = r.S1 + 0.35;

# Continue simulating from last end point
m2 = r.simulate (20, 40, 50, ['time', 'S1']);
# Merge the data sets
m3 = np.vstack ((m1, m2))
# Do a negative perturbation in S1
r.S1 = r.S1 - 0.35;

# Continue simulating from last end point
m4 = r.simulate (40, 60, 50, ['time', 'S1']);

# Merge and plot the final two halves of the simulation
r.plot (np.vstack ((m3,m4)))
```

**Listing 4.5** Script for Figure 4.5.

```
import tellurium as te
import numpy as np
r = te.loada ('''
    $Xo -> S1; k1*Xo;
    S1 -> $X1; k2*S1;

    k1 = 0.2; k2 = 0.4;
```

```
Xo = 1; S1 = 0.5;
''')

# Simulate the first part up to 20 time units
m1 = r.simulate (0, 20, 5, ['time', 'S1']);

# Perturb the parameter k1
r.k1 = r.k1*1.7;

# Simulate from the last point
m2 = r.simulate (20, 50, 40, ['time', 'S1']);

# Restore the parameter back to ordinal value
r.k1 = 0.2;

# Carry out final run of the simulation
m3 = r.simulate (50, 80, 40, ['time', 'S1']);

# Merge all data sets and plot
m4 = np.vstack((m1, m2, m3));
r.plot (m4);
```

**Listing 4.6** Script for Figure 4.6.



# 5

## *Traditional Concepts in Metabolic Regulation*

### **5.1 Introduction**

---

This book is about the control of biochemical systems with a focus on metabolic pathways. The ability to control reaction rates and concentrations in a changing environment is one of the characteristics of living systems. Cells must monitor prevailing conditions and make appropriate decisions. Cells make sure, for example, that adequate phosphate and redox potentials are available at all times. They also have to ensure that major transitions from one state to another (for example cell division) avoid any disruption to subsystems that are essential to cell viability. These activities presumably require a great deal of coordination and control. Indeed, years of research has uncovered a myriad number of feedback and feedforward control loops together with many less obvious means of control.

It is worth examining some of the history of how we came to understand control in biological cells. The first thing to note is that understanding control in any complex system is challenging. It was difficult in the past and it is difficult now. Man's propensity to grasp the many factors involved in a complex system is limited. As a result, reasoning about complex systems cannot be done by intuition alone, it requires expertise and the application of approaches from mathematics, engineering and computer science.

## 5.2 Early Quantitative Efforts

---

During the early part of the 20th century, it became apparent that chemical processes in biological cells were the result of sequences of separate chemical transformations. The first sequence of steps discovered, later to be called a ‘pathway’, was yeast glycolysis. Subsequently, many other pathways were discovered such as the Calvin and Krebs cycle, including pathways involved in amino acid biosynthesis and degradation. As early as the 1930s, a number of individuals began taking a theoretical interest in the dynamic properties of such pathways. Much of the early work focused on the question of limiting factors. This may have originated from a statement by Blackman [8] in 1905 who stated as an axiom: “when a process is conditioned as to its rapidity by a number of separate factors, the rate of the process is limited by the pace of the slowest factor”. This implied that the understanding of a complex system could be accomplished by identifying the limiting factor; and so the idea of the rate-limiting step, the pacemaker, the bottleneck, and master reaction was born.

### The Pacemaker

Although the idea of a pacemaker reaction in a pathway was extremely attractive, there were opponents to the idea even as early as the 1930s. Burton [18] was probably one of the first to point out that: “In the steady state of reaction chains, the principle of the master reaction has no application”. Hearon [50] made a more general mathematical analysis and developed strict rules for the prediction of mastery in a linear sequence of enzyme-catalysed reaction. Webb [149] gave a severe criticism of the concept of the pacemaker and of its blind application to solving problems of regulation in metabolism. Waley [146] made a simple but clear analysis of linear reaction chains that showed that rate-limitingness was a shared commodity. Later authors from the biochemical community, such as Higgins [54] but particularly Heinrich and Rapoport [52], supported the same conclusion with more advanced analysis. In parallel with this work, other communities were coming to the same conclusion. Most notably Sewell Wright, a geneticist, wrote a treatise on ‘Physiological and Evolutionary Theories of Dominance’ [153], where he discussed the limiting factors in relation to hypothesized networks controlled by ‘genes’. This work was taken up by Kacser and Burns [63] in Edinburgh and was developed into a major theory of control in pathways. Heinrich and Rapoport [52] simultaneously accomplished the same feat but from a more biochemical perspective. Finally, Savageau [113] in the United States, an engineer by training, developed the same approach and reached similar conclusions.

## 5.3 Prevailing Ideas

---

Nevertheless, while there was considerable theoretical and some experimental work that suggested that the concept of the pacemaker was erroneous, the biochemical community, for what ever reason, ignored these results. Instead, the biochemistry community, which

had largely morphed into molecular biology in the 1970s, developed its own framework for understanding the operating principles of cellular networks. This framework was derived largely through an intuitive approach, based neither on experimental evidence or mathematical reasoning. This ultimately led to a number of unfortunate misunderstandings in how cellular networks operate, misconceptions that still prevail today.

One of the chief concepts in the traditional control framework is the pacemaker or rate-limiting step. The rate-limiting step is thought to be located near the start of a pathway and because it is rate-limiting, the pathway is controlled entirely by this one key step. In addition, it is proposed that rate-limiting steps are likely to be the site for allosteric regulation. There are a number of criteria that are used to identify the possible rate-limiting step, though there is no real definitive test. These criteria include:

- The rate-limiting step is the slowest step in the pathway.
- The rate-limiting step has the lowest substrate-affinity (highest  $K_m$ ), meaning that the reaction velocity is the lowest when saturating substrate concentrations are present for all enzymes.
- The rate-limiting step will be the regulated step.
- The rate-limiting step is an irreversible reaction.
- The rate-limiting step is usually the first step in the pathway.
- The rate-limiting step is far from equilibrium.

No single criterion could positively identify a rate limiting step, but the cross-over theorem is one that was considered important. The technique worked as follows. A metabolic pathway is perturbed by adding an inhibitor that changes the activity of one of the enzyme catalyzed steps, and the metabolite concentrations before and after the inhibited step are measured. If the inhibited step is rate limiting, then the metabolites upstream would increase and those downstream decrease. The technique was originally developed by Britton Chance [20] in the 1950s as a means to study the electron transport chain in mitochondria. The advantage here was that many of the intermediates had characteristic absorption spectra so could be easily measured. The method was used to identify the sites where electron transfer was being coupled to ATP production. Although applicable to the electron transport chain (Fell, 1996), its subsequent use to identify sites of regulation in metabolic pathways has been considered on theoretical grounds to be untrustworthy (Heinrich et al, 1974).

## Regulatory Enzymes and Feedback Regulation

A key concept in traditional metabolic control is that the reaction step where feedback regulation acts is the rate-limiting step. Rate-limiting enzymes could therefore be identified

by locating the regulated steps. For example, a classic rate-limiting step in glycolysis is phosphofructokinase. Understanding how pathways are controlled however is much more subtle than this. Repeated measurements [51, 116, 27, 17, 134, 106, 89, 140, 80, 88] have shown for example that phosphofructokinase is in fact not rate-limiting even though it is heavily regulated. Since the development of recombinant technology in the late 70s, the ability to control enzyme levels has become relatively easy. There are many experiments reported where over expression of a regulated step resulted in no change in the pathway flux even though such steps were considered rate-limiting. Even in the face of considerable experimental evidence, the idea that regulated steps are rate limiting continues to persist. As we will see in subsequent chapters, such an idea is inconsistent with logic and experimental evidence even though most undergraduate textbooks (except for Lehninger) and Wikipedia (a major source of information for students) continue to support the concept.

## **5.4 A Modern Understanding of Metabolism** ---

Although the metabolic parts list is almost complete, as witnessed by the development of genomic scale metabolic reconstructions [94], our understanding of how metabolism operates is still conceptually primitive and incomplete. The last four decades has seen some progress but still many researchers use the classical conceptual framework of metabolic control. The reminder of the book will focus on modern ideas of metabolic control.

This completes the preliminary chapters to the book. In Part II we will delve more deeply into the current approaches to understanding metabolic control and regulation.

## **Further Reading** ---

1. Fell D A (1996) Understanding the Control of Metabolism. Ashgate Publishing. ISBN-10: 185578047X
2. Fell D A (1992) Metabolic control analysis: a survey of its theoretical and experimental development. *Biochemical J*, 286, 313-330
3. Morandini P (2009) Rethinking metabolic control. *Plant Science*, 176(4), 441-451

## **Part II**

# **Metabolic Control Analysis**



## Preface to Part II

---

Metabolic Control Analysis (MCA) is a general approach used to understand how perturbations propagate through a biochemical network. Although it is called metabolic control analysis, the approach applies equally well to gene regulatory and protein signaling networks. Historically, during the inception of MCA, the focus was on metabolic systems and the name metabolic control analysis seems to have stuck. Ideally it should be called biochemical control analysis, but it's difficult to change something after it has been used for so long. Savageau developed an almost identical approach to MCA called Biochemical Systems Theory in the United States. He was forward thinking enough to realize that the approach was quite general, and not restricted to just metabolic systems. In this book I will use the notation that the Kacser/Heinrich groups in Europe developed rather than the notation developed by Savageau. There are two reasons for this, one is my own familiarity with the European notation; secondly, and this is no criticism of Savageau's work, the European model has been developed much further. I think this has partly to do with cultural differences. Having lived the United States for over 18 years, I have found that the American culture tends to place more emphasis on the practical and empirical rather than the theoretical. Savageau's work, which was largely theoretical, therefore was not as widely known or developed as the European efforts.



# 6

## *Elasticities*

### **6.1 Introduction**

---

Enzymes catalyze virtually all the chemical transformations of metabolism. They coordinate all the primary activities of a cell, ranging from energy transformations and storage, through to maintenance of cellular structure and integrity. They directly manage the expression and maintenance of host DNA, including replication. Enzymes clearly serve an essential and fundamental role in the activity of a cell, and for this reason we can regard them as fundamental units of life. If we are to understand how cellular systems work, an appreciation of the properties of these fundamental units is obviously essential. This first section will focus on the properties of the isolated enzymes, and later we will consider intact pathways.

Since enzymes are the functional units of metabolism, it is important to understand how an enzyme responds to changes in its environment. An important part of Metabolic Control Analysis (MCA) is a consideration of this question. In MCA the measure that describes this response is called the **elasticity coefficient**. Elasticity coefficients are so important to MCA that the remainder of the chapter is devoted to their discussion.

Elasticities describe how sensitive a reaction rate is to changes in reactant, product and effector concentrations. They represent the degree to which changes are transmitted from the immediate environment to the reaction rate. From a systems perspective they are critical components in understanding how a disturbance, such as the introduction of a drug applied at one or more points in a cellular pathway, propagates to the rest of the system. It is the magnitude and signs of the elasticities that determine how far and at what strength the disturbance travels. Elasticities are therefore central in helping us understand how networks function. In this chapter we will focus on describing the properties of elasticities, and how

they can be computed and used to describe changes at a reaction step.

To study the properties of an individual enzyme, the usual experimental procedure is to purify the enzyme and study it *in vitro*. Once purified and isolated, the environment of the enzyme can be controlled and in principle, the concentrations of all the participating molecules manipulated at will. Individual substrates, effectors etc., can be selectively changed and any change in rate recorded. In this manner, the response of the reaction rate to changes in factors that might affect the reaction rate can be studied.

Consider an experiment where we investigate the response of the reaction rate to changes in substrate concentration. In the experiment the concentration of substrate can be changed and the change in reaction rate observed. A plot of the reaction rate versus the substrate concentration would form a continuous curve. Let us denote the concentration of substrate by the symbol  $s$ , and the rate of reaction by  $v$ . The experiment proceeds in two steps. The first step involves measuring the rate of reaction,  $v$ , at some substrate concentration of interest, say  $s$ . In the second part of the experiment, the concentration of substrate is increased by an amount given by  $\delta s$ ,<sup>1</sup> and the experiment repeated at the new concentration of  $s + \delta s$ . The increase in  $s$  is likely to cause a change in the rate of reaction from  $v$  to  $v_{\text{new}}$ . The difference between the two rates,  $v_{\text{new}} - v$  is the change in rate as a result of the change,  $\delta s$ . We denote this change in rate by  $\delta v$ . Depending on the particular enzyme, the effectors, the substrates and products, the change we observe in the rate may be small or large. In order to judge the relative effectiveness of any particular modifier, we can form the ratio:

$$\frac{\delta v}{\delta s}$$

This will give us the change in  $v$  per unit change in  $s$ . By measuring this ratio for each factor that might affect the rate, we can gauge which ones have more or less of an effect.

There are however, two problems with this ratio. The first is that its value depends on the size of the change we make to  $s$ . This is particularly true if the response of  $v$  to changes in  $s$  is non-linear (as most enzyme rate responses are). The second problem is that the ratio depends on the units we choose to measure the rate and concentration. A possible solution to the later problem would be for all experimenters to employ a standard set of units, but this would be almost impossible to achieve in practice. A much easier way around this problem is to eliminate the units altogether by scaling the ratio with the rate and concentration. We can eliminate the concentration units by dividing the change,  $\delta s$ , by the concentration of  $s$ , i.e.  $\delta s/s$ . Likewise, we can eliminate the reaction units by dividing by  $v$ . Therefore, rather than measure  $\delta v/\delta s$ , it is more sensible to measure:

$$\frac{\delta v}{\delta s} \frac{s}{v}$$

This still leaves us with the first problem, which is that the value of the ratio varies with the amount of change we make to  $s$ . We could all decide on a standard change to make in  $s$ , say

---

<sup>1</sup>The  $\delta$  means ‘a small change’

doubling  $s$ , and measuring the change in  $v$ . This would be difficult to achieve in practice, however, and ultimately has limited value. A better way is as follows.

Assume that the substrate concentration has been set to a value  $s$ . At this concentration, the enzyme will show a reaction rate of  $v$ . If we make a change  $\delta s$  to  $s$ , then this will cause a change in rate  $\delta v$  and we can compute the ratio,  $\delta v / \delta s$ . We could make the change smaller and remeasure the change in  $v$  and compute the ratio again. If we were to continue making  $\delta s$  smaller and smaller, the ratio given by  $\delta v / \delta s$  will slowly approach a limiting slope. This slope is the tangent to the curve at the point  $s$ . Those familiar with the calculus will recognize that in reducing  $\delta s$  to a smaller and smaller increment, the ratio,  $\delta v / \delta s$ , has reached a limiting value called the derivative:

$$\frac{\delta v}{\delta s} \longrightarrow \frac{dv}{ds} \quad \text{as } \delta s \rightarrow 0 \quad (6.1)$$

The ratio,  $\delta v / \delta s$  tends to the differential,  $dv/ds$ , as  $\delta s$  tends to zero. The differential has a precise meaning, it is the slope of the curve at the point  $s$  and, significantly for us, it has a *unique* value at this point.

As before, we can scale  $dv/ds$  to eliminate the measuring units to obtain:

$$\frac{dv}{ds} \frac{s}{v}$$

This expression represents the scaled slope of the response curve at  $s$ , and is called the **elasticity coefficient** of the rate of reaction,  $v$ , with respect to the concentration of metabolite  $S$ . It measures how responsive a reaction rate is to changes in the concentration of a modifier, in this case the concentration of substrate,  $S$ . We could also have changed the concentration of the product,  $P$ , or the concentration of an effector. In either case we would be able to measure an elasticity. This means there will be as many elasticity coefficients for a particular enzyme as there are modifiers that might affect its reaction rate. Thus, not only will an enzyme be characterized by a substrate elasticity, but also by a product elasticity and any effector elasticities. In addition, other factors which might affect the reaction rate, such as pH, ionic strength and so on, will also have associated elasticity coefficients. Any particular enzyme will thus be fully characterized when all its elasticities have been measured or computed.

### Elasticities must be measured under *in vivo* conditions

In practice, if an enzyme is purified with the intention of measuring its elasticities, then the concentrations of the substrates and products, the pH, ionic strength and so on should be faithfully recreated in order to mimic the *in vivo* condition. If this is not done, the measured values for the elasticities will not reflect the elasticities *in vivo* and their usefulness is lost. As will be revealed in the next chapter, the elasticities are the building blocks with which we can begin to understand the properties of intact pathways.

## 6.2 Elasticity Coefficients

The **elasticity coefficient** is defined according to the following expression:

$$\varepsilon_{s_i}^v = \left( \frac{\partial v}{\partial s_i} \frac{s_i}{v} \right)_{s_j, s_k, \dots} = \frac{\partial \ln v}{\partial \ln s_i} \approx v\% / s_i\% \quad (6.2)$$

The symbol for an elasticity is the Greek epsilon,  $\varepsilon$ .

**Unitless:** Due to the scaling, the elasticity is a dimensionless quantity.

The elasticity measures how responsive a reaction rate is to changes in the concentration of a modifier, in this case the concentration of modifier  $s_i$ . Any modifier can be changed to see how it affects the reaction rate. We could have changed the concentration of the product, effector or anything else that might affect the reaction rate. The larger the elasticity, the greater the effect it has on the reaction rate<sup>2</sup>.

Sometimes the **unscaled elasticity** is useful (See Chapter 13) and we designate this using the symbol  $\mathcal{E}$ .

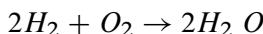
**Unscaled Elasticity:**

$$\mathcal{E}_{s_i}^v = \left( \frac{\partial v}{\partial s_i} \right)_{s_j, s_k, \dots} \quad (6.3)$$

When writing the elasticity symbol,  $\varepsilon$  or  $\mathcal{E}$ , a subscript is often used to indicate the modulating factor ( $s$ ), and a superscript to indicate the effect that is being measured ( $v$ )

The subscripts,  $s_j, s_k, \dots$  in the definition (6.2) indicate that any species or factor that could also influence the reaction rate **must be held constant** at their current value when species  $s_i$  is changed. This is also implied in the use of the **partial derivative** symbol,  $\partial$ , rather than the derivative symbol,  $d$ . In normal usage, these subscripts are often left out as the partial derivative symbol is usually sufficient to indicate what is meant.

The elasticity is closely related to the **kinetic order**, sometimes called the reaction order. For simple mass-action chemical kinetics, the kinetic order is the power to which a species is raised in the kinetic rate law. Reactions with zero-order, first-order, and second-order are commonly found in chemistry, and in each case the kinetic order is zero, one and two, respectively. For a reaction such as:



<sup>2</sup>Those familiar with quantitative economics will have come across a similar concept.

assuming that the irreversible mass-action rate law is given by:

$$v = k h_2^2 \cdot o_2$$

the kinetic order with respect to hydrogen is two and oxygen one. In this case the kinetic order also corresponds to the stoichiometric amount of each molecule although this may not always be true.

Example 6.1 shows the elasticities for zero, first, second, and  $n^{th}$  order reactions. Given that the elasticity is defined in terms of a derivative, it is possible, if the rate law is known, to compute an elasticity by differentiation (See Example 6.1). From the example we see that the elasticity reduces to the kinetic order for simple mass-action kinetics.

In biochemical systems theory, elasticities are also called the **apparent kinetic order**.

**Kinetic Order:** The elasticity for a reactant in an elementary reaction is equal to the kinetic order of the reactant.

### Example 6.1

Determine the elasticities with respect to species  $a$ , for the following mass-action rate laws by differentiating and scaling each rate law:

1.  $v = k$

**Elasticity:**  $\varepsilon_a^v = \frac{\partial v}{\partial a} \frac{a}{v} = 0$

2.  $v = ka$

**Elasticity:**  $\varepsilon_a^v = \frac{\partial v}{\partial a} \frac{a}{v} = \frac{a}{ka} = 1$

3.  $v = ka^2$

**Elasticity:**  $\varepsilon_a^v = \frac{\partial v}{\partial a} \frac{a}{v} = \frac{2kaa}{ka^2} = 2$

4.  $v = ka^n$

**Elasticity:**  $\varepsilon_a^v = \frac{\partial v}{\partial a} \frac{a}{v} = \frac{nka^{n-1}a}{ka^n} = n$

### Operational Interpretation

The definition of the elasticity (6.2) also gives us a useful operational interpretation.

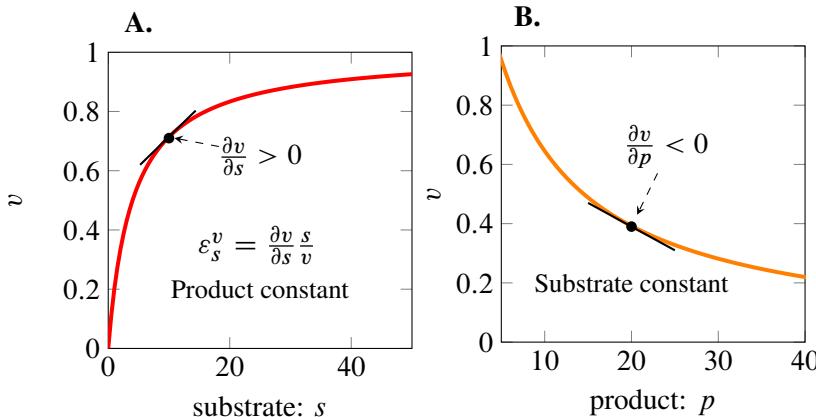
**Operational Definition:** The elasticity is the fractional change in reaction rate in response to a fractional change in a given reactant or product, while keeping all other reactants, products, and other modifiers constant.

Since the elasticity is expressed in terms of fractional changes, it is also possible to get an approximate value for the elasticity by considering **percentage changes**. For example, if we increase the substrate concentration of a particular reaction by 2% and the reaction rate increases by 1.5%, then the elasticity is given approximately by  $1.5/2 = 0.75$ . The elasticity is however only strictly defined (See equation (6.2)) for infinitesimal changes and not finite percentage changes. However, so long as the changes are small, the finite approximation is a good estimate for the true elasticity.

For species that cause reaction rates to increase, the elasticity is **positive**, while for species that cause the reaction rate to decrease, the elasticity is **negative**. Therefore, reactants generally have positive elasticities and products generally have negative elasticities (Figure 6.1).

Modifiers that **increase** a reaction rate will have a **positive** elasticity.

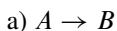
Modifiers that **decrease** a reaction rate will have a **negative** elasticity.



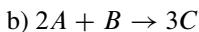
**Figure 6.1** **A.** Reaction rate versus substrate. Increases in the substrate cause an increase in the rate. A positive slope yields a positive elasticity. **B.** Reaction rate versus product (assuming a positive rate from reactant to product). Increases in product result in a decrease in reaction rate; a negative slope yields a negative elasticity. Curves generated by assuming  $v = s/(2 + s)$  and  $2/(1 + (0.1 + 0.2p))$ , respectively.

### Example 6.2

How many elasticities are there for the following reversible mass-action reactions and what are their likely signs?



There are two elasticities,  $\varepsilon_a^v$  which will be positive and  $\varepsilon_b^v$  which will be negative.



There are three elasticities,  $\varepsilon_a^v$  which will be positive,  $\varepsilon_b^v$  which will also be positive, and  $\varepsilon_c^v$  which will be negative.

---

## Numerical Estimation

We saw in example (6.1) how elasticities can be computed algebraically by differentiating the rate law and scaling. Numerically, the elasticity can be estimated by making a small change (say 5%) to the chosen reactant concentration and measuring the change in reaction rate. For example, assume that the reference reaction rate is  $v_o$ , and the reference reactant concentration,  $s_o$ . If we increase the reactant concentration by  $\Delta s_o$ <sup>3</sup> and observe the new reaction rate as  $v_1$ , then the elasticity can be estimated by using Newton's difference quotient:

$$\varepsilon_s^v \simeq \frac{v_1 - v_o}{\Delta s_o} \frac{s_o}{v_o} = \frac{v_1 - v_o}{v_o} \left/ \frac{s_1 - s_o}{s_o} \right.$$

Newton's quotient method relies on making one perturbation to  $s_o$ ,  $\Delta s_o$ . A much better estimate for the elasticity can be obtained by doing two separate perturbations in  $s_o$ . One perturbation to **increase**  $s_o$ , and another to **decrease**  $s_o$ . In each case the new reaction rate is recorded; this is called the three-point estimation method. For example, if  $v_1$  is the reaction rate when we increase  $s_o$ , and  $v_2$  is the reaction rate when we decrease  $s_o$ , then we can use the following three-point formula to estimate the elasticity:

$$\varepsilon_s^v \simeq \frac{1}{2} \frac{v_1 - v_2}{s_1 - s_o} \left( \frac{s_o}{v_o} \right)$$

These approximate numerical methods are particularly useful when computing elasticities using software.

### Example 6.3

---

Estimate the elasticity using Newton's difference quotient and the three-point estimation method. Compare the results with the exact value derived algebraically:

Let  $v = s/(0.5 + s)$ . Find the elasticity when  $s = 0.6$ .

#### a) Algebraic Evaluation

Differentiation and scaling the rate law gives the elasticity as  $0.5/(0.5 + s)$ . At a value of 0.6 for  $s$ , the exact value for the elasticity is: 0.4546

#### b) Difference Quotient

Let us use a step size of 5%. Therefore  $h = 0.05 \times 0.6 = 0.03$  from which  $s_1 = 0.63$ .  $s_o = 0.6$ . From these values we can compute  $v_1$  and  $v_o$ .  $v_o = 0.6/(0.5 + 0.6) = 0.5454$ ,  $v_1 = 0.63/(0.5 + 0.63) = 0.5575$ . From these values the estimated elasticity is given by:  $\varepsilon_s^v = ((0.5575 - 0.5454)/0.5454) / ((0.63 - 0.6)/0.6) = 0.443$

---

<sup>3</sup>  $\Delta s_o$  means a change to  $s_o$

Compared to the exact value the error is 0.0116, or **2.55 % error**.

### c) Three-Point Estimation

In addition to calculating  $v_1$  in the last example, we must also compute  $v_2$ . To do this we subtract  $h$  from  $s_o$  to give  $v_2 = 0.533$ . The three-point estimation formula gives us:

$$\varepsilon_s^v = 0.5 \frac{0.5575 - 0.5327}{0.03} \frac{0.6}{0.5454} = 0.4549$$

Compared to the exact value the error is only 0.0033, or **0.7 % error**, a significant improvement over the difference quotient method.

The degree of error in the difference quotient method will depend on the value of  $s$ , which in turn determines the degree of curvature (or nonlinearity) at the chosen point. The more curvature there is, the more inaccurate the estimate. The value in this example was chosen where the curvature is high, therefore the error was larger.

In example (6.1), the elasticities were constant values. However for more complex rate law expressions, this need not be the case (see example (6.4)), and the elasticity will change in response to changes in the reactant and product concentrations. Consequently, when measuring the elasticity numerically or experimentally, one has to choose a particular operating point, most commonly the *in vivo* state.

### **Example 6.4**

Determine the elasticities for the following rate laws by differentiating and scaling:

1.  $v = k(a + 1)$

$$\text{Elasticity: } \varepsilon_a^v = \frac{\partial v}{\partial a} \frac{a}{v} = k \frac{a}{k(a + 1)} = \frac{a}{(a + 1)}$$

2.  $v = k/(a + 1)$

$$\text{Elasticity: } \varepsilon_a^v = \frac{\partial v}{\partial a} \frac{a}{v} = -\frac{k}{(1 + a)^2} \frac{a}{k/(a + 1)} = -\frac{a}{a + 1}$$

3.  $v = a/(a + 1)$

$$\text{Elasticity: } \varepsilon_a^v = \frac{\partial v}{\partial a} \frac{a}{v} = \frac{1}{(a + 1)^2} \frac{a}{a/(a + 1)} = \frac{1}{a + 1}$$

4.  $v = ka(a + 1)$

$$\text{Elasticity: } \varepsilon_a^v = \frac{\partial v}{\partial a} \frac{a}{v} = k(1 + 2a) \frac{a}{ka(a + 1)} = 1 + \frac{a}{a + 1}$$

The examples illustrate that for more complex rate laws, the elasticity becomes a function of the reactant concentrations.

### **Experimental Estimation**

Experimentally, we can measure an elasticity using the following procedure. Consider a simple reaction such as  $A \rightarrow B$ , and measure the elasticity of reaction with respect to  $A$ .

We must first select an operating point for  $A$  and  $B$ . This choice will depend on the system under study. For example, perhaps we are interested in the value of the substrate elasticity for an enzyme catalyzed reaction when the substrate and product concentration are at their  $K_m$  levels. Once the operating point is chosen, the reaction is started and the rate of reaction measured. It is important that during the measurement, only a small amount of substrate is consumed and product produced, otherwise the estimate for the elasticity will not be accurate. We now begin the experiment again but this time the substrate concentration is increased by a small amount, and the product concentration is reset to its value in the first experiment. The reaction is started and the new reaction rate measured. The fractional change in reaction rate and substrate is recorded and the ratio computed to give the substrate elasticity. In principle, the same kind of experiment could be performed on the product, this time keeping the substrate concentration constant.

### Simple protocol for estimating the substrate elasticity

1. Set substrate and product concentrations to their operating points.
2. Record the reaction rate at the operating point.
3. Restore all concentrations to their original starting points.
4. Increase the concentration of substrate by a small amount.
5. Record the new reaction rate.
6. Compute the elasticity by dividing the fractional change in reaction rate by the fractional change in substrate concentration.
7. At all times, maintain other substrate, product and effector concentrations at the operating point.

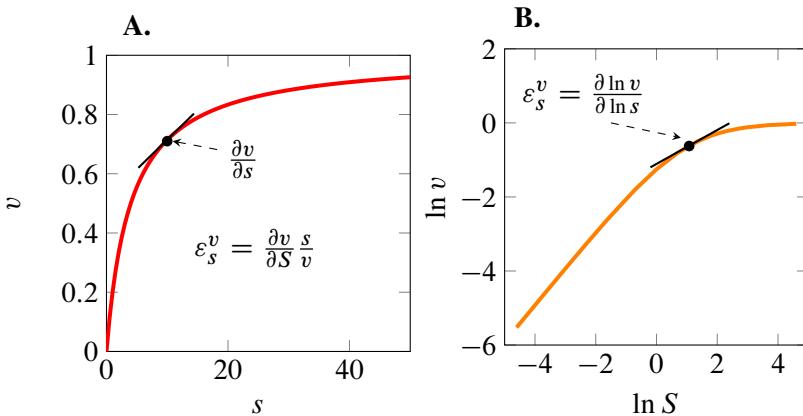
The algebraic definition of the elasticity automatically suggests ways to estimate their values that include algebraic differentiation of the rate law (if the rate laws is available), and numerical computation of values by simulation or even by experiment. In the next section the use of log/log plots will be used to empirically determine an elasticity from an enzyme kinetics experiment.

### Log Form

The definition of the elasticity in equation (6.2) shows the elasticity expressed using a log notation:

$$\varepsilon_s^v = \frac{\partial \ln v}{\partial \ln s}$$

This notation is frequently used in the literature. In Figure 6.2 the left panel is a simple plot of rate versus substrate concentration and shows a typical response for an enzyme as it



**Figure 6.2** **A.** The slope of the reaction rate versus the reactant concentration scaled by both the reactant concentration and reaction rate yields the elasticity,  $\varepsilon_s^v$ . **B.** If the log of the reaction rate and log of the reactant concentration are plotted, the elasticity can be read directly from the slope of the curve. Curves are generated by assuming  $v = s/(2 + s)$ .

saturates at high substrate concentration. The right-hand panel shows the same plot but now the axes are logarithmic. As we will see, the slope of the curve on the right-panel is a direct measure of the elasticity. What it clearly shows is that the slope (or elasticity) changes as the substrate changes. The origin of this effect is worth explaining to those unfamiliar with the calculus.

If we examine the growth pattern of a micro-organism, we will often find that it follows a pattern of the kind,  $y = a^x$ . What this means is that the number of microorganisms increases by a fixed proportion per unit time. Often such data is plotted on a semi-logarithmic scale rather than the usual linear scale as it helps to emphasize the fact that the relative growth under these conditions is the same throughout the growth phase. To explain this statement, a numerical example will be useful. Of the two sequences of numbers:

$$\begin{array}{ccccccc} 100, & 150, & 200, & 250, & 300, & \dots \\ 100, & 150, & 225, & 337.5, & 506.25, & \dots \end{array}$$

the first shows a regular increase of 50 units and the second a regular increase of 50 percent from one number to the next. On a linear scale, the points representing the first sequence appear as equal distances from each other and those representing the second sequence at increasing distances. If instead, we take the logarithms to base 10 of these numbers as in the following sequence:

$$\begin{array}{ccccccc} 2, & 2.176, & 2.301, & 2.398, & 2.477, & \dots \\ 2, & 2.176, & 2.352, & 2.528, & 2.704, & \dots \end{array}$$

then on the logarithmic scale, it is the second sequence that gives points at equal distances

from each other while the first sequence shows points at decreasing distances along the axis. It would seem, therefore, that equal distances between points on a linear scale indicate equal *absolute* changes in the variable, and equal distances between points on a logarithmic scale indicate equal *proportional* changes in the variable. Before taking the logarithm, the second sequence increased by 50% each time. In log form however, it increased by a constant absolute amount of 1.176.

More formally we can describe this effect as follows. Consider a variable  $y$  to be some function  $f(x)$ , that is  $y = f(x)$ . If  $x$  increases from  $x$  to  $(x + h)$ , then the change in the value of  $y$  will be given by  $f(x + h) - f(x)$ . The **proportional** change however, is given by:

$$\frac{f(x + h) - f(x)}{f(x)}$$

The **rate of proportional change** at the point  $x$  is given by the above expression divided by the step change in the  $x$  value, namely  $h$ :

Rate of proportional change =

$$\lim_{h \rightarrow 0} \frac{f(x + h) - f(x)}{hf(x)} = \frac{1}{f(x)} \lim_{h \rightarrow 0} \frac{f(x + h) - f(x)}{h} = \frac{1}{y} \frac{dy}{dx}$$

From calculus we know that  $d \ln y / dx = (1/y) dy / dx$ , therefore the rate of proportional change equals:

$$\frac{d \ln y}{dx}$$

This is a measure of the rate of *proportional* change of the variable  $y$ , or function  $f(x)$ . Just as  $dy/dx$  measures the gradient of the curve,  $y = f(x)$  plotted on a linear scale,  $d \ln y / dx$  measures the slope of the curve when plotted on a semi-logarithmic scale, that is the rate of proportional change. For example, a value of 0.05 means that the curve increases at 5% per unit  $x$ .

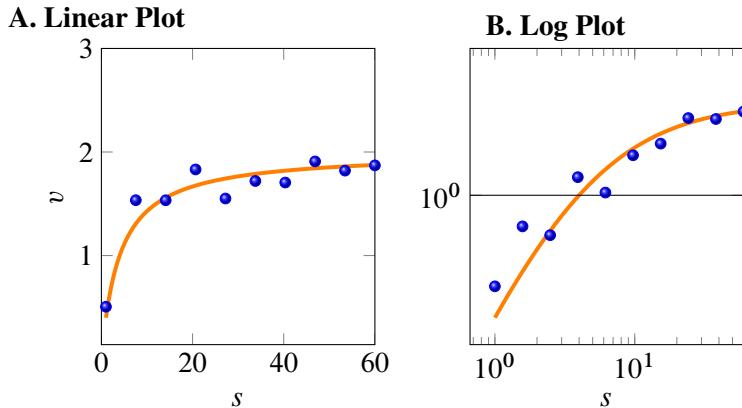
We can apply the same argument to the case when we plot a function on both  $x$  and  $y$  logarithmic scales. In such a case, the following result is true:

$$\frac{d \ln y}{d \ln x} = \frac{x}{y} \frac{dy}{dx}$$

This shows the relationship between the log form and non-log form of the elasticity.

If experimental data is derived from a rate experiment on an isolated enzyme, the data can be plotted in log space (Figure 6.3) and the elasticity read directly from the slope of the line.

As already mentioned, the log form of the elasticity expresses the ratio of relative changes. In approximate terms, we can say that for an  $x$  % change in the concentration of a molecular species, the elasticity will give the percentage change,  $v$  %, in the reaction rate. For this reason the elasticity is sometimes expressed as a ratio of percentage changes:



**Figure 6.3 A:** Plot of reaction rate versus substrate concentration. Measurements include errors. **B:** Same data but plotted in log space. The elasticity can be read directly from the slope of the curve. Curves are generated by assuming  $v = 2s/(4 + s)$ .

$$\varepsilon_{s_i}^v \approx \frac{\% \text{ change in } v}{\% \text{ change in } s_i} \quad (6.4)$$

For example, if the concentration of a substrate is increased from 1.5 mM to 1.95 mM, then the percentage increase in substrate concentration is 30 %. If at the same time, the reaction rate of the enzyme increased from  $55 \mu M g^{-1} min^{-1}$  to  $12 \mu M g^{-1} min^{-1}$ , then the percentage increase in rate must be 24 %. Therefore the elasticity can be estimated approximately from the ratio 24/30, which is equal to 0.8; that is, the enzyme rate changes almost in proportion to a change in substrate. If the enzyme were acting *in vivo* and a disturbance upstream caused the concentration of substrate to rise, then this enzyme would respond by increasing its rate almost in proportion to the change in substrate concentration.

### 6.3 Mass-action Kinetics

Computing the elasticities for mass-action kinetics is straightforward. For a reaction such as  $v = ks$ , it was shown earlier (6.1) that  $\varepsilon_s^v = 1$ . For a generalized irreversible mass-action law such as:

$$v = k \prod s_i^{n_i}$$

the elasticity for species  $s_i$  is  $n_i$ . For simple mass-action kinetic reactions, the kinetic order and elasticity are therefore identical and independent of species concentration.

For a simple irreversible mass-action reaction rate law such as:

$$v = k_1 s - k_2 p \quad (6.5)$$

The elasticities for the substrate and product can be determined as before by differentiating and scaling:

$$\varepsilon_s^v = \frac{k_1 s}{k_1 s - k_2 p} = \frac{v_f}{v} \quad (6.6)$$

$$\varepsilon_p^v = -\frac{k_2 p}{k_1 s - k_2 p} = -\frac{v_r}{v} \quad (6.7)$$

In the above equations  $v_f$  is the forward rate,  $v_r$  is the reverse rate, and  $v$  is the net rate. Note that  $\varepsilon_s^v$  is **positive** and  $\varepsilon_p^v$  **negative**.  $\varepsilon_p^v$  is negative because increases in product concentration will slow down the net forward rate.

If we divide top and bottom of equation (6.6) by  $k_1$  and  $s$ , and equation (6.7) by  $k_2$  and  $p$ , and noting that the ratio  $k_1/k_2 = K_{eq}$  (1.11),  $p/s = \Gamma$  (1.14), and  $\Gamma/K_{eq} = \rho$  (1.13), we can express the elasticities in the form:

$$\begin{aligned} \varepsilon_s^v &= \frac{1}{1 - \Gamma/K_{eq}} = \frac{1}{1 - \rho} \\ \varepsilon_p^v &= -\frac{\Gamma/K_{eq}}{1 - \Gamma/K_{eq}} = -\frac{\rho}{1 - \rho} \end{aligned} \quad (6.8)$$

These expressions can vary over a wide range of values. Far from equilibrium ( $\rho \approx 0$ )  $\varepsilon_s^v$  will lie close to 1.0, while  $\varepsilon_p^v$  will be close to -0.0. When operating close to equilibrium however ( $\rho \approx 1$ ), the same elasticities will tend to  $+\infty$  and  $-\infty$ , respectively. This behavior is depicted in Figure 6.4.

Of interest is the following relation, see equation (1.15):

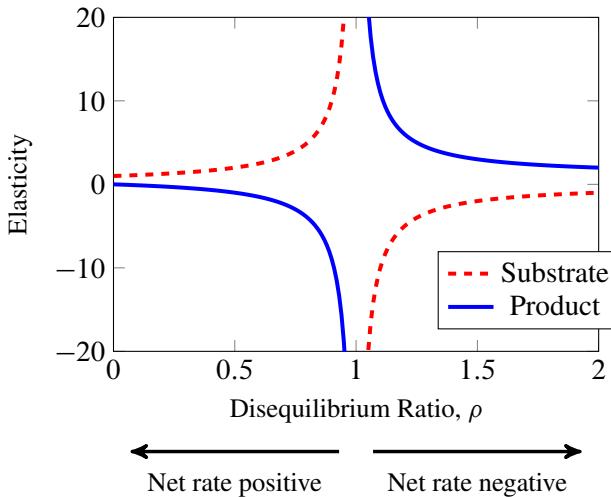
$$\frac{\varepsilon_p^v}{\varepsilon_s^v} = \frac{v_r}{v_f} = -\rho$$

which connects the ratio of the elasticities to the disequilibrium ratio. It also follows from the above equations that the sum of the elasticities for mass-action kinetic rate laws is always one:

$$\varepsilon_s^v + \varepsilon_p^v = 1 \quad (6.9)$$

This means that if one of the elasticities is known, the other can be easily determined by subtraction.

Equation (6.9) is significant for another reason. Since  $\varepsilon_p^v$  is negative, the absolute magnitude of  $\varepsilon_s^v$  will **always** be larger than the absolute value for  $\varepsilon_p^v$  when dealing with mass-action kinetics. That is:



**Figure 6.4** Elasticities as a function of the disequilibrium ratio,  $\rho$ .

$\rho$	$\varepsilon_s^v = 1/(1-\rho)$	$\varepsilon_p^v = -\rho/(1-\rho)$
0.9	10	-9
0.5	2	-1
0.2	1.2	-0.25
0.1	1.111	-0.111

**Table 6.1** Selected values for the elasticities and the disequilibrium ratio,  $\rho$ : Note :  $\|\varepsilon_s^v\| > \|\varepsilon_p^v\|$ .

$$\|\varepsilon_s^v\| > \|\varepsilon_p^v\|$$

For small elasticity values, the relative difference between the elasticities can be significant. This means that changes in substrate concentrations will have a much greater effect on the reaction velocity than changes in product concentrations.

For simple mass-action kinetics, changes in substrate concentrations will have a much greater effect on the reaction velocity than changes in product concentrations.

The propagation of signals along a pathway is **determined** by the elasticity values. Given that substrate elasticities are larger than product elasticities, signal propagation tends to amplify when traveling downstream compared to signals traveling upstream which tend to be attenuated. For the general reversible mass-action rate law:

$$v = k_1 \prod s_i^{n_i} - k_2 \prod p_i^{m_i} \quad (6.10)$$

The elasticities can be shown to equal:

$$\begin{aligned}\varepsilon_{s_i}^v &= \frac{n_i}{1 - \rho} \\ \varepsilon_{p_i}^v &= -\frac{m_i \rho}{1 - \rho}\end{aligned}\quad (6.11)$$

## 6.4 Enzyme Catalyzed Reactions

The irreversible Briggs-Haldane equation is given by:

$$v = \frac{V_m s}{K_m + s}$$

where  $V_m$  is the maximal rate and  $K_m$  the substrate concentration that yields half the maximal rate. It is straightforward to determine the algebraic elasticity with respect to the substrate concentration. The derivative  $\partial v / \partial s$  is given by:

$$\frac{\partial v}{\partial s} = \frac{V_m K_m}{(K_m + s)^2}$$

Scaling by  $v$  and  $s$  yields the elasticity equation:

$$\varepsilon_s^v = \frac{K_m}{K_m + s}$$

The substrate elasticity shows a range of values (Figure 6.5) from zero at high substrate concentrations to one at low substrate concentrations. When the enzyme is near saturation it is naturally unresponsive to further changes in substrate concentration, hence the elasticity is near zero. The reaction behaves as a zero-order reaction at this point. When the elasticity is close to one at low  $s$ , the reaction behaves with first-order kinetics. In addition, the reaction order changes depending on the substrate concentration.

It is interesting to note that when  $s = K_m$ , the elasticity is equal to one half.

**Enzyme Elasticity.** We can also compute the elasticity with respect to enzyme concentration since *in vivo* enzyme concentrations can change. Given that  $V_m = e_t k_{cat}$ , where  $e_t$  is the total enzyme concentration and  $k_{cat}$  the catalytic constant, the enzyme elasticity is derived as follows:

$$\frac{\partial v}{\partial e_t} = \frac{k_{cat} K_m}{K_m + s}$$

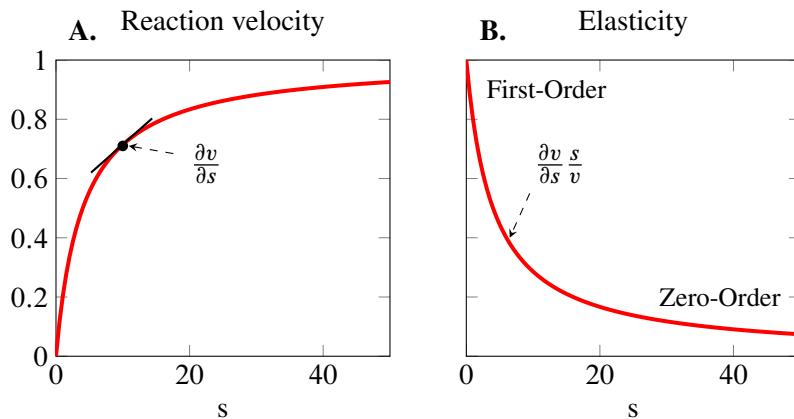
Scaling by  $e_t$  and  $v$  yields:

$$\frac{\partial v}{\partial e_t} \frac{e_t}{v} = \frac{k_{\text{cat}} s}{K_m + s} e_t \frac{K_m + s}{e_t k_{\text{cat}} s} = 1$$

Hence the **enzyme elasticity is one**:

$$\varepsilon_e^v = 1 \quad (6.12)$$

This is not surprising because the reaction rate is first-order with respect to the enzyme concentration.



**Figure 6.5** A. Left panel: the reaction velocity for an irreversible Michaelis-Menten rate law as a function of substrate concentration. The curve is also marked by the slope  $\partial v / \partial s$ . B. Right panel: the substrate elasticity is plotted as a function of substrate concentration.  $K_m = 4$  and  $V_m = 1$ . Note the elasticity starts at one, then decreases to zero as  $s$  increases.

For the reversible Briggs-Haldane equation, the pattern of elasticities is more complex. It is still however the case that  $\varepsilon_e^v = 1$ . Details can be found in the companion book [109] but Table 6.2 summarizes the results. Two aspects are worth pointing out. Near equilibrium, as with simple mass-action kinetics, the substrate and product elasticities approach positive and negative infinity, respectively. Secondly, there are subtle competition effects between the substrate and product that effect the substrate and product elasticities.

## 6.5 Cooperativity

Given the Hill equation:

$$v = \frac{V_m s^n}{K_d + s^n} = \frac{V_m s^n}{K_H^n + s^n} \quad (6.13)$$

**Table 6.2 Values of Elasticities Depending on Saturation and Equilibrium Conditions**

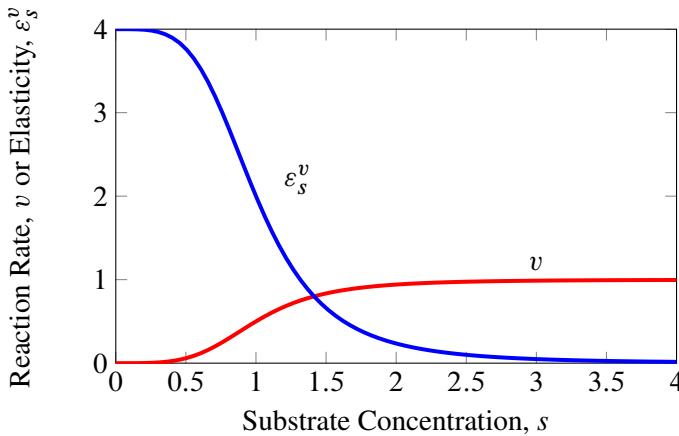
Equilibrium State	Degree of Saturation	Elasticities
Near Equilibrium	All degrees of saturation	$\varepsilon_s^v \gg 1; \varepsilon_p^v \ll -1; \varepsilon_s^v + \varepsilon_p^v \approx 1$
Far from Equilibrium	$s \ll K_s$ and $p \ll K_p$	$\varepsilon_s^v \approx 1; \varepsilon_p^v \approx 0$
Far from Equilibrium	$s \gg K_s$ and $p \ll K_p$	$\varepsilon_s^v \approx 0; \varepsilon_p^v \approx 0$
Far from Equilibrium	$p \gg K_p$ (Any substrate level)	$\varepsilon_s^v \approx -\varepsilon_p^v$
	if $s/K_s \ll p/K_p$	$\varepsilon_s^v \approx 1 \quad \varepsilon_p^v \approx -1$
	$s/K_s \approx p/K_p$	$\varepsilon_s^v \approx 0.5 \quad \varepsilon_p^v \approx -0.5$
	$s/K_s > p/K_p$	$\varepsilon_s^v < 0.5 \quad \varepsilon_p^v > -0.5$

where  $K_H$  is the concentration of ligand that yields half the maximal rate. The elasticity coefficient,  $\varepsilon_s^v$  may be derived directly from the Hill equation (6.13). Differentiating and scaling the Hill equation yields the following elasticity both in terms of the dissociation constant,  $K_d$  and the half maximal activity constant,  $K_H$ :

$$\varepsilon_s^v = \frac{n K_d}{K_d + s^n} = \frac{n}{1 + \left(\frac{s}{K_H}\right)^n} \quad (6.14)$$

The elasticity of a reaction obeying the Hill equation has a value equal to  $n$  at low substrate concentrations ( $s \ll K_d$ ). In contrast, irreversible Michaelian enzymes at low substrate concentrations have an elasticity value of one. Therefore an enzyme obeying the Hill equation shows a much higher elasticity to the substrate concentration compared to a Michaelian enzyme. Like a Michaelian enzyme, the value of the elasticity falls off rapidly as the substrate concentration increases, reaching zero as the enzyme becomes saturated. Figure 6.6 illustrates this response for  $n = 4$  and  $K_d = 1$ . An interesting feature in Figure 6.6 is the delayed fall in the elasticity at low substrate concentrations.

Similar equations can be derived for other cooperative models as well as allosteric control. The important message is that cooperative and allosteric control illicite high values for



**Figure 6.6** Plot showing the response of the rate and elasticity for the Hill model, with  $n = 4$  and  $K_d = 1$ .

elasticity coefficients which can have a significant impact on pathway behavior. Some of these effects will be covered in detail in later chapters.

Much more detail on the elasticities of a great variety of rate laws can be found in the companion book [109].

## 6.6 Local Equations

The elasticity coefficient is of central importance to metabolic control analysis (MCA). Just as the Michaelian constants are essential to describing the rate of an enzyme-catalysed reaction, the elasticities are equally essential to describing the behavior of whole pathways. Before we discuss this topic in the next chapter, some direct uses of the elasticities will be given here.

Recall that the elasticity coefficient is given by:

$$\varepsilon_s^v = \frac{\partial v}{\partial s} \frac{s}{v}$$

This definition can be rearranged and an approximate equation written in the form:

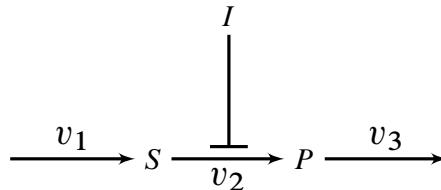
$$\frac{\delta v}{v} \approx \varepsilon_s^v \frac{\delta s}{s}$$

This relation is approximate because the changes considered are finite, and the definition of an elasticity applies strictly to infinitesimal changes. The equation describes how, given a fractional change in some effector S, the resulting fractional change in rate can be computed.

For example, if the elasticity of an enzyme reaction towards an effector S is 0.8, then given a fractional change in  $s$  of 0.05 (a 5% change in  $s$ ), the fractional change in rate is given by:

$$0.8 \times 0.05 = 0.04$$

In other words, a 5% change in  $s$  leads to a 4% change in reaction rate.



**Figure 6.7** Species  $I$  inhibits reaction  $v_2$  in addition to potential affects from  $S$  and  $P$ .

The diagram (Figure 6.7) shows a fragment from a larger pathway. The central reaction step has three effectors which could potentially change the rate  $v_2$  – these are  $S$ ,  $P$  and an inhibitor,  $I$ . Let us consider a disturbance<sup>4</sup> somewhere in the pathway but *not* originating at the reaction step under consideration. This disturbance will ultimately cause changes in each of the effectors by amounts  $\delta s$ ,  $\delta p$  and  $\delta i$ . These changes will also be accompanied by a change to the reaction rate by an amount,  $\delta v$ . There are two immediate questions we can ask: (i) what is the relationship between the change in the effectors and the change in rate? and, (ii) what is the contribution that each change in effector makes to the final change in rate?

The answers to these questions are straightforward to obtain. Provided that the changes are small, then the fractional change in rate,  $\delta v/v$ , is defined by the sum of the individual contributions:

$$\frac{\delta v}{v} \approx \varepsilon_s^v \frac{\delta s}{s} + \varepsilon_p^v \frac{\delta p}{p} + \varepsilon_I^v \frac{\delta i}{i}$$

For example, let us assume the following values for the elasticities,  $\varepsilon_s^v = 0.4$ ;  $\varepsilon_p^v = -0.5$ ;  $\varepsilon_i^v = -0.2$ , and assuming the following changes in effectors,  $\delta s/s = 0.05$ ;  $\delta p/p = 0.03$ ;  $\delta i/i = 0.01$ , then the fractional change in rate through the step is given by:

$$\frac{\delta v}{v} \approx 0.4 \times 0.05 + (-0.5) \times 0.03 + (-0.2) \times 0.01 = 0.003$$

The rate has only changed by 0.3 %, much of the potential increase that could have been obtained by the change in  $S$  has been reduced by strong product inhibition. To answer the question – what contribution does each effector make to the final change in rate – is

<sup>4</sup>This could be due to one of a number of causes, a change in enzyme expression, nutrient supply change, hormonal change, etc.

given simply by examining the individual changes. Thus out of the total absolute change in rate, the change brought about by S contributed 54 %, while the change in product and inhibitor contributed -41 % and -5 %, respectively. Clearly the change in inhibitor was not an important factor.

The degree to which each modifier has an affect on the reaction rate is an important consideration. Although in pathway diagrams we may see many feedback loops, we rarely see the quantitative contribution that the modifier makes in relation to other potential modifiers.

In general, for a reaction step embedded in a pathway and acted upon by  $m$  modifiers, the change in rate due to changes in all modifiers is given by the relation:

$$\frac{dv}{v} = \sum_{j=1}^m \varepsilon_{s_j}^v \frac{ds_j}{s_j}$$

where the symbol,  $\sum$  means ‘sum of’. In the equation the small but finite changes have been replaced by differentials so that the relation is exact. If the concentration of enzyme is also changed, then we may also add the enzyme elasticity to the sum, as in:

**Local Equation:**

$$\frac{dv}{v} = \sum_{j=1}^m \varepsilon_{s_j}^v \frac{ds_j}{s_j} + \varepsilon_e^v \frac{de}{e} \quad (6.15)$$

This relation is probably one of the most important mathematical relations used in MCA, and we will come across its application in subsequent chapters. In mathematical terms, the relationship is a modification of the standard total derivative. This states that if  $y$  is some function  $f$  of  $m$  variables,  $x_i$ , then:

$$y = f(x_1, x_2, \dots, x_m),$$

The total derivative of  $y$  is therefore given by:

$$dy = \frac{\partial y}{\partial x_1} dx_1 + \frac{\partial y}{\partial x_2} dx_2 + \dots + \frac{\partial y}{\partial x_m} dx_m$$

Scaling both sides of this equation will lead a form similar to the locate equation (6.15).

## 6.7 General Elasticity Rules

Just as there are rules for differential calculus, there are similar rules for computing elasticities. These rules can be used to simplify the derivation of elasticities for complex rate law expressions. Table 6.3 shows some common elasticity rules, where  $a$  designates a constant, and  $x$  the variable. For example, the first rule says that the elasticity of a constant is zero.

- 
1.  $\varepsilon(a) = 0$
  2.  $\varepsilon(x) = 1$
  3.  $\varepsilon(f(x) \pm g(x)) = \varepsilon(f(x)) \frac{f'(x)}{f(x)+g(x)} \pm \varepsilon(g(x)) \frac{g'(x)}{f(x)+g(x)}$
  4.  $\varepsilon(x^a) = a$
  5.  $\varepsilon(f(x)^a) = a\varepsilon(f(x))$
  6.  $\varepsilon(f(x)g(x)) = \varepsilon(f(x)) + \varepsilon(g(x))$
  7.  $\varepsilon(f(x)/g(x)) = \varepsilon(f(x)) - \varepsilon(g(x))$
- 

**Table 6.3** Transformation rules for determining the elasticity of a function,  $a = \text{constant}$ ,  $x = \text{variable}$ .

We can illustrate the use of these rules with a simple example. Consider the reversible mass-action rate law (6.5):

$$v = k_1 s - k_2 p$$

To determine the elasticity we first apply rule 3 to give:

$$\varepsilon_s^v = \varepsilon_s(k_1 s) \frac{k_1 s}{k_1 s - k_2 p} - \varepsilon_s(k_2 p) \frac{-k_2 p}{k_1 s - k_2 p}$$

where  $\varepsilon_s(f)$  means the elasticity of expression  $f$  with respect to variable  $s$ .

We transform the elasticity terms by applying additional rules. Apply rule 6 to the expression  $\varepsilon_s(k_1 s)$  to give:

$$\varepsilon_s(k_1 s) = \varepsilon_s(k_1) + \varepsilon_s(s)$$

We can now apply rule 1 to the first term on the right, and rule 2 to the second term on the right to give:

$$\varepsilon_s(k_1 s) = 0 + 1$$

Since we're evaluating the elasticity of  $s, p$  in this situation is a constant, therefore:

$$\varepsilon_s(k_2 p) = \varepsilon_s(k_2) + \varepsilon_s(p) = 0 + 0$$

Combining these results yields:

$$\varepsilon_s^v = \frac{k_1 s}{k_1 s - k_2 p}$$

which corresponds to the first equation in (6.6).

Now consider a simple irreversible enzyme kinetic rate equation:

$$v = \frac{V_m s}{K_m + s}$$

where  $V_m$  is the maximal velocity and  $K_m$  the substrate concentration at half maximal velocity.

The elasticity for this equation can be derived by first using the quotient rule (rule 7) which gives:

$$\varepsilon_s^v = \varepsilon(V_m s) - \varepsilon(K_m + s)$$

The rules can now be applied to each of the sub-elasticity terms. For example, we can apply rule 6 to the first term,  $\varepsilon(V_m s)$ , and rule 3 to the second term,  $\varepsilon(K_m + s)$ , to yield:

$$\varepsilon_s^v = (\varepsilon(V_m) + \varepsilon(s)) - \left( \varepsilon(K_m) \frac{K_m}{K_m + s} + \varepsilon(s) \frac{s}{K_m + s} \right)$$

Applying rules 1 and 2 allows us to simplify ( $\varepsilon(V_m) = 0; \varepsilon(K_m) = 0; \varepsilon(s) = 1$ ) the equation to:

$$\varepsilon_s^v = 1 - \left( \frac{s}{K_m + s} \right)$$

or:

$$\varepsilon_S^v = \frac{K_m}{K_m + S}$$

### **Example 6.5**

---

Determine the elasticity expression for the rate laws using log-log rules:

1.  $v = k(A + 1)$

Begin with the product rule 6:

$$\varepsilon_A^v = \varepsilon(k) + \varepsilon(A + 1) = \varepsilon(A + 1)$$

Next use the summation rule 3 and rule 2:

$$\begin{aligned} \varepsilon_a^v &= \varepsilon(a + 1) = \varepsilon(A) \frac{a}{a + 1} + \varepsilon(1) \frac{1}{a + 1} \\ &= \frac{a}{a + 1} + 0 = \frac{a}{a + 1} \end{aligned}$$

2.  $v = k/(a + 1)$

Begin with the quotient rule 6 followed by Rule 3 and 2:

$$\begin{aligned} \varepsilon_a^v &= \varepsilon(k) - \varepsilon(a + 1) = -\frac{1}{a + 1} \\ &= -\frac{a}{a + 1} \end{aligned}$$

3.  $v = a(a + 1)$

Begin with the quotient rule 6:

$$\varepsilon_a^v = \varepsilon(a) + \varepsilon a + 1$$

Next use Rule 2, 3 and 2:

$$\varepsilon_a^v = 1 + \frac{1}{a+1}$$


---

To make matters even simpler, we can define the elasticity rules using an algebraic manipulation tool such as Mathematica (<http://www.wolfram.com/>) to automatically derive the elasticities [152]. To do this we must first enter the rules in Table 6.3 into Mathematica. The script shown in Figure 6.8 shows the same rules (with a few additional ones) in Mathematica format.

```
(* Define elasticity evaluation rules *)
el[x_, x_] := 1
el[k_, x_] := 0
el[Log[u_, x_]] := el[Log[u], x] = el[u, x]/Log[u]
el[Sin[u_], x_] := el[Sin[u], x] = u el[u, x]Cos[u]/Sin[u]
el[Cos[u_], x_] := el[Sin[u], x] = -u el[u, x]Sin[u]/Cos[u]
el[u_*v_, x_] := el[u*v, x] = el[u, x] + el[v, x]
el[u_/v_, x_] := el[u/v, x] = el[u, x] - el[v, x]
el[u_+v_, x_] := el[u+v, x] = el[u, x]u/(u+v) + el[v, x]v/(u+v)
el[u_-v_, x_] := el[u-v, x] = el[u, x]u/(u-v) - el[v, x]v/(u-v)
el[u^v_, x_] := el[u^v, x] = v (el[u, x] + el[v, x] Log[u])
```

**Figure 6.8** Elasticity rules expressed as a Mathematica script.

The notation `f[x_, y_] := g()` means define a function that takes two arguments, `x_` and `y_`. The underscore character in the argument terms is essential. Note also the symbol ‘`:`’ in the assignment operator.

Typing `el[k1 s - k2 p, s]` into Mathematica will result in the output:

`k1 s / (-k2 p + k1 s)`

## 6.8 Summary

---

The elasticity coefficient is a measure of how sensitive the rate of a reaction is to changes in its environment. The factors of interest are the concentrations of modifiers such as substrates, products, and effectors and the concentration of enzymes.

There will be as many elasticities as there are modifiers of the reaction. The elasticity is strictly defined in terms of a partial derivative, which means that it measures the change in rate when one modifier is changed. For example, the substrate elasticity is measured when all other modifiers are held constant except for the substrate concentration. Algebraically

this is achieved by partial differentiation, and experimentally by clamping the appropriate modifier concentrations.

The elasticity coefficient can be written in various equivalent forms, each reflecting a different emphasis:

$$\varepsilon_{s_j}^v = \left( \frac{\partial v_i / v_i}{\partial s_j / s_j} \right)_{s_k, s_l, \dots} = \frac{s_j}{v_i} \left( \frac{\partial v_i}{\partial s_j} \right)_{s_k, s_l, \dots} = \left( \frac{\partial \ln v_i}{\partial \ln s_j} \right)_{s_k, s_l, \dots}$$

The first form is the ratio of fractional changes, the second form the scaled slope on a linear plot, and the third form the slope of a log/log plot.

The notation,  $s_k, s_l \dots$  means that these modifiers are held constant during the measurement of the partial derivative. An approximate form of the elasticity is given by:

$$\varepsilon_s^v \approx \frac{\% \text{ change in } v}{\% \text{ change in } s}$$

which can be used to estimate an elasticity when changes in reaction rate and modifier are known. Elasticities have a number of important properties:

- The elasticity coefficient is **not** a constant but depends on the concentrations of all modifiers that might affect the reaction rate. An elasticity is not like a  $K_m$  or  $K_i$ ; the Michaelian constants are characteristic for a particular enzyme and modifier, reflecting the enzymes' kinetic mechanism and interaction energy with the modifier. Kinetic constants do not in general depend on the concentrations of the modifiers; elasticities do.
- In general, an elasticity is a function of both the kinetic characteristics of an enzyme, and the concentration of all the various modifiers that might interact with the enzyme.
- For the standard irreversible Michaelian mechanism, the elasticity of a substrate at saturating levels is zero, and when the substrate is below its  $K_m$ , the elasticity is unity. When the substrate concentration is equal to the  $K_m$ , the elasticity has a value of 0.5.

Given a change in concentration of a modifier, it is possible to use the elasticity coefficient to predict (approximately) the change in rate, thus:

$$\frac{\delta v}{v} \approx \varepsilon_s^v \frac{\delta s}{s}$$

If more than one modifier is changing at a time, then the approximate change in rate is given by the sum of the individual contributions:

$$\frac{\delta v}{v} \approx \sum_j \varepsilon_{s_j}^v \frac{\delta s_j}{s_j}$$

It is very important to appreciate that the elasticities used in the above equation must be measured at the prevailing state of the modifiers. It makes no sense to use an elasticity that has previously been measured at a substrate concentration of 2mM, and then to use the same elasticity *value* at a substrate concentration of 20mM.

## Further Reading

---

1. Fell D A (1996) Understanding the Control of Metabolism. Portland Press, ISBN: 185578047X
2. Heinrich R and Schuster S (1996) The Regulation Of Cellular Systems. Springer; 1st edition, ISBN: 0412032619
3. Sauro HM (2012) Enzyme Kinetics for Systems Biology. 2nd Edition, Ambrosius Publishing ISBN: 978-0982477335

## Exercises

---

1. What is the relevance of elasticity coefficients in understanding network dynamics?
2. State the operational interpretation of an elasticity.
3. Why is the elasticity coefficient expressed in terms of a partial derivative? What does it mean in terms of an experimental operation?
4. An experiment indicates that a given molecule  $x$  has an elasticity of -0.5 with respect to the rate of a reaction. State two key aspects that this elasticity describes.
5. What is the elasticity with respect to the species  $a$  given the rate law  $v = ka^3$ ?
6. Work out algebraically the elasticity for the rate law,  $v = k_1 s + k_2$ . Describe its properties at high and low levels of  $s$ .
7. Derive the elasticity expression with respect to  $x$  for the following:
  - a)  $v = x^2 + 1$
  - b)  $v = x^2 + x$
  - c)  $v = x/(x^2 + 1)$
8. Describe one technique for numerically estimating an elasticity.
9. Given a change in  $\delta v/v$  equal to 0.04, and if  $\varepsilon_s^v = 0.1$ , what is the change in  $\delta s/s$ ? If the concentration of  $s$  was 2.5mM, what is the absolute change in  $s$ ?

10. If the concentration of  $s$  is 3mM, and is the elasticity of an enzyme with a rate law,  $V_{max} s / (K_m + s)$  is 0.6, what is the  $K_m$  of the enzyme? What would be the elasticity at 8mM? Why does the elasticity change?
11. What does the term  $\Gamma / K_{eq}$  measure?
12. Describe the value of disequilibrium ratio as a reaction nears equilibrium.
13. For a mass-action reversible reaction, describe what happens to the substrate and product elasticities as the reaction approaches equilibrium.
14. Derive the two equations in (6.8).
15. Describe the significance of equation (6.9).
16. Using the elasticity rules in Table 6.3, derive the elasticity for the following equation indicating all intermediate steps.

$$v = s^n / (K_m + s^n)$$

# 7

## ***Introduction to Biochemical Control***

### **7.1 What do we mean by Control?**

---

For most people the word control means the ability to influence, command or to restrain a situation or process<sup>1</sup>. In this chapter the term control will be used in a similar sense to describe how much influence a given reaction step in a network has on the system. To make matters simpler, the system will be considered at steady state so that control will refer to how much influence a given reaction step has on the steady state. That is, how fluxes and concentrations are influenced. The measure of much control a reaction step has over the steady state will be called the **control coefficient**. Our initial definition of a control coefficient is as follows:

The amount of control (or influence) that a particular reaction step has on a flux or species concentrations is called the **control coefficient**.

Most reaction steps in a cell are controlled by proteins. One question to ask is, how much influence does a given protein have on the system's steady state? Experimentally, such control can be measured by changing the concentration of an enzyme or changing its activity via an inhibitor and measuring the effect on the steady state flux and species concentrations. We can change the concentration of a protein in various ways such as using irreversible inhibitors, changing the promoter consensus sequence on the gene that codes for the protein,

---

<sup>1</sup>In engineering, control theory refers to the body of knowledge concerned with the design and study of systems that can perform specific tasks or achieve a particular objective.

employing antisense RNA to reduce the expression level, or using dCas9 guided by a specific guide RNA.

In addition to investigating how individual reaction steps control the fluxes and concentrations in a network, we are also interested in how external factors influence the network. Examples of external factors include the level of nutrients, hormones, and of particular interest to human health, therapeutic drugs. In these situations, rather than using the word control, we will use the term response. Thus, a biological cell will have a response to a particular infusion of a drug. The degree of influence an external factor has on a biological system will be described using **response coefficients**.

The degree of influence a particular external input has on a pathway is called the **response coefficient**.

We will consider response coefficients in a later chapter, the focus here will be on control coefficients.

### A Reminder on Notation

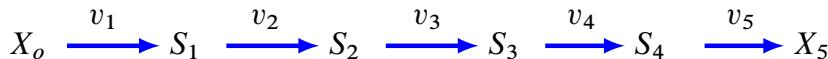
It is worth describing again the notation that will be used in this and subsequent chapters.

1. Upper case letters such as  $S$  or  $X$  refer to the name of a molecular species.
2. Lower case letters such as  $s$  or  $x$  refer to the concentration of the molecules species.
3. The upper case letter  $E$  refers to the name of an enzyme.
4. The lower case letter  $e$  refers to the concentration of the enzyme,  $E$ .
5. The symbol  $\Delta$  means a change, for example  $\Delta s$  means a change in the concentration of  $S$ .
6. The symbol  $\delta$  means a small change, for example  $\delta e_1$  means a small change to enzyme  $E_1$ .
7. The symbol  $J$  refers to the steady state flux through a pathway.
8. The symbol  $v_i$  refers to the rate of reaction through the  $i^{th}$  reaction step.

## 7.2 Control Coefficients

---

Control coefficients are used to describe how much influence (i.e. control) a given reaction step has on the steady state flux or species concentration level. It is common to measure



**Figure 7.1** Five step linear pathway.  $X_o$  and  $X_5$  are assumed to be fixed in the pathway model.

this influence by changing the concentration of the enzyme that catalyzes the reaction. To describe control coefficients in more detail, let us consider a thought experiment.

The following discussion will be centered on the simple linear pathway shown in Figure 7.1. Let us assume that the species pools,  $S_1$  to  $S_4$  in the pathway are empty (zero concentration), and that  $X_o$  and  $X_5$  are *fixed* species forming the system boundary.

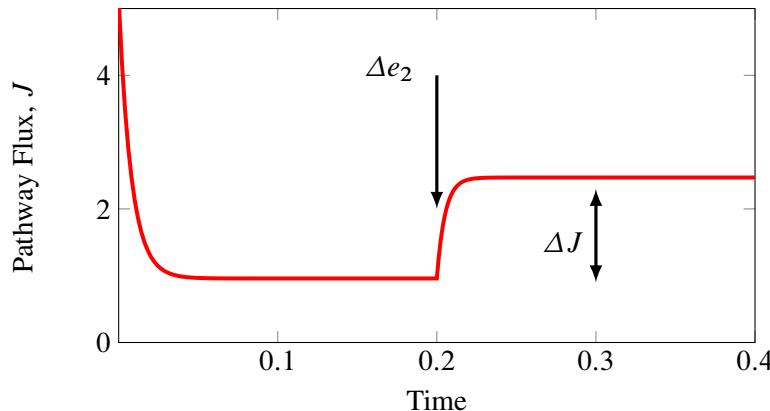
To make matters simpler, assume that the concentration of right-hand boundary pool,  $X_5$  is set to zero. In order to have a net flux through the pathway, the external metabolite,  $X_o$ , must have a positive value, perhaps 1 mM. This is the situation at time zero. Let us allow the pathway to evolve in time. The first thing that happens is that the reaction catalyzed by the first enzyme begins to convert  $X_o$  into product  $S_1$ . Since we assume that  $X_o$  is fixed, the concentration of  $X_o$  is unaffected by this rate of consumption. However, the product  $S_1$  is a floating species and as time goes on, its concentration will rise. As the concentration of  $S_1$  increases, two things will happen. First, the second enzyme will begin to convert  $S_1$  into  $S_2$ , and secondly  $S_1$  will begin to inhibit its own production rate by the first enzyme on account of product inhibition. The first reaction will therefore begin to rise at a **slower** rate.

Since the second enzyme is now generating  $S_2$ ,  $S_2$  starts to increase.  $S_2$  in turn will stimulate the third enzyme to begin making  $S_3$ , but it also begins to inhibit the second enzyme. And so on down the chain, all concentrations begin to rise and all enzymes show a positive rate. The concentrations of the floating species and the reaction rates cannot however go on rising forever. We have already seen that as the species concentrations rise, they begin to inhibit the enzymes that produce them. The net effect of these many interactions is that the concentrations slowly settle to a constant value such that the rates at which they are being made is exactly balanced by the rates at which they are being consumed. The rate of the first enzyme must balance the rate of the second enzyme, that is  $v_1 = v_2$ , but the second and third rates will also be in balance, so that  $v_2 = v_3$ . This must mean that the rate through the first enzyme must be the same as the rate through the third enzyme,  $v_1 = v_3$ . In fact all rates across each enzyme will equal each other, that is:

$$v_1 = v_2 = v_3 = v_4 = v_5$$

This state is the steady state, where the concentrations of all the metabolites settle to some value and no longer evolve in time, and the rate through each step is the **same**.

The fact that the rate across each enzyme is the same also means that there is a constant flow of material through the pathway, which we call the **flux**, symbolized by  $J$ . At steady state, there are no ‘slow’ rates or ‘fast’ rates, they are all the **same**.



**Figure 7.2** Effect of a perturbation in  $E_2$  at  $t = 0.2$  on the flux through the pathway, Figure 7.1. Note the initial transition to steady state between  $t = 0$  and  $t = 0.2$ . At  $t = 0.2$  a change  $\Delta e_2$  is made to  $E_2$  resulting in a change  $\Delta J$  in the steady state pathway flux. Parameters are given in the Tellurium Script: 8.1 at the end of the chapter.

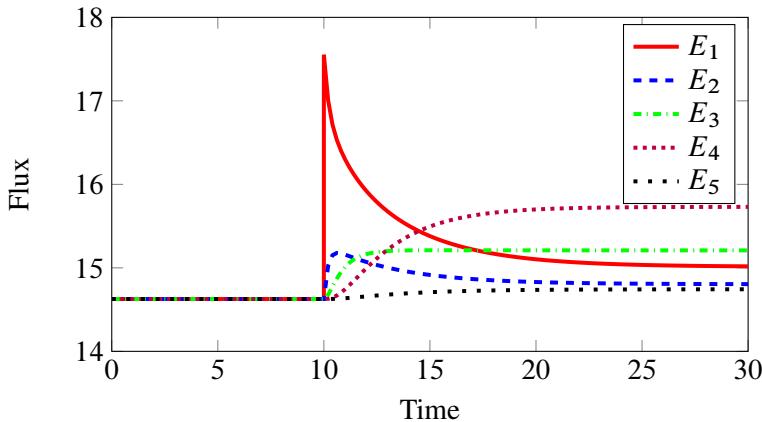
In a linear pathway at steady state, the rates of reaction are equal to each other and non-zero. At the same time all floating species are unchanging.

Given the pathway at steady state, we can consider some additional thought experiments such as the effect of perturbations on the steady state. Let us change the concentration of one of the enzymes and see what happens to the steady state concentrations and the flux,  $J$ . Let us double the concentration of enzyme,  $E_2$ , that catalyzes the second step. The immediate effect is to increase the rate,  $v_2$ , through the step. This in turn results in more  $S_2$  being produced and more  $S_1$  consumed,  $S_2$  will therefore rise and  $S_1$  will fall. The rise in  $S_2$  will cause the reaction rates through each step downstream to increase as  $S_3$  and  $S_4$  start to rise. Assuming that the first enzyme is product inhibited by  $S_1$ , the fall in  $S_1$  will cause a **rise** in the rate through  $v_1$ . The net effect of all these changes is that the net flux through the pathway will **increase**, all species concentration downstream of  $v_2$  will **increase** and  $S_1$  will **decrease**. Figure 7.2 illustrates a simulation that shows the change in flux through  $E_2$  as the pathway approaches steady state, followed by the effect of a perturbation in  $E_2$  at  $t = 0.2$ .

To see how effective the change in enzyme concentration is, we can take the ratio of the change in flux or species concentration to the change in enzyme:

$$\frac{\Delta J}{\Delta e_2}, \quad \frac{\Delta s_1}{\Delta e_2}, \dots \frac{\Delta s_4}{\Delta e_2}$$

where  $\Delta$  means ‘a change in’. However, because enzyme kinetic rate laws are usually nonlinear, the degree of influence we measure will depend on the size of the  $\Delta E$ . Therefore instead of making large changes to the enzyme concentration, we should make small changes,



**Figure 7.3** Effect of perturbing each enzyme by 20% in the linear pathway, Figure 7.1. Note that each enzyme affects the system differently both in the transient response and in the final steady state response. Parameters given in the Tellurium script: 7.3.

for example:

$$\frac{\delta J}{\delta e_2}, \quad \frac{\delta s_1}{\delta e_2}, \dots \frac{\delta s_4}{\delta e_2}$$

where  $\delta$  means ‘a small change’. We can be more precise mathematically if we make the changes infinitesimally small. Our measurement of influence then becomes:

$$\frac{dJ}{de_2}, \quad \frac{ds_1}{de_2}, \dots \frac{ds_4}{de_2}$$

Finally, if we want to make the measurement useful to experimentalists, we can remove the units by scaling the derivatives, such that:

$$\frac{dJ}{de_2} \frac{e_2}{J}, \quad \frac{ds_1}{de_2} \frac{e_2}{s_1}, \dots \frac{ds_4}{de_2} \frac{e_2}{s_4}$$

Obviously in an experiment we cannot make infinitesimal changes, but we can make changes sufficiently small (but still measurable) that we can approximate the derivatives. The scaled derivatives are called **control coefficients**, and we will define both the flux and concentration control coefficients as follows:

**Definition of the Flux Control Coefficient:**

$$C_{e_i}^J = \frac{dJ}{de_i} \frac{e_i}{J} = \frac{d \ln J}{d \ln e_i} \approx \frac{J \%}{e_i \%} \quad (7.1)$$

**Definition of the Concentration Control Coefficient:**

$$C_{e_i}^{s_j} = \frac{ds_j}{de_i} \frac{e_i}{s_j} = \frac{d \ln s_j}{d \ln e_i} \approx \frac{s_j \%}{e_i \%} \quad (7.2)$$

**Example 7.1**

A given enzyme catalyzed reaction in a metabolic pathway has a flux control coefficient equal to 0.2:

$$C_e^J = 0.2$$

What does this mean?

A flux control coefficient of 0.2 means that increasing the enzyme activity of the step by 1% will increase the steady state flux through the pathway by 0.2%.

In expression (7.1),  $J$  is the flux through the pathway and  $e_i$  the enzyme concentration of the  $i^{th}$  step. Operationally, an individual  $C_{e_i}$  is measured by making a small change in  $E_i$ , waiting for the system to reach a new steady state, and then taking the ratio of the change. Before moving on to another step, the level of  $E_i$  must be restored back to its original value. From a practical standpoint we see that the control coefficients can also be approximated by the ratio of **percentage changes** which is a useful interpretation for measurement purposes. The other point to note is that like elasticities, we can express the control coefficients in log form (See section 6.2).

The flux control coefficient measures the fractional change in flux brought about by a given fractional change in enzyme concentration. The concentration control coefficients measure the fractional change in species concentration given a fractional change in enzyme concentration. Control coefficients are useful because they tell us how much influence each enzyme or protein has in a biochemical reaction network.

It is important to note however, that knowing the values of the control coefficients does not tell us why certain enzymes or proteins have more influence than others. To answer the ‘why’ question we must consider the theorems associated with how control is distributed, and the relationship of the control coefficients to the elasticities of the network.

### 7.3 Distribution of Control

Flux control coefficients are a useful measure to judge the degree to which a particular step influences the steady state flux. One key question is how the influences are distributed across a pathway.

Consider the simple two step pathway:

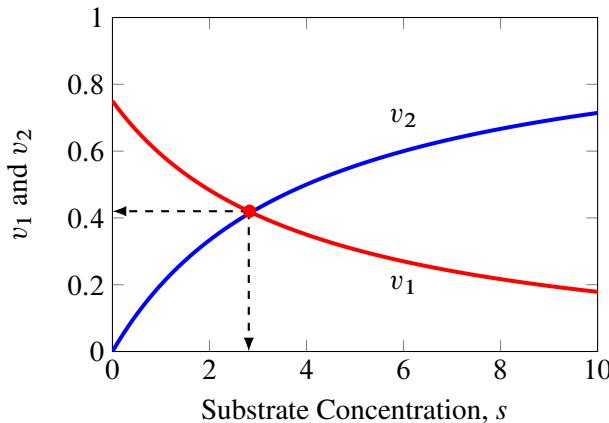


There is a simple graphical technique we can use to study how the enzyme concentrations,  $e_1$  and  $e_2$ , control the steady state concentration,  $s$ , and the steady state flux,  $J$ . In this system, the steady state flux,  $J$ , will be numerically equal to the reaction rates  $v_1$  and  $v_2$ :

$$J = v_1 = v_2$$

It is important to recall that for an isolated enzyme where all reactants, products and effectors are held constant, the reaction rate  $v$  is proportional to the concentration of enzyme,  $E$ , i.e  $v \propto e$ .

Let us plot both reaction rates,  $v_1$  and  $v_2$ , against the substrate concentration,  $s$ , Figure 7.4.



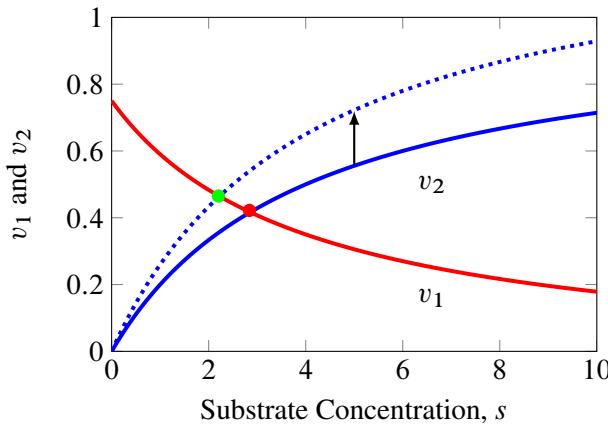
**Figure 7.4** Plot of  $v_1$  and  $v_2$  versus the concentration of  $S$  for a simple two step pathway. The intersection of the two curves marks the point when  $v_1 = v_2$ , that is steady state. A perpendicular dropped from this point gives the steady state concentration of  $S$ .

Note the response of  $v_1$  to changes in  $s$ .  $v_1$  falls as  $s$  increases due to product inhibition by  $S$ . The intersection point of the two curves marks the point when  $v_1 = v_2$ , that is, the steady state. A line dropped perpendicular from the intersection point marks the steady state concentration of  $S$ .

Let us increase the concentration of  $E_2$  by 30% by adding more enzyme (Figure 7.5). Because the reaction rate is proportional to  $E_2$ , the curve is scaled upwards although its general shape stays the same. Note how the intersection point moves to the left, indicating that the steady state concentration of  $S$  **decreases** relative to the reference state. This is understandable because with a higher  $v_2$ , more  $S$  is consumed, therefore  $S$  decreases.

In the next experiment, restore  $E_2$  back to its original level and instead increase the amount of  $E_1$  by 30% (Figure 7.6). Again, changing  $E_1$  scales the  $v_1$  curve but because of the negative curvature, the  $v_1$  curve shifts right. This moves the intersection point to the right, indicating that the steady state concentration of  $S$  **increases** relative to the reference state.

Let us now change the activity of **both**  $E_1$  and  $E_2$  by 30% (Figure 7.7). Note that the curves for  $v_1$  and  $v_2$  are both scaled upwards, moving the intersection point vertically upwards and



**Figure 7.5**  $v_2$  has been increased by 30% (dotted line) by increasing the enzyme concentration on  $v_2$ . This results in a displacement of the steady state curve to the left, leading to a decrease in the steady state concentration of  $S$ .

therefore **does not** change the steady state concentration of  $S$ . This happens because both curves move vertically by the same fraction so that the intersection point can only move vertically.

This experiment highlights an important result, when all enzyme concentrations are increased by the same fraction, the flux increases by that same fraction but the species or metabolite levels remain **unchanged**. We can summarize this with the following statement:

If all  $E_i$  are increased by a factor  $\alpha$ , then the steady state change in  $J$  and  $S_i$  is:

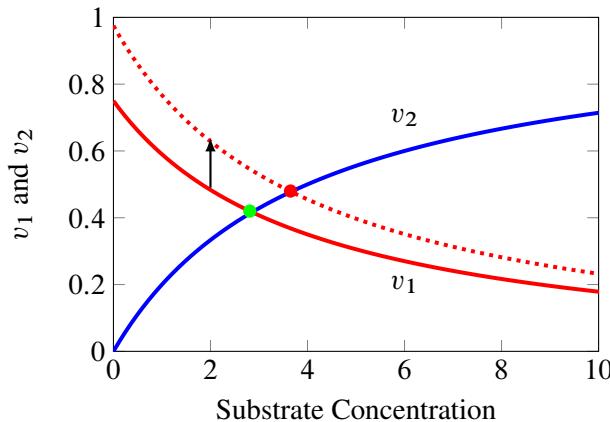
Given that all enzymes are increased by  $\alpha e_i$  then

$$\delta J = \alpha J \quad \text{and} \quad \delta S_j = 0$$

This is such an important result that it will be repeated again:

Increasing the activities of both enzymes by the **same fraction** will increase the flux through the pathway by the **same fraction** but will **not change** the concentration of the pathway species,  $S$ .

Given a pathway with  $n$  steps where every enzyme concentration is raised by the same factor,  $\alpha$ , the species concentrations will remain unchanged but the flux increases by a factor  $\alpha$ . This observation is true no matter how complex the pathway topology and doesn't just apply to linear chains of reaction steps.

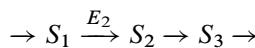


**Figure 7.6**  $v_1$  has been increased by 30% (dotted line) by increasing the enzyme concentration on  $v_1$ . This results in a displacement of the steady state curve to the right, leading to an increase in the steady state concentration of  $S$ .

A way to understand why  $J$  increases by  $\alpha J$  is as follows. Since  $\delta s = 0$ , the only change that could possibly effect the flux is the change in enzyme concentration, since the enzyme concentration has increased by a given proportion ( $\alpha$ ), the flux must also have increased by the same proportion since the rate is proportional to the enzyme concentration (i.e.  $v_i \propto e_i$ ), hence  $J \rightarrow \alpha J$ .

### Example 7.2

In the following pathway, an increase in  $E_2$  by 20% results in a 5% increase in the steady state flux. Estimate the value of  $C_{e_2}^J$ .



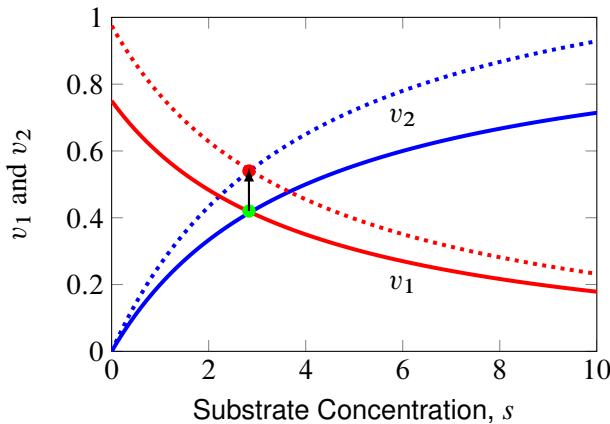
$C_{e_2}^J$  is the ratio of the fractional change in flux divided by the fraction change in the enzyme concentration. Therefore, an estimate for the control coefficient is given by:

$$C_{e_2}^J = \frac{0.05}{0.2} = 0.25$$

## 7.4 Predicting Flux and Concentration Changes

We can also rearrange equations (7.1), and (7.2) into the following form:

$$\frac{dJ}{J} = C_{e_i}^J \frac{de_i}{e_i}$$



**Figure 7.7** In this experiment, both  $E_1$  and  $E_2$  are increased by 30 % (dotted lines). Because both rates are increased by the same amount, the rate of change of  $S$  does not change. This means that there is no resulting change to the steady state concentration of  $S$ . The net flux through the pathway has however increased by 30 %.

$$\frac{ds_j}{s_j} = C_{e_i}^{s_j} \frac{de_i}{e_i}$$

These simple relations allow us to compute the change in flux or concentration given a change in enzyme concentration. These relations only hold true if the changes in enzyme concentration are infinitesimal. For practical purposes the relationships will approximately hold provided the changes in  $E_i$  are small. Of more interest is if we make changes to multiple enzymatic steps, the overall change will be the sum of the individual changes. The technical reason for this is that small changes in  $E_i$  mean that the response is linear so that multiple responses can be summed to obtain the total response. In general, if we make changes to  $n$  reaction steps, then the overall change in flux and species concentrations is given by:

$$\frac{dJ}{J} = \sum_{i=1}^n C_{e_i}^J \frac{de_i}{e_i} \quad (7.3)$$

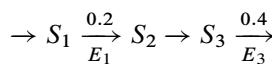
$$\frac{ds}{s} = \sum_{i=1}^n C_{e_i}^s \frac{de_i}{e_i} \quad (7.4)$$

See also Box 7.1.

#### Example 7.3

In the following pathway, the numbers refer to the flux control coefficients for the respective reaction

step:



What is the percentage change in flux if we increase  $E_1$  by 10% and  $E_3$  by 20%?

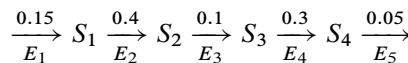
To calculate this we use equation (7.3):

$$\frac{\delta J}{J} = 0.1 \times 0.2 + 0.2 \times 0.4 = 0.1 \text{ or } 10\%$$


---

#### **Example 7.4**

In the following pathway, the numbers refer to the flux control coefficients for the respective reaction steps:



Given no other information, if you could increase enzyme concentrations by 20%, which two steps would you engineer to increase the flux the most?

Engineering the second and fourth enzymes will have the most effect on the steady state flux since they have the highest flux control coefficients. If we increased the second and fourth step by 20% the percentage change in flux will be:

$$\frac{\delta J}{J} = 0.2 \times 0.4 + 0.2 \times 0.3 = 0.14 \text{ or } 14\%$$


---

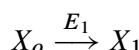
## 7.5 Summation Theorems

---

In this section we will introduce the concept of an **operational proof**. These proofs rely on carrying out thought experiments on a system and then casting the experiments in algebraic form from which new results (or theorems) can be derived. Although perhaps not as rigorous as a purely algebraic approach, operational proofs offer insight into the underlying biology and dynamics of the system, and are therefore very useful exercises in their own right.

### One step pathway

Although perhaps somewhat trivial, let us first consider a one step pathway:



The pathway is a single reaction catalyzed by a single enzyme,  $E_1$ . There are no floating metabolites, just two fixed external pools,  $X_o$  and  $X_1$ . There is a flux from  $X_o$  to  $X_1$ , denoted  $J$ , and is equal to the rate through the enzyme,  $E_1$ . If we make a change to the

**Box 7.1 Total Derivative. Proof of (7.3) and (7.4).**

If:

$$J = J(e_1, e_2, \dots)$$

then the total derivative is given by:

$$dJ = \frac{\partial J}{\partial e_1} de_1 + \frac{\partial J}{\partial e_2} de_2 + \dots$$

Dividing both sides by  $J$ , and for each term on the right, multiply top and bottom by the appropriate  $e_i$ , yields:

$$\frac{dJ}{J} = \frac{\partial J}{\partial e_1} \frac{e_1}{J} \frac{de_1}{e_1} + \frac{\partial J}{\partial e_2} \frac{e_2}{J} \frac{de_2}{e_2} + \dots$$

$$\frac{dJ}{J} = C_{e_1}^J \frac{de_1}{e_1} + C_{e_2}^J \frac{de_2}{e_2} + \dots = \sum_{i=1}^n C_{e_i}^J \frac{e_i}{J}$$

The same applies to  $s = s(e_1, e_2, \dots)$ .

concentration of  $E_1$ , there will be a change in flux, given by  $\delta J$ , which is equal to the change in rate,  $\delta v$ .

In order to work out the change in flux (although trivial in this case) we will use a standard procedure that will be employed in subsequent examples. This procedure involves the use of equation (6.15) given towards the end of the last chapter. This equation is repeated here:

$$\frac{dv}{v} = \sum_{j=1}^m \varepsilon_{s_j}^v \frac{ds_j}{s_j} + \varepsilon_e^v \frac{de}{e} \quad (7.5)$$

The equation enables us to compute the total change in rate through a step given changes to all modifiers that might affect the rate, including the enzyme catalyzing the reaction. In the one step pathway, the only thing that is changing is the concentration of  $E_1$  which means there are no  $\delta s_j$  terms to consider. Inserting the change in  $E_1$  into the above equation and dropping the  $\delta s_j$  terms, gives us the following relation:

$$\frac{dv}{v} = \varepsilon_{e_1}^v \frac{de_1}{e_1} \quad (7.6)$$

The elasticity,  $\varepsilon_{e_1}^v$ , equals one. This can be shown as follows. Assume that the reaction rate can be described using a standard Michaelis-Menten type rate law:

$$v = \frac{V_m x_o}{K_m + x_o}$$

where  $V_m$  is the maximal velocity, and  $K_m$  the concentration of substrate,  $X_o$ , that yields half the maximal rate. The maximal velocity can be further described in terms of the catalytic rate constant,  $k_{cat}$ , and the concentration of enzyme,  $e_1$ . We can therefore rewrite the rate law as:

$$v = \frac{e_1 k_{cat} x_o}{K_m + x_o}$$

To compute the enzyme elasticity,  $\varepsilon_{e_1}^v$ , the rate law can be differentiated with respect to  $e_1$  and scaled:

$$\varepsilon_{e_1}^v = \frac{k_{cat} x_o}{(K_m + x_o)} \frac{(K_m + x_o)}{e_1 k_{cat} x_o} e_1 = 1 \quad (7.7)$$

This yields an elasticity of one. Another way to look at this is to realize that  $v \propto e_1$  which is a first-order reaction response, and as we saw in the last chapter, first-order responses yield elasticities of one. Setting the enzyme elasticity to one allows us to simplify equation (7.6) to:

$$\frac{dv}{v} = \frac{de_1}{e_1}$$

The change in rate on the left-hand side takes into account all changes that have occurred. The left-hand side must therefore equal the change in the pathway's flux, and so:

$$\frac{dJ}{J} = \frac{de_1}{e_1}$$

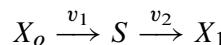
Dividing both sides by  $de_1/e_1$  yields the flux control coefficient:

$$C_{e_1}^J = 1$$

We can conclude that the enzyme of a single step pathway has complete proportional control over the flux. The total amount of control is equal to one.

## Two step pathway

Consider next a two step pathway:



where  $X_o$  and  $X_1$  are fixed. Let the pathway be at steady state and imagine increasing the concentration of enzyme,  $E_1$ , catalyzing the first step by an amount,  $\delta e_1$ . The effect of this is to increase the steady state levels of  $S$  and flux,  $J$ . Let us now increase the level of  $E_2$  by  $\delta e_2$  such that the change in  $S$  is restored to its original value it had in steady state.

The net effect of these two changes is that by definition,  $\delta s = 0$ .

There are two ways to look at this thought experiment, from the perspective of the **system** and from the perspective of **local changes**. For the system we can compute the overall change in flux or species concentration by adding the two control coefficient terms using equation (7.3), thus:

$$\begin{aligned}\frac{\delta J}{J} &= C_{e_1}^J \frac{\delta e_1}{e_1} + C_{e_2}^J \frac{\delta e_2}{e_2} \\ \frac{\delta s}{s} &= C_{e_1}^s \frac{\delta e_1}{e_1} + C_{e_2}^s \frac{\delta e_2}{e_2} = 0\end{aligned}\tag{7.8}$$

We can also look at what is happening locally at every reaction step using equation (7.5) for which there will be two: one for  $v_1$ , and another for  $v_2$ . Since the thought experiment guarantees that  $\delta s = 0$ , the local equations are quite simple:

$$\begin{aligned}\frac{\delta e_1}{e_1} &= \frac{\delta v_1}{v_1} \\ \frac{\delta e_2}{e_2} &= \frac{\delta v_2}{v_2}\end{aligned}$$

Because the pathway is linear, at steady state,  $v_1 = v_2 = J$ . We can substitute these expressions into (7.8) and rewrite the system equations as:

$$\frac{\delta J}{J} = C_{e_1}^J \frac{\delta v_1}{v_1} + C_{e_2}^J \frac{\delta v_2}{v_2}$$

$$\frac{\delta s}{s} = C_{e_1}^s \frac{\delta v_1}{v_1} + C_{e_2}^s \frac{\delta v_2}{v_2} = 0$$

Since  $\delta J/J = \delta v_1/v_1 = \delta v_2/v_2$ , we can rewrite the above equations as:

$$\alpha = C_{e_1}^J \alpha + C_{e_2}^J \alpha = \alpha(C_{e_1}^J + C_{e_2}^J)$$

$$0 = C_{e_1}^s \alpha + C_{e_2}^s \alpha = \alpha(C_{e_1}^s + C_{e_2}^s)$$

where  $\alpha = \delta J/J$ . We then conclude through cancelation of  $\alpha$  that:

$$\begin{aligned}1 &= C_{e_1}^J + C_{e_2}^J \\ 0 &= C_{e_1}^s + C_{e_2}^s\end{aligned}$$

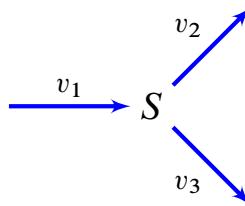
It can be shown that these relations can be extended to longer pathways so that for a linear pathway of  $n$  enzyme catalyzed reactions, the summations will include all  $n$  steps, that is:

$$\begin{aligned}\sum_{i=1}^n C_{e_i}^J &= 1 \\ \sum_{i=1}^n C_{e_i}^s &= 0\end{aligned}$$

Without proof (see later), the summation theorems apply to pathways of any shape or size with any number of regulatory loops. To help justify this sweeping statement, let us consider one more example, a simple branched pathway.

### Simple branched pathway

Consider the branched pathway shown in Figure 7.8.



**Figure 7.8** Simple Branched Pathway.

At steady state the following statement is true:

$$v_1 = v_2 + v_3$$

As before, let us make a positive perturbation in the concentration of  $E_1$ , that is  $\delta e_1$  to the reaction step  $v_1$ . This will cause the steady state level of  $S$  and all reactions rates downstream to increase. We now make changes to the concentrations of  $E_2$  and  $E_3$  such that  $\delta s = 0$ . This can be done by decreasing the levels of  $E_2$  and  $E_3$  until the steady state concentration of  $S$  is restored back to where it was before the initial perturbation in  $E_1$ . We now ask what are the magnitudes of the perturbations in  $e_1$ ,  $e_2$ , and  $e_3$ , needed to ensure that  $\delta s = 0$ ?

The local equations for each reaction step can be written as follows, note there are no  $\delta s$  terms because  $\delta s = 0$  (again assuming that  $v_i \propto e_i$ ):

$$\frac{\delta e_1}{e_1} = \frac{\delta v_1}{v_1}, \quad \frac{\delta e_2}{e_2} = \frac{\delta v_2}{v_2}, \quad \frac{\delta e_3}{e_3} = \frac{\delta v_3}{v_3}$$

It is also true that to satisfy the steady state condition, the sum of the changes in  $\delta v_2$  and  $\delta v_3$  must equal the change  $\delta v_1$ , that is:

$$\delta v_1 = \delta v_2 + \delta v_3$$

We divide both sides by  $v_1$  and adjusting the denominators of  $v_2$  and  $v_3$  to obtain:

$$\frac{\delta v_1}{v_1} = \frac{\delta v_2}{v_2} \frac{v_2}{v_1} + \frac{\delta v_3}{v_3} \frac{v_3}{v_1}$$

To constrain the possible solutions, we impose the condition  $\delta v_1/v_1 = \delta v_2/v_2$ , allowing us to rewrite the above equation as:

$$\frac{\delta v_1}{v_1} - \frac{\delta v_1}{v_1} \frac{v_2}{v_1} = \frac{\delta v_3}{v_3} \frac{v_3}{v_1}$$

That is:

$$\frac{\delta v_1}{v_1} \left(1 - \frac{v_2}{v_1}\right) = \frac{\delta v_3}{v_3} \frac{v_3}{v_1}$$

where  $v_2/v_1$  is the fraction of flux going down the upper branch, which we will call,  $\alpha$ , and  $v_3/v_1$  the fraction of flux going down the lower arm, that is  $1 - \alpha$ . Therefore:

$$\frac{\delta v_1}{v_1} (1 - \alpha) = \frac{\delta v_3}{v_3} (1 - \alpha)$$

$$\text{and so } \frac{\delta v_1}{v_1} = \frac{\delta v_3}{v_3}$$

In other words  $\delta v_1/v_1 = \delta v_2/v_2 = \delta v_3/v_3$ . We now write out the system equations for the flux through the first step,  $J$ :

$$\begin{aligned} \frac{\delta J}{J} &= C_{e_1}^J \frac{\delta v_1}{v_1} + C_{e_2}^J \frac{\delta v_2}{v_2} + C_{e_3}^J \frac{\delta v_3}{v_3} \\ \frac{\delta s}{s} &= C_{e_1}^s \frac{\delta v_1}{v_1} + C_{e_2}^s \frac{\delta v_2}{v_2} + C_{e_3}^s \frac{\delta v_3}{v_3} \end{aligned}$$

Using the previous result that  $\delta v_1/v_1 = \delta v_2/v_2 = \delta v_3/v_3$  we can write:

$$\begin{aligned} \frac{\delta J}{J} &= C_{e_1}^J \frac{\delta v_1}{v_1} + C_{e_2}^J \frac{\delta v_1}{v_1} + C_{e_3}^J \frac{\delta v_1}{v_1} \\ \frac{\delta s}{s} &= C_{e_1}^s \frac{\delta v_1}{v_1} + C_{e_2}^s \frac{\delta v_1}{v_1} + C_{e_3}^s \frac{\delta v_1}{v_1} \end{aligned}$$

Note that the flux change  $\delta J$  through  $v_1$  is non-zero while  $\delta s = 0$ :

$$\begin{aligned} \frac{\delta J}{J} &= \frac{\delta v_1}{v_1} (C_{e_1}^J + C_{e_2}^J + C_{e_3}^J) \\ \frac{\delta s}{s} &= \frac{\delta v_1}{v_1} (C_{e_1}^s + C_{e_2}^s + C_{e_3}^s) = 0 \end{aligned}$$

therefore:

$$\begin{aligned} 1 &= C_{e_1}^J + C_{e_2}^J + C_{e_3}^J \\ 0 &= C_{e_1}^s + C_{e_2}^s + C_{e_3}^s \end{aligned}$$

Once again the flux control coefficients sum to one, and the concentration control coefficients sum to zero. Such operational proofs can be extended to other pathway configurations and a more formal approach using a combination of implicit differentiation together with the application of linear algebra techniques will show that the summation theorems apply to networks of arbitrary complexity. Given a pathway of arbitrary complexity, the following relationships (7.9) are true:

**Summation Theorems:**

$$\sum_{i=1}^n C_{e_i}^J = 1 \quad (7.9)$$

$$\sum_{i=1}^n C_{e_i}^{sj} = 0$$

where  $n$  equals the number of reaction steps in the pathway.

The one caveat is that if the control coefficients are expressed in terms of changes in enzyme concentrations, then there is the implicit assumption that  $v_i \propto E_i$  and changes in a particular  $E_i$  has no effect on other enzyme concentrations. Later on we will see that even these assumptions can be relaxed by using an alternative definition for the control coefficient.

### Interpreting the summation theorems

How can we interpret the summation theorems described in the last section? Consider is the flux summation theorem, repeated below:

$$\sum_{i=1}^n C_{e_i}^J = 1$$

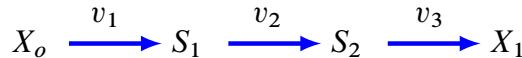
What biological insight can be obtained from this result? The flux summation theorem implies that if any one of the enzyme steps has a high control coefficient, then the remaining steps must have small control coefficients. This assumes that there are no negative control coefficients (a situation that is not guaranteed for branched pathways). For a linear chain of  $n$  steps, the average control coefficient will be  $1/n$ . For a pathway with many steps, either the control coefficient for each is small, or one or two steps have significant control and the remaining steps very little. It has been suggested that this property is the molecular basis for the existence of metabolic dominant and recessive genes. Heterozygotes which carry one normal copy and one mutant copy of a gene are found on average to have only 50 % of the enzyme activity of the homozygous individual. Yet heterozygous individuals are generally ‘normal’ in the sense that they appear indistinguishable from the homozygote. In such situations, the mutant gene is termed ‘recessive’ and the normal gene dominant. Such behavior can be explained if we assume that many of the control coefficients in a pathway are small, so that even a 50 % reduction in activity has little or no effect.

Another important aspect of the summation theorem is that the flux summation implies a ‘total’ or unit amount of control available in a pathway, the total being one. The total control is distributed amongst the different enzymes according to each enzyme’s control coefficient. Since there is a total amount of control to distribute in a pathway, an effect which causes the control coefficient of one enzyme to change will mean that the control coefficients of other steps **must change** to compensate and maintain the total control at unity. This implies that

the control distribution in a pathway is a dynamic process, changing as the conditions of the pathway change.

### Example

Consider the three step pathway shown in Figure 7.9. A simple mathematical model for this pathway is given in the Tellurium script listing 7.1.



**Figure 7.9** Three step linear pathway.

$$\begin{aligned} C_{Vm_1}^J &= 0.3677 \\ C_{Vm_2}^J &= 0.1349 \\ C_{Vm_2}^J &= 0.4989 \end{aligned}$$

**Table 7.1** Flux Control Coefficients.

```
import tellurium as te

r = te.loada '''
# Assume the kcat values equal 1
J1: $Xo -> S1; e1/Km1*(Xo-S1/Keq1)/(1 + Xo/Km1 + S1/Km2);
J2: S1 -> S2; e1/Km2*(S1-S2/Keq2)/(1 + S1/Km3 + S2/Km4);
J3: S2 -> $X1; e1*S2/(Km5 + S2);

Xo = 2; X1 = 0;
Keq1 = 1.2; Keq2 = 2.5;
e1 = 3.4; e2 = 8.2; e3 = 2.3;
Km1 = 0.6; Km2 = 0.78;
Km3 = 0.9; Km4 = 1.2;
Km5 = 0.5;
S1 = 0; S2 = 0;
''')

r.steadyState()
print " C1 = ", r.getCC ("J1", "e1"),
      " C2 = ", r.getCC ("J1", "e2"),
      " C3 = ", r.getCC ("J1", "e3")
```

**Listing 7.1** Three step pathway model.

The command `getCC()` is used to compute the control coefficient. The first argument is the variable we wish to observe, in this case the flux through step one. The second argument is the parameter we which to perturb, in this case one of the enzymes. Table 7.1 shows the values for the flux control coefficients computed from this model. Flux control is distributed across all three reaction steps. Almost 50% of control is located on the last step. This shows that in a linear pathway the committed step (i.e. the first step) is not necessarily the step with the most control. By varying the values of the various kinetic parameters, it is possible to obtain almost any pattern of control.

### What determines the value of a flux control coefficient?

Many of the subsequent chapters will focus on the question of what determines the value of a particular flux or concentration control coefficient. In this section a brief description will be given.

Consider the three step pathway shown in Figure 7.9. Assume the pathway is at steady state. What happens to the steady state flux when the concentration of the enzyme that catalyzes the second step is increased? A change,  $\delta e_2$ , is made resulting in new steady state concentrations to  $s_1$  and  $s_2$ , as well as a new steady state flux,  $J$ .

The change in reaction rate,  $v_2$ , can be described using the local equation introduced in equation (6.15) and repeated here using the  $\delta$  notation:

$$\frac{\delta v}{v} = \sum_{j=1}^m \varepsilon_{s_j}^v \frac{\delta s_j}{s_j} + \varepsilon_e^v \frac{\delta e}{e}$$

The local equation for  $v_2$  (7.10) will have three terms, one involving  $e_2$ , and two other terms involving  $s_1$  and  $s_2$ :

$$\frac{\delta v_2}{v_2} = \varepsilon_{s_1}^{v_2} \frac{\delta s_1}{s_1} + \varepsilon_{s_2}^{v_2} \frac{\delta s_2}{s_2} + \varepsilon_{e_2}^{v_2} \frac{\delta e_2}{e_2} \quad (7.10)$$

Collectively the three terms determine how the rate through  $v_2$  changes. The enzyme,  $e_2$ , contributes a positive change in rate given by  $\varepsilon_{e_2}^{v_2} \delta e_2 / e_2$ . Since the enzyme elasticity  $\varepsilon_{e_2}^{v_2}$  is one, the change in  $v_2$  is in proportion to the change in  $e_2$ . For the sake of argument assume that  $e_2$  is increased by 2% and the resulting net change in the final steady state flux increases by 1%. The above equation can be written as:

$$0.01 = \varepsilon_{s_1}^{v_2} \frac{\delta s_1}{s_1} + \varepsilon_{s_2}^{v_2} \frac{\delta s_2}{s_2} + 0.02 \quad (7.11)$$

It should be clear from the equation that the terms that include  $s_1$  and  $s_2$  must make a negative contribution in order to reduce the contribution from  $e_2$  so that the overall sum equals 0.01.

As a result of the increase in  $v_2$ ,  $s_1$  will be consumed faster and  $s_2$  made faster.  $s_1$  will therefore **decrease** by an amount,  $\delta s_1$ , and  $s_2$  will **increase** by an amount,  $\delta s_2$ . The drop in  $s_1$  will result in a reduction in the reaction rate, and the increase in  $s_2$  will likewise reduce the rate via product inhibition. Both changes **oppose** the increase in rate caused by  $e_2$ .

The degree to which changes in  $s_1$  and  $s_2$  change the rate will be determined by the two elasticities,  $\varepsilon_{s_1}^{v_2}$  and  $\varepsilon_{s_2}^{v_2}$ . If both elasticities are large, then the reduction in rate caused by the change in  $s_1$  and  $s_2$  will be equally great, meaning that any change brought about by changes in  $e_2$  will be effectively zeroed out. The net effect is that the flux control coefficient for this step will be small.

If both  $\varepsilon_{s_1}^{v_2}$  and  $\varepsilon_{s_2}^{v_2}$  are small then the reduction in rate caused by the change in  $s_1$  and  $s_2$  is likely also to be small. In this case most of the contribution is coming from  $e_2$  and therefore the flux control coefficient will be higher.

To summarize:

1. Control coefficients describe the influence a given reaction step has on a flux or species concentration at steady state.
2. Control coefficients are dimensionless quantities.
3. Flux control is shared throughout a pathway; that is, the degree to which flux is limited in a pathway is shared. It is unlikely that a single step has exclusive control over the flux.
4. If one step gains flux control, one or more other steps must lose control.
5. Control coefficients are system properties; they can only be computed or measured in an intact system. Inspection of a single enzymatic step will not reveal its degree of control (or influence) on the pathway.

## Rate-limiting steps

In much of the metabolic literature and many contemporary biochemistry textbooks, one will often find a brief discussion of an idea called the rate-limiting step. In text books we will find statements such as:

*'It is of course a truism to say that every metabolic pathway can and must have only one rate-limiting step'*, Denton and Pogson, 1976.

*'In order to exert control on the flux of metabolites through a metabolic pathway, it is necessary to regulate its rate-limiting step'* Voet and Voet Biochemistry (p522), 3rd edition, 2004.

*'The rate-limiting step occurs near the beginning of the pathway and is regulated by feedback inhibition...'* Wikipedia (April 16, 2018), [https://en.wikipedia.org/wiki/Metabolic\\_pathway](https://en.wikipedia.org/wiki/Metabolic_pathway)

The implication is that a given metabolic pathway has one and only one rate-limiting step,

often assigned to either the so-called committed step or regulated step. However this is inconsistent with the previously presented view that control, or rate-limitingness can be shared among all steps and not confined to just one. Other than a few books such as the 4th edition of Lehninger Biochemistry, most textbooks and unfortunately many research articles continue to cling to the idea of a single rate-limiting step. The 4th edition of Lehninger states more correctly:

*'Metabolic control analysis shows that control of the rate of metabolite flux through a pathway is distributed among several of the enzymes in that path'.*

The literature also uses an overabundance of similar terms such as rate-determining, pace-maker, choke point, bottleneck, master reaction or key enzyme. These terms are meant to convey a similar idea to the rate-limiting step. One of the earliest references to the concept of the rate-limiting step is a quote from Blackman [8]:

*'When a process is conditioned as to its rapidity by a number of separate factors, the rate of the process is limited by the pace of the slowest factors.'*

This sentence started a century long love affair with the idea of the rate-limiting step in biochemistry, a concept that has lasted to this very day. From the 1930s to the 1950s there were however a number of published papers which were highly critical of the concept, most notably Burton [18], Morales [87] and in particular Hearon [50]. Unfortunately much of this work did not find its way into the rapidly expanding fields of biochemistry and molecular biology after the second world war. Instead, the intuitive idea first pronounced by Blackman remains today as one of the cornerstones in understanding cellular regulation. The concept drives much of metabolic engineering and drug targeting of metabolism. What is most surprising however, is that a simple quantitative analysis shows that it cannot be true, and there is ample experimental evidence [51, 17] to support the alternative notion of shared control. The concept of the rate-limiting step is both inconsistent with logic and more importantly, experimental evidence.

The confusion over the existence of rate-limiting steps stems from a failure to realize that rates in cellular networks are governed by the law of mass-action. That is, if a concentration changes, then so does its rate of reaction. Some researchers try to draw analogies between cellular pathways and human experiences such as traffic congestion on roads or customer lines at shopping store checkouts. In each of these analogies, the rate of traffic and the rate of customer checkouts does not depend on how many cars are in the traffic line or how many customers are waiting in line. In a sense, these systems are not governed by 'mass-action' kinetics. In these analogies the use of the phrase rate-limiting step is reasonable. Traffic congestion and the customer line are rate-limiting because the only way to increase the flow is to either widen the road or increase the number of cash tills, i.e. there is a single factor that determines the rate of flow. In reaction networks the flow is governed by many factors including substrate/product/effecter concentrations as well as the capacity of the reaction ( $V_{max}$ ) itself. In biological pathways, rate-limiting steps are therefore the exception rather than the rule. It is highly unlikely for a single reaction to be fully rate limiting because it can be influenced by many factors. Many hundreds of measurements of control coefficients

have been born out of this prediction.

Most biochemistry and molecular biology literature interpret the rate-limiting step to be the single step in a pathway which limits the flux. In terms of our control coefficients we can interpret the rate-limiting step as the step with a flux control coefficient of one. This means, by the summation theorem, that all other steps (at least in a linear chain) must have flux control coefficients of zero. However, when we consider branched and cyclic systems it is possible to have flux control coefficients much greater than one (other control coefficient must then be negative to satisfy the summation theorem). In these cases, what adjective should we use: hyper-rate-limiting steps, super-bottleneck, extreme-choke points? In the final analysis, it is better to assign a **value** to the rate-limitingness of a particular step in a pathway rather than designate a given reaction step as either rate-limiting or not.

## Further Reading

---

1. Rafael Moreno-Sanchez, Emma Saavedra, Sara Rodrguez-Enriquez, and Viridiana Olin-Sandoval (2008) Metabolic Control Analysis: A Tool for Designing Strategies to Manipulate Metabolic Pathways, Journal of Biomedicine and Biotechnology, Volume 2008, Article ID 597913, doi:10.1155/2008/597913

## Exercises

---

1. At steady state, all reaction rates are equal in a linear chain of reactions. Explain this statement.
2. The control coefficients are defined in terms of infinitesimal relative changes. An alternative would be to define them using large finite changes, that is  $\Delta J / \Delta E$  which could be more easily measured. What is the main disadvantage to defining control coefficients in terms of large finite changes?
3. List three properties of control coefficients.
4. In a given reaction step  $E_i$ , the enzyme concentration is increased by 15%. The steady state change in flux was found to be 5% and the change in a species,  $S_j$ , changed by -3%. Estimate the values for the flux control coefficient,  $C_{e_i}^J$  and the concentration control coefficient,  $C_{e_i}^{S_j}$ .
5. In the last question you were asked to find estimates for the control coefficients. Why were you asked to estimate the control coefficients and not their precise values?
6. A given reaction step has a flux control coefficient of 0.6. If the enzyme concentration is increased by 40%, what is the approximate change in the steady state flux?

7. Two reactions have flux control coefficients of 0.2 and 0.3, respectively. The concentration of the first enzyme is changed by 10%, and the second enzyme by 30%. What is the approximate change in the steady state flux if both changes are made?
8. What assumption(s) are made in the derivation of the summation theorems?
9. In a linear pathway, the concentration control coefficient for a give species,  $S$ , is found to be negative with respect to one enzyme but positive with respect to every other enzymatic step. Where is the species,  $S$ , located in the pathway. Explain your answer.
10. The last four steps in a five step pathway are found to have concentration control coefficients for a species,  $S$ , of -0.1, -0.2, -0.5 and -0.05. What is the concentration control coefficient with respect to the first step?
11. Locate five biochemistry and molecular cell biology textbooks and explain how the books describe regulation in pathways with respect to flux control. If they mention rate-limiting steps or rate-determining steps, describe how they justify these statements, if at all.
12. Why are rate-limiting steps unlikely to be found in natural pathways?

## 7.A Python/Tellurium Scripts

```
import tellurium as te
import numpy as np
r = te.loada ('''
J0: $Xo -> S1; (E1/0.3)*(0.5*Xo-100*S1)/(1+Xo+S1);
J1: S1 -> S2; E2*(3*S1-0.2*S2)/(1+S1+S2);
J2: S2 -> S3; E3*(500*S2-10*S3)/(1+S2+S3);
J3: S3 -> S4; E4*(200*S3-2*S4)/(1+S3+S4);
J4: S4 -> $X1; E5*(200*S4-2*X1)/(1+S4+X1);

Xo = 10; X1 = 0;
E1 = 3.4; E2 = 8.2;
E3 = 2.3; E4 = 1.8;
E5 = 4.5;
S1 = 0; S2 = 0; S3 = 0; S4 = 0;
''')

m1 = r.simulate (0, 0.2, 100, ['time', 'J0']);
r.E2 = r.E2*4;
m2 = r.simulate (0.2, 0.5, 200, ['time', 'J0']);
alldata = np.vstack ((m1, m2));
```

```
r.plot (alldata)
```

**Listing 7.2** Script for Figure 7.2.

```
import tellurium as te
import numpy as np
import pylab
r = te.loada (''''
J0: $Xo -> S1; E1*(10*Xo-2*S1)/(1+Xo+S1);
J1: S1 -> S2; E2*(10*S1-2*S2)/(1+S1+S2);
J2: S2 -> S3; E3*(10*S2-2*S3)/(1+S2+S3);
J3: S3 -> S4; E4*(10*S3-2*S4)/(1+S3+S4);
J4: S4 -> $X1; E5*(10*S4-2*X1)/(1+S4+X1);

Xo = 10; X1 = 0;
E1 = 3.4; E2 = 8.2;
E3 = 2.3; E4 = 1.8;
E5 = 4.5;
S1 = 8.359; S2 = 17.68; S3 = 6.938; S4 = 0.4816;
'''')

# ----- E1 -----
r.simulate (0, 80, 500);
m1 = r.simulate (0, 10, 100, ['time', 'J0']);
r.E1 = r.E1*1.2;
m2 = r.simulate (10, 30, 100, ['time', 'J0']);
alldata = np.vstack((m1, m2)); # Combine the two segments
pylab.plot (alldata[:,0], alldata[:,1], linewidth=2, label='E1')

# ----- E2 -----
r.resetAll()
r.simulate (0, 80, 500);
m1 = r.simulate (0, 10, 100, ['time', 'J0']);
r.E2 = r.E2*1.2;
m2 = r.simulate (10, 30, 100, ['time', 'J0']);
alldata = np.vstack((m1, m2)); # Combine the two segments
pylab.plot (alldata[:,0], alldata[:,1], linewidth=2, label='E2')

# ----- E3 -----
r.resetAll()
r.simulate (0, 80, 500);
m1 = r.simulate (0, 10, 100, ['time', 'J0']);
r.E3 = r.E3*1.2;
m2 = r.simulate (10, 30, 100, ['time', 'J0']);
alldata = np.vstack((m1, m2)); # Combine the two segments
pylab.plot (alldata[:,0], alldata[:,1], linewidth=2, label='E3')
```

```
# ----- E4 -----
r.resetAll()
r.simulate (0, 80, 500);
m1 = r.simulate (0, 10, 100, ['time', 'J0']);
r.E4 = r.E4 *1.2;
m2 = r.simulate (10, 30, 100, ['time', 'J0']);
alldata = np.vstack((m1, m2)); # Combine the two segments
pylab.plot (alldata[:,0], alldata[:,1], linewidth=2, label='E4')

# ----- E5 -----
r.resetAll()
r.simulate (0, 80, 500);
m1 = r.simulate (0, 10, 100, ['time', 'J0']);
r.E5 = r.E5 *1.2;
m2 = r.simulate (10, 30, 100, ['time', 'J0']);
alldata = np.vstack((m1, m2)); # Combine the two segments
pylab.plot (alldata[:,0], alldata[:,1], linewidth=2, label='E4')

pylab.legend(loc='upper left')
pylab.show()
```

**Listing 7.3** Script for Figure 7.3.



# 8

## *Linking the Parts to the Whole*

A deeper understanding of metabolism (or any other network) requires us to relate the parts, that is the enzymes, to the whole pathway. We need to understand how the properties of the parts contribute to the behavior we see in intact systems. To put it more grandly, we seek to understand phenotype from genotype.

So far two types of measures have been introduced: (i) the elasticities which describe how individual reactions respond to changes in their participating reactants and other modifiers, and (ii) control coefficients which describe how much influence individual reactions have on the response of whole pathways. In this chapter the aim is to bridge these two descriptions, and in order to do that, the relationship between the elasticities and control coefficients must be described.

### **8.1 Control Coefficients in Terms of Elasticities**

---

The relationship between the control coefficients and elasticities is best introduced by way of example. Consider the pathway:



Assume  $v_1$  is catalyzed by an enzyme  $E_1$ , and  $v_2$  by an enzyme  $E_2$ .  $X_o$  and  $X_1$  are fixed species. Assume the pathway is at steady state with a steady state flux of  $J$  and concentration  $s$ . An experiment is carried out where  $E_1$  is increased by an amount  $\delta e_1$ . This will result in a change to the steady state where the flux will change by an amount,  $\delta J$ , and the concentration of  $S$  changed by  $\delta s$ . The experiment can be described using two local (6.15) and two system equations (7.1) and (7.2). The local equations are concerned with changes at  $v_1$  and  $v_2$ ,

while the system equations describe how the change  $\delta e_1$  changes the steady state flux and concentration of  $S$ .

$$\left. \begin{array}{l} \frac{\delta v_1}{v_1} = \frac{\delta e_1}{e_1} + \varepsilon_1^1 \frac{\delta s}{s} \\ \frac{\delta v_2}{v_2} = \varepsilon_1^2 \frac{\delta s}{s} \end{array} \right\} \text{Local equations}$$

$$\left. \begin{array}{l} \frac{\delta s}{s} = C_{e_1}^s \frac{\delta e_1}{e_1} \\ \frac{\delta J}{J} = C_{e_1}^J \frac{\delta e_1}{e_1} \end{array} \right\} \text{System equations}$$

In this an subsequent chapter the notation for the elasticities will be simplified. In previous changes elasticities have been expressed using notation such as  $\varepsilon_{s_1}^{v_2}$ . From now on we will drop the  $v$  and  $s$  lettering in the sub and superscript and instead refer to the elasticity using the simpler notation  $\varepsilon_1^2$ . This will reduce the amount of clutter in the equations. For example,  $\varepsilon_3^5$  refers the elasticity of reaction  $v_5$  with respect to species  $s_3$ .

Because  $\delta v_1/v_1 = \delta v_2/v_2$  at steady state, the two local equations can be set equal to each other:

$$\frac{\delta e_1}{e_1} + \varepsilon_1^1 \frac{\delta s}{s} = \varepsilon_1^2 \frac{\delta s}{s}$$

Rearranging, and replacing  $\delta s/s$  with the system equation,  $C_{e_1}^s \delta e_1/e_1$  yields:

$$\frac{\delta e_1}{e_1} + \varepsilon_1^1 C_{e_1}^s \frac{\delta e_1}{e_1} = \varepsilon_1^2 C_{e_1}^s \frac{\delta e_1}{e_1}$$

Canceling and rearranging gives:

$$C_{e_1}^s = \frac{1}{\varepsilon_1^2 - \varepsilon_1^1} \quad (8.1)$$

The steady state change in  $\delta v_1/v_1$  is the same as the change in pathway flux,  $\delta J/J$ , therefore we can equate the first system equation to the first local equation:

$$C_{e_1}^J \frac{\delta e_1}{e_1} = \frac{\delta e_1}{e_1} + \varepsilon_1^1 \frac{\delta s}{s} \quad (8.2)$$

Dividing both sides of the equation by  $\delta e_1/e_1$  gives:

$$C_{e_1}^J = 1 + \varepsilon_1^1 \frac{\delta s/s}{\delta e_1/e_1}$$

Noting that in the limit  $(\delta s/s)/(\delta e_1/e_1)$  is equal to  $C_{e_1}^s$  yields:

$$C_{e_1}^J = 1 + \varepsilon_1^1 C_{e_1}^s$$

But from (8.1)  $C_{e_1}^s = 1/(\varepsilon_1^2 - \varepsilon_1^1)$ , so that:

$$C_{e_1}^J = 1 + \varepsilon_1^1 \frac{1}{\varepsilon_1^2 - \varepsilon_1^1}$$

Rearranging yields:

$$C_{e_1}^J = 1 + \frac{\varepsilon_1^1}{\varepsilon_1^2 - \varepsilon_1^1} = \frac{\varepsilon_1^2}{\varepsilon_1^2 - \varepsilon_1^1} \quad (8.3)$$

Given the flux and concentration summation theorems, it is easy to obtain the other two control coefficients by subtraction ( $C_{e_2}^J = 1 - C_{e_1}^J$  and  $C_{e_2}^s = -C_{e_1}^s$ ):

$$\begin{aligned} C_{e_2}^s &= -C_{e_1}^s = -\frac{1}{\varepsilon_1^2 - \varepsilon_1^1} \\ C_{e_2}^J &= 1 - C_{e_1}^J = -\frac{\varepsilon_1^1}{\varepsilon_1^2 - \varepsilon_1^1} \end{aligned} \quad (8.4)$$

The resulting equations describe how the control coefficients are determined by the elasticities. Using these equations it is possible to understand why particular steps are more or less rate limiting than others, and why some reactions may have more or less influence over metabolite levels. One interesting result is that the ratio of the flux control coefficients is related to the ratio of the flanking elasticities:

$$\frac{C_{e_1}^J}{C_{e_2}^J} = -\frac{\varepsilon_1^2}{\varepsilon_1^1}$$

Similarly, the ratio of the concentration control coefficients is minus one:

$$\frac{C_{e_1}^s}{C_{e_2}^s} = -1$$

These results are true for all pairs of adjacent steps in a pathway. These properties are described by a set of additional theorems that relate the control coefficients with the elasticities, and will be covered in the next section.

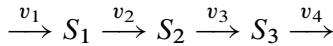
## 8.2 Connectivity Theorems

In this section an additional set of theorems are introduced that relate the control coefficients to the substrate, product and effector elasticities. The theorems are called the **connectivity theorems** and represent one of the most important results in metabolic control analysis.

In deriving the summation theorems in the previous chapter, certain operations were performed on the pathway such that the flux changed value but the concentrations of the species were unchanged, that is,  $\delta J/J \neq 0$  and  $\delta s/s = 0$ .

The constraints on the flux and concentration variables when deriving the summation theorems suggest a complementary set of constraints. That is, one or more operations can be performed on the enzymes such that the opposite is true,  $\delta J/J = 0$  and  $\delta s/s \neq 0$ . A set of operations which preserve the flux but change the species concentrations lead to the second set of theorems, called the **connectivity theorems**.

Consider the following pathway fragment:



Let us make a change to the rate through  $v_2$  by increasing the concentration of enzyme  $E_2$ . Assume  $E_2$  is increased by an amount,  $\delta e_2$ . This will result in a change to the steady state of the pathway. The concentration of  $S_2$ ,  $S_3$ , and the flux through the pathway will rise, and the concentration of  $S_1$  will decrease because it is upstream of the disturbance.

Let us now impose a second change to the pathway such that the flux is restored to what it was before the original change. Since the flux increased when  $E_2$  was changed, the flux can be decreased by decreasing one of the other enzyme levels. If the concentration of  $E_3$  is decreased, this will reduce the flux. Decreasing  $E_3$  will also cause the concentration of  $S_2$  to further increase. However,  $S_1$  and  $S_3$  will change in the opposite direction compared to when  $E_2$  was increased.

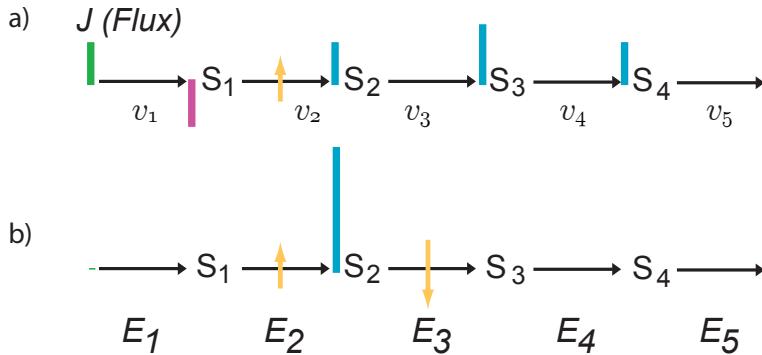
When  $E_3$  is sufficiently changed so that the flux is restored to its original value, the concentrations of  $S_1$  and  $S_3$  will also be restored to their original values. It is only  $S_2$  that will differ. This is true because the flux through  $v_1$  is now the same as it was originally (since we've restored the flux), and  $E_1$  has not been manipulated in anyway. This means that the concentration of  $S_1$  and all species upstream of  $S_1$  must be the same as they were before the modulations occurred. The same arguments apply to  $S_3$  and all species downstream of  $v_4$ .

To summarize:  $E_2$  has been **increased** by  $\delta e_2$ , resulting in a change,  $\delta J$ , to the flux. The concentration of  $E_3$  is **decreased** such that the flux is restored to its original value. In the process,  $S_2$  has changed by  $\delta s_2$  and neither  $S_1$  nor  $S_3$  have changed. In fact *no other* species in the entire system has changed other than  $S_2$ .

If a particular species is made and consumed by many steps, it would still be possible to perform the necessary manipulations on all the adjacent enzymes such that only the shared species changed in concentration and the flux was unaltered.

### Flux Connectivity Theorem

Using the thought experiment described in the last section, two sets of equations can be derived which apply simultaneously to the pathway, a local equation and a system equation. The system equation will describe the effect of the enzyme changes on the flux. Since the net change in flux is zero, and the fact that only  $E_2$  and  $E_3$  were changed, the system flux



**Figure 8.1** Connectivity Theorem: a) A change is made to  $E_2$  that catalyzes  $v_2$ . This causes changes to species concentrations upstream and downstream including a change in flux indicated by the vertical bars. Note that  $S_1$  decreases. b) Another change is made to  $E_3$  that catalyzes  $v_3$  to oppose the change in flux. This results in only three net changes, a change in  $E_2$ ,  $E_3$  and  $S_2$ ; no other changes occur.

can be written using the following system equation:

$$\frac{\delta J}{J} = 0 = C_{e_2}^J \frac{\delta e_2}{e_2} + C_{e_3}^J \frac{\delta e_3}{e_3} \quad (8.5)$$

To determine the local equations, we must focus on what is happening at the reaction steps  $v_2$  and  $v_3$ . As a result of making changes to  $E_2$  and  $E_3$ , the change in rate at  $v_2$  is given by:

$$0 = \frac{\delta v_2}{v_2} = \frac{\delta e_2}{e_2} + \varepsilon_2^2 \frac{\delta s_2}{s_2}$$

and at  $v_3$ :

$$0 = \frac{\delta v_3}{v_3} = \frac{\delta e_3}{e_3} + \varepsilon_2^3 \frac{\delta s_2}{s_2}$$

Note that  $\delta e_2/e_2$  will not necessarily equal  $\delta e_3/e_3$ . No other changes took place so these are the only local equations to consider. We can rearrange the local equations so that:

$$\frac{\delta e_2}{e_2} = -\varepsilon_2^2 \frac{\delta s_2}{s_2} \quad (8.6)$$

$$\frac{\delta e_3}{e_3} = -\varepsilon_2^3 \frac{\delta s_2}{s_2} \quad (8.7)$$

$\delta e_2/e_2$  and  $\delta e_3/e_3$  can be inserted from the local equations into the system equations (8.5) to obtain:

$$0 = \frac{\delta J}{J} = - \left( C_{e_2}^J \varepsilon_2^2 \frac{\delta s_2}{s_2} + C_{e_3}^J \varepsilon_2^3 \frac{\delta s_2}{s_2} \right)$$

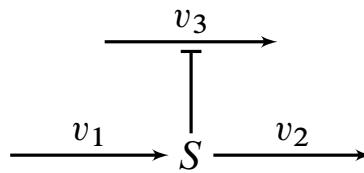
and therefore:

$$0 = \frac{\delta s_2}{s_2} \left( C_{e_2}^J \varepsilon_2^2 + C_{e_3}^J \varepsilon_2^3 \right)$$

Since  $\delta s_2/s_2$  is not equal to zero it must be true that:

$$0 = C_{e_2}^J \varepsilon_2^2 + C_{e_3}^J \varepsilon_2^3$$

This derivation can be applied to a species that interacts with any number of steps. In general, the number of terms will equal the number of interactions a species makes. For example, consider the pathway fragment in Figure 8.2.



**Figure 8.2** Pathway Fragment used in Connectivity Theorem Proof.

$S$  interacts with its production rate,  $v_1$ , a consumption rate,  $v_2$ , and an inhibitory interaction with  $v_3$ . The connectivity may therefore be written as:

$$C_{e_1}^J \varepsilon_s^1 + C_{e_2}^J \varepsilon_s^2 + C_3^J \varepsilon_s^3 = 0$$

For a species  $S$  that interacts with  $r$  other steps, the flux connectivity theorem is written as:

$$0 = \sum_{i=1}^r C_{e_i}^J \varepsilon_s^i$$

### Concentration Connectivity Theorem

To derive the flux connectivity theorem we used the system equation that was related to the flux. It is however possible to use a different set of system equations, those with respect to the species concentrations. In the case of the species, there will be two distinct system equations. One of these will describe the effect the modulation has on the common species ( $S_2$  in the example), and a second will describe the effect on any other species ( $S_1, S_3$ , etc.). Consider first the system equation involving the common species. For the pathway under consideration, this equation is given by:

$$\frac{\delta s_2}{s_2} = C_{e_2}^{s_2} \frac{\delta e_2}{e_2} + C_{e_3}^{s_2} \frac{\delta e_3}{e_3}$$

Given that the change in the common species,  $\delta s_2/s_2$ , is non-zero (Figure 8.1), substituting in the local equations, equation (8.7), leads to:

$$\frac{\delta s_2}{s_2} = -C_{e_2}^{s_2} \varepsilon_2^2 \frac{\delta s_2}{s_2} - C_{e_3}^{s_2} \varepsilon_2^3 \frac{\delta s_2}{s_2}$$

Since  $\delta s_2/s_2 \neq 0$ , the term  $\delta s_2/s_2$  can be canceled which leads to the **first** concentration connectivity theorem:

$$-1 = C_{e_2}^{s_2} \varepsilon_2^2 + C_{e_3}^{s_2} \varepsilon_2^3$$

A **second** theorem can be derived by considering the effect of the modulation on a distant species, for example  $S_3$ . In this case the system equation with respect to  $S_3$  becomes:

$$0 = \frac{\delta s_3}{s_3} = C_{e_2}^{s_3} \frac{\delta e_2}{e_2} + C_{e_3}^{s_3} \frac{\delta e_3}{e_3}$$

Note that the equation equals zero because the operations ensure that species other than the common species,  $S_2$ , do not change in concentration.

Substituting once again the local equations into the above system equation leads to:

$$0 = \frac{\delta s_3}{s_3} = -C_{e_2}^{s_3} \varepsilon_2^2 \frac{\delta s_2}{s_2} - C_{e_3}^{s_3} \varepsilon_2^3 \frac{\delta s_2}{s_2}$$

or:

$$0 = -\frac{\delta s_2}{s_2} (C_{e_2}^{s_3} \varepsilon_2^2 + C_{e_3}^{s_3} \varepsilon_2^3)$$

However,  $\delta s_2/s_2$  is not zero, therefore it must be the case that:

$$0 = C_{e_2}^{s_3} \varepsilon_2^2 + C_{e_3}^{s_3} \varepsilon_2^3$$

That completes the proof for the concentration connectivity theorems. As with the flux connectivity theorems, the concentration connectivity theorems can be generalized to any number of steps that a species might interact with.

To summarize:

### Flux Connectivity Theorem

With respect to a common metabolite,  $S_k$ , where  $r$  is the number of interactions  $S_k$  makes with neighboring reaction steps:

$$\sum_{i=1}^r C_{e_i}^J \varepsilon_{S_k}^{v_i} = 0 \quad (8.8)$$

### Concentration Connectivity Theorem

With respect to the common metabolite,  $S_k$ , where  $r$  is the number of interactions  $S_k$  makes with neighboring reaction steps:

$$\sum_{i=1}^r C_{e_i}^{s_k} \varepsilon_{s_k}^{v_i} = -1 \quad (8.9)$$

### Concentration Connectivity Theorem

With respect to the common metabolite,  $S_k$ , and a distant metabolite,  $S_m$ , where  $r$  is the number of interactions  $S_k$  makes with neighboring reaction steps. **Note:**  $k \neq m$ :

$$\sum_{i=1}^r C_{e_i}^{s_m} \varepsilon_{s_k}^{v_i} = 0$$

### Interpretation

Why are the connectivity theorems important? First and foremost, the theorems link local effects, in terms of the elasticities, to global effects, in terms of the control coefficients. Consider for example the following linear pathway:



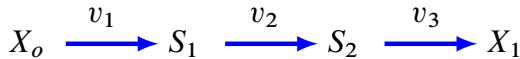
The flux connectivity can be written in the form:

$$\frac{C_{e_1}^J}{C_{e_2}^J} = -\frac{\varepsilon_1^2}{\varepsilon_1^1}$$

The ratio of two adjacent flux control coefficients is inversely proportional to the ratio of the corresponding elasticities. This means that high flux control coefficients tend to be associated with small elasticities, and small flux control coefficients with large elasticities.

This effect is explained by species opposing changes in rates. Since species with high elasticities are able to oppose rate changes more effectively than small elasticities, it follows that large elasticities are associated with small flux control coefficients and *vice versa*.

A good example of this is the case of a reaction operating near equilibrium where the elasticities are very high relative to adjacent elasticities on neighboring enzymes. In such situations the flux control coefficients of near equilibrium enzymes are *likely* to be small. However, one must bear in mind that it is the ratio of elasticities which is important and not



**Figure 8.3** Three step linear pathway.

their absolute values. Simply examining the elasticity of a single reaction is not sufficient to draw a firm conclusion. Even more so, one must also consider all the ratios of the elasticities along a pathway. Even though one elasticity ratio may suggest a high or low flux control coefficient on a particular enzyme, the other ratios coupled to the flux summation theorem must be considered because they will give an absolute value to a particular flux control coefficient. This reinforces the point that control coefficients are system-wide properties and cannot be reliably estimated from the properties of the step alone.

The examination of a single enzyme will not give an absolute indication of the ability of that enzyme to control the flux or species concentrations.

A second application of the connectivity theorems is for making simple statements about the properties of specific pathway configurations. Chapter 10 will delve more deeply into the properties of linear pathways, but here we can make some simple statements using the connectivity theorems. Consider the three-step pathway in Figure 8.3.

Two flux and four concentration connectivity theorems can be written for this pathway:

$$C_1^J \varepsilon_1^1 + C_2^J \varepsilon_1^2 = 0 \quad C_2^J \varepsilon_2^2 + C_3^J \varepsilon_2^3 = 0 \quad (8.10)$$

$$C_1^1 \varepsilon_1^1 + C_2^1 \varepsilon_1^2 = -1 \quad C_2^2 \varepsilon_2^2 + C_3^2 \varepsilon_2^3 = -1 \quad (8.11)$$

$$C_1^2 \varepsilon_1^1 + C_2^2 \varepsilon_1^2 = 0 \quad C_2^1 \varepsilon_2^2 + C_3^1 \varepsilon_2^3 = 0 \quad (8.12)$$

Consider the situation where there is no product inhibition on the first step, i.e.  $\varepsilon_1^1 = 0$ . Looking at the flux connectivity theorems and inserting zero for  $\varepsilon_1^1$ , the flux theorems can be rewritten as:

$$\begin{aligned} C_2^J \varepsilon_1^2 &= 0 \\ C_2^J \varepsilon_2^2 + C_3^J \varepsilon_2^3 &= 0 \end{aligned}$$

Assuming that  $\varepsilon_1^2$  is not zero, it must then be true that  $C_2^J = 0$ . Since  $C_2^J = 0$ , the second theorem is reduced to:

$$C_3^J \varepsilon_2^3 = 0$$

Again,  $\varepsilon_2^3$  is assumed to be non-zero, therefore it must be true that  $C_3^J = 0$ . From this simple analysis one can deduce that since  $C_2^J = C_3^J = 0$ , then by the flux summation theorem,  $C_1^J = 1$ .

If  $\varepsilon_2^2 = 0$ , one can similarly show that  $C_3^J = 0$  and  $C_1^J + C_2^J = 1$ . From this it can be concluded that if there is no product inhibition or reversibility in the first reaction then the first step fully controls the flux.

The concentration connectivity theorems can also be used to derive properties of the pathway. Assume as before that  $\varepsilon_1^1 = 0$ . Using the two connectivity theorems that include  $\varepsilon_1^1$ , and removing the term containing  $\varepsilon_1^1$  from each yields:

$$\begin{aligned} C_2^2 \varepsilon_1^2 &= 0 \\ C_2^1 \varepsilon_1^2 &= -1 \quad \Rightarrow \quad C_2^1 = -\frac{1}{\varepsilon_1^2} \end{aligned}$$

The first equation indicates that without product inhibition on the first step, the second step has no influence on  $S_2$ . Logically this makes sense because without product inhibition it is impossible to change the steady state flux. Since the flux through  $v_3$  must be unchanged it must be the case the  $S_2$  is also unchanged. The second equation is more interesting, it states that without product inhibition, the ability of the second step to influence  $S_1$  is inversely proportional to  $\varepsilon_1^2$ . For example, if  $\varepsilon_1^2$  is small, then the influence of the second step on  $S_1$  will be large. Note that the relationship also indicates via the sign, that an increase in activity of step two **reduces** the concentration of  $S_1$ , which makes logical sense.

Other patterns can be found in behavior by setting some of the other elasticities to zero. For example, if  $\varepsilon_2^3$  is set to zero (reaction step three is saturated), then  $C_2^J = 0$  and  $C_1^J = 0$ , meaning that  $C_3^J = 1$  and  $C_2^1 = 0$ .

### 8.3 Response Coefficients

---

Control coefficients measure the response of a pathway to changes in enzyme activity. What about the effect of external factors such as inhibitors, pharmaceutical drugs or boundary species? Such effects are measured by another coefficient called the **response coefficient**. The **flux response coefficient** is defined by:

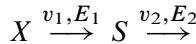
$$R_x^J = \frac{dJ}{dx} \frac{x}{J}$$

and the **concentration response coefficient** by:

$$R_x^s = \frac{ds}{dx} \frac{x}{s}$$

where  $x$  is the concentration of the external factor. The response coefficient measures how sensitive a pathway is to changes in external factors other than enzyme activities. What is the relationship of the response coefficients with respect to the control coefficients and elasticities of a pathway?

Like many of the proofs in this chapter, we can carry out a thought experiment to investigate the response coefficients more closely. Consider the pathway fragment below:



where  $X$  is the fixed boundary species. Let us increase the concentration of  $E_1$  by an amount  $\delta e_1$ . This will cause the steady state flux and concentration of  $S$  and in fact all downstream species beyond  $v_2$  to increase. Now decrease the concentration of  $X$  such that we restore the flux and steady state concentration of  $S$  back to their original values. From this thought experiment we can write the operations in terms of the local response equation and a system response equation as follows:

$$\frac{\delta v_1}{v_1} = \varepsilon_x^1 \frac{\delta x}{x} + \varepsilon_{e_1}^1 \frac{\delta e_1}{e_1} = 0 \quad \left. \right\} \text{Local equation}$$

$$\frac{\delta J}{J} = R_x^J \frac{\delta x}{x} + C_{e_1}^J \frac{\delta e_1}{e_1} = 0 \quad \left. \right\} \text{System equation}$$

Note that the right-hand sides are zero because the thought experiment guarantees that the flux has not changed. We can eliminate the  $\delta e_1/e_1$  term in the system response equation by substituting the term from the local response equation. In addition, if we assume that the reaction rate for an enzyme catalyzed reaction is proportional to the enzyme concentration, then we know from previous considerations that  $\varepsilon_{e_1}^1 = 1$ , see (7.7).

Therefore:

$$0 = R_x^J \frac{\delta x}{x} - C_{e_1}^J \varepsilon_x^1 \frac{\delta x}{x}$$

Since  $\delta x/x \neq 0$ , we can cancel  $\delta x/x$  to yield:

$$R_x^J = C_{e_1}^J \varepsilon_x^1 \quad (8.13)$$

This gives use a useful relationship. It can be generalized for multiple external factors acting simultaneously by summing up individual responses:

$$R_x^J = \sum_{i=1}^n C_{e_i}^J \varepsilon_x^{v_i}$$

Likewise, the response of a species,  $S$ , to an external factor,  $X$ , is given by:

$$R_x^s = \sum_{i=1}^n C_{e_i}^s \varepsilon_x^{v_i} \quad (8.14)$$

These relationships apply to any external factor that can affect one or more reaction rates. The relationship between the response coefficient, control and elasticity coefficients carries an important message. The response of an external factor,  $X$ , is a function of two things:

- 1) The effect the factor has on the step it acts upon.
- 2) The effect that the step itself has on changing the system.

This means than an effective external factor, such as a pharmaceutical drug, must not only be able to bind and inhibit the enzyme or protein being targeted, but the step itself must be able to transmit the disturbance to the rest of the pathway and ultimately affect the phenotype.

The ability of an external factor to influence a given species or flux depends on:

1. The ability of the external factor to influence its immediate target.
2. The ability of the target to influence the network it is connected to.

## 8.4 Canonical Control Coefficients

Let us write the response coefficient equation in a different way. Consider:

$$C_{e_i}^J = \frac{R_{e_i}^J}{\varepsilon_{e_i}^{v_i}}$$

Expand the terms and replace  $e_i$  with a general parameter  $p$  to yield:

$$C_{v_i}^J = \frac{\frac{dJ}{dp} \frac{p}{J}}{\frac{\partial v_i}{\partial p} \frac{p}{v_i}} = \frac{dJ}{dv_i} \frac{v_i}{J} \quad (8.15)$$

Equation (8.15) defines a control coefficient as a **parameterless control coefficient**,  $C_{v_i}^J$ . In this text such control coefficients will be called **canonical control coefficients** to distinguish them from enzyme based control coefficients. These coefficients describe the effect of a change in the local reaction rate on the steady state level of the pathway flux. This may seem like an odd definition; isn't the change in reaction rate the same as the change in steady state pathway flux? In this case, no. We have to be clear what the derivative,  $dv_i$ , in the denominator actually means. Operationally, the change indicated by  $dv_i$  refers to a change in  $v_i$  by some unspecified means under the conditions where the reactants, products and any other effectors remain constant.  $dv_i$  in this context is also sometimes referred to as the local rate, i.e. the change in the reaction rate we could impose if the reaction were not connected to the rest of the network. The way  $dv_i$  is changed is unspecified, but it must be done via a parameter of the system, not by a variable quantity such as one of the reactants or products.

The identity of the parameter will depend on the constraints imposed by the system, but common parameters could be the concentration of expressed enzyme, the catalytic constant of the enzyme, or an external inhibitor.

For example, suppose that the change to alter  $v_i$  is via a change in the enzyme concentration,  $E_i$ . Such a change will cause an immediate change,  $dv_i$ , in the reaction rate. The system is now allowed to evolve to its new steady state. The  $dv_i$  will cause changes in the immediate environment of the reaction, causing the enzyme's substrate to decrease and its product to increase. These changes in turn will propagate throughout the system. Once the system has settled to the new steady state, an inspection of  $v_i$  will reveal that the final change in rate does not equal the original  $dv_i$  (because the local environment has now changed). We refer to the final change in  $v_i$  as  $dJ$ , that is the change in flux through the system. Taking the ratio of  $dJ$  and  $dv_i$  and scaling, we obtain the canonical control coefficient. In many situations, the enzyme elasticity,  $\varepsilon_{e_i}^{v_i} = 1$ , which means that:

$$C_{e_i}^J = C_{v_i}^J$$

In other words, the control coefficients that are measured,  $C_e^J$ , are often identical to the canonical control coefficients. Strictly speaking, the summation and connectivity theorems only apply to the canonical control coefficients, but because  $\varepsilon_{e_i}^{v_i} = 1$ , we can often safely express the theorems using the control coefficients with respect to enzyme concentration.

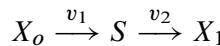
## 8.5 How to Derive Control Equations

---

A key aspect of MCA is being able to derive the control equations, that is the equations that relate the elasticities to the control coefficients. Section 8.1 illustrates one way that used the local and system equations. In this section a number of other approaches will be described, some of which can be automated by computer.

### Derivation using the Theorems

One way to derive the control equations is to combine the summation and connectivity theorems. For example, consider a two-step pathway such as:



where  $X_o$  and  $X_1$  are fixed species. There is one flux connectivity theorem with respect to every species in a pathway so that in the above example, there will only be one connectivity theorem centered around  $S$ :

$$C_{e_1}^J \varepsilon_s^1 + C_{e_2}^J \varepsilon_s^2 = 0$$

In addition, there will be a flux summation theorem:

$$C_{e_1}^J + C_{e_2}^J = 1$$

These two equations can be combined and solved for  $C_{e_1}^J$  and  $C_{e_2}^J$  to give expressions relating the control coefficients in terms of the elasticities, thus:

$$\begin{aligned} C_{e_1}^J &= \frac{\varepsilon_s^2}{\varepsilon_s^2 - \varepsilon_s^1} \\ C_{e_2}^J &= -\frac{\varepsilon_s^1}{\varepsilon_s^2 - \varepsilon_s^1} \end{aligned} \quad (8.16)$$

The concentration control coefficients can be derived in similar fashion by using the concentration control coefficient theorems. For a two-step pathway these would be:

$$\begin{aligned} C_{e_1}^s + C_{e_2}^s &= 0 \\ C_{e_1}^s \varepsilon_s^1 + C_{e_2}^s \varepsilon_s^2 &= -1 \end{aligned}$$

Combining both equations and solving for  $C_{e_1}^s$  and  $C_{e_2}^s$  yields the following control equations:

$$\begin{aligned} C_{e_1}^s &= \frac{1}{\varepsilon_s^2 - \varepsilon_s^1} \\ C_{e_2}^s &= -\frac{1}{\varepsilon_s^2 - \varepsilon_s^1} \end{aligned}$$

These are the same equations that were derived at the start of the chapter. For more complex pathways such as branches and moiety conserved cycles, additional theorems are required to solve the equations.

As a convenience, the theorems can be rendered in matrix form and solved using standard matrix methods. Both the flux and concentration control coefficient theorems can be recast in matrix form as follows:

$$\begin{bmatrix} 1 & 1 & 1 \\ \varepsilon_1^1 & \varepsilon_1^2 & 0 \\ 0 & \varepsilon_2^2 & \varepsilon_2^3 \end{bmatrix} \begin{bmatrix} C_1^J & C_1^{s_1} & C_1^{s_2} \\ C_2^J & C_2^{s_1} & C_2^{s_2} \\ C_3^J & C_3^{s_1} & C_3^{s_2} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad (8.17)$$

For example, the first row of the first matrix multiplied by the first column of the elasticity matrix yields the flux summation theorem. It is a simple matter to extend the matrix to any size linear pathway by following the pattern. For example, a four-step pathway is represented by:

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ \varepsilon_1^1 & \varepsilon_1^2 & 0 & 0 \\ 0 & \varepsilon_2^2 & \varepsilon_2^3 & 0 \\ 0 & 0 & \varepsilon_3^3 & \varepsilon_3^4 \end{bmatrix} \begin{bmatrix} C_1^J & C_1^{s_1} & C_1^{s_2} & C_1^{s_3} \\ C_2^J & C_2^{s_1} & C_2^{s_2} & C_2^{s_3} \\ C_3^J & C_3^{s_1} & C_3^{s_2} & C_3^{s_3} \\ C_4^J & C_4^{s_1} & C_4^{s_2} & C_4^{s_3} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

Multiplying both sides by the inverse of the matrix containing the elasticities yields:

$$\begin{bmatrix} C_1^J & C_1^{s_1} & C_1^{s_2} & C_1^{s_3} \\ C_2^J & C_2^{s_1} & C_2^{s_2} & C_2^{s_3} \\ C_3^J & C_3^{s_1} & C_3^{s_2} & C_3^{s_3} \\ C_4^J & C_4^{s_1} & C_4^{s_2} & C_4^{s_3} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ \varepsilon_1^1 & \varepsilon_1^2 & 0 & 0 \\ 0 & \varepsilon_2^2 & \varepsilon_2^3 & 0 \\ 0 & 0 & \varepsilon_3^3 & \varepsilon_3^4 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

One advantage of writing the equations in matrix form is that it makes it easy to evaluate the control coefficients by inverting the elasticity matrix. For example, if a three-step linear pathway has the following elasticities:

$$\begin{bmatrix} 1 & 1 & 1 \\ -0.6 & 1.2 & 0 \\ 0 & -0.2 & 0.5 \end{bmatrix}$$

Then the control coefficient matrix is given by:

$$\begin{bmatrix} 1 & 1 & 1 \\ -0.6 & 1.2 & 0 \\ 0 & -0.2 & 0.5 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} = \begin{bmatrix} 0.192 & 1.347 & 0.385 \\ 0.577 & -0.962 & 1.154 \\ 0.231 & -0.385 & -1.538 \end{bmatrix}$$

Note how the left column of values sum to one, reflecting the flux summation theorem. The second and third columns sum to zero, corresponding to the concentration summation theorem.

It is possible to rearrange the control matrix so that the flux control coefficients are positioned along the top row, and concentration control coefficients along the subsequent rows. Also, one can move the negative signs on the right-hand side of equation (8.17) to the elasticity terms so that the right-hand side becomes an identity matrix. See equation below, (8.18):

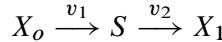
$$\begin{bmatrix} C_1^J & C_2^J & C_3^J \\ C_1^{s_1} & C_2^{s_1} & C_3^{s_1} \\ C_1^{s_2} & C_2^{s_2} & C_3^{s_2} \end{bmatrix} \begin{bmatrix} 1 & -\varepsilon_1^1 & 0 \\ 1 & -\varepsilon_1^2 & -\varepsilon_2^2 \\ 1 & 0 & -\varepsilon_2^3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (8.18)$$

For more complex pathways that include branches and cycles, this arrangement becomes more convenient.

### Pure Algebraic Method - Advanced Topic

The pure algebraic method relies on the use of implicit differentiation of the system equation and is a formalization of the thought experiments presented at the start of the chapter.

Consider the simplest two-step pathway already considered:



At steady state the rate of change of  $S$  is given by:

$$\frac{ds}{dt} = v_1 - v_2 = 0$$

Assuming that  $v_1$  can be changed by perturbations to the concentration of catalyzing enzyme,  $E_1$ , we can write the rate of change as:

$$\frac{ds}{dt} = v_1(s(e_1), e_1) - v_2(s(e_1)) = 0$$

In this equation each reaction rate,  $v_i$ , is a function of both the steady state species concentration and the perturbing parameter,  $E_1$ . Note that  $v_1$  is both a function of  $S$  and  $E_1$ , and  $S$  in turn is a function of  $E_1$ .  $v_2$  is only a function of  $S$  and not *directly* a function of  $E_1$  but indirectly via  $s$ . We can implicitly differentiate this equation with respect to  $e_1$  to yield:

$$0 = \frac{\partial v_1}{\partial s} \frac{ds}{de_1} + \frac{\partial v_1}{\partial e_1} - \frac{\partial v_2}{\partial s} \frac{ds}{de_1}$$

We can scale each of the derivatives by multiplying by the appropriate factors, that is:

$$0 = \frac{\partial v_1}{\partial s} \frac{s}{v_1} \frac{ds}{de_1} \frac{e_1}{s} + \frac{\partial v_1}{\partial e_1} \frac{e_1}{v_1} - \frac{\partial v_2}{\partial s} \frac{s}{v_1} \frac{ds}{de_1} \frac{e_1}{s}$$

which can be simplified to:

$$0 = C_{e_1}^s \varepsilon_s^1 + \varepsilon_{e_1}^1 - C_{e_1}^s \varepsilon_s^2$$

Solving for  $C_{e_1}^s$  and assuming  $v_1$  is first-order with respect to  $e_1$  so that  $\varepsilon_{e_1}^1 = 1$ , yields:

$$C_{e_1}^s = \frac{1}{\varepsilon_s^2 - \varepsilon_s^1} \quad (8.19)$$

We can derive  $C_{e_2}^s$  in the same way by implicitly differentiating:

$$\frac{ds}{dt} = v_1(s(e_2)) - v_2(s(e_2), e_2) = 0$$

The flux control coefficients can be computed in a similar way. For example, to find  $C_{e_1}^J$  we can implicitly differentiate:

$$\begin{aligned} J &= v_1(s(e_1), e_1) \\ \frac{dJ}{de_1} &= \frac{\partial v_1}{\partial s} \frac{ds}{de_1} + \frac{\partial v_1}{\partial e_1} \end{aligned}$$

Scaling yields:

$$C_{e_1}^J = C_{e_1}^s \varepsilon_s^1 + 1$$

Substituting  $C_{e_1}^s$  gives:

$$C_{e_1}^J = -\frac{1}{\varepsilon_s^1 - \varepsilon_s^2} \varepsilon_s^1 + 1 = \frac{\varepsilon_s^1}{\varepsilon_s^2 - \varepsilon_s^1}$$

The use of implicit differentiation can also be cast into matrix form. Recall that the system equation is given by:

$$\frac{ds}{dt} = \mathbf{N} \mathbf{v}(\mathbf{s}(\mathbf{e}), \mathbf{e}) \quad (8.20)$$

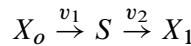
where  $\mathbf{N}$  is the stoichiometry matrix,  $\mathbf{v}$  the rate vector, and  $ds/dt$  the rate of change vector. The rate vector has been made an explicit function of the concentrations,  $\mathbf{s}$ , and the vector of enzyme concentrations,  $\mathbf{e}$ . At steady state this equation is equal to zero and differentiation with respect to  $\mathbf{e}$  yields the following:

$$0 = \mathbf{N} \frac{\partial \mathbf{v}}{\partial \mathbf{s}} \frac{d\mathbf{s}}{d\mathbf{e}} + \mathbf{N} \frac{\partial \mathbf{v}}{\partial \mathbf{e}}$$

Rearrangement gives:

$$\frac{d\mathbf{s}}{d\mathbf{e}} = \left( \mathbf{N} \frac{\partial \mathbf{v}}{\partial \mathbf{s}} \right)^{-1} \mathbf{N} \frac{\partial \mathbf{v}}{\partial \mathbf{e}} \quad (8.21)$$

The left-hand side represents the unscaled concentration control coefficient.  $\partial \mathbf{v} / \partial \mathbf{s}$  is a  $n$  by  $m$  matrix of unscaled elasticity coefficients, where  $n$  is the number of reactions and  $m$  the number of floating species.  $\partial \mathbf{v} / \partial \mathbf{e}$  is a  $n$  by  $n$  matrix of elasticities with respect to each enzyme,  $e$ . To illustrate how this equation can be used, consider the two-step pathway:



where  $X_o$  and  $X_1$  are fixed species. In this model,  $n = 2$ ,  $m = 1$ . With two steps the  $\mathbf{e}$  vector will have two entries,  $e_1$  and  $e_2$ . The elasticity matrix will be a 2 by 1 matrix:

$$\begin{bmatrix} \frac{\partial v_1}{\partial s} \\ \frac{\partial v_2}{\partial s} \end{bmatrix}$$

and  $\partial \mathbf{v} / \partial \mathbf{e}$  is a 2 by 2 matrix:

$$\begin{bmatrix} \frac{\partial v_1}{\partial e_1} & 0 \\ 0 & \frac{\partial v_2}{\partial e_2} \end{bmatrix}$$

Note that  $e_1$  has no direct affect on  $v_1$  and  $e_2$  has no direct effect on  $v_1$ . Combined together:

$$\begin{bmatrix} \frac{ds}{de_1} \\ \frac{ds}{de_2} \end{bmatrix} = \left( \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} \frac{\partial v_1}{\partial s} \\ \frac{\partial v_2}{\partial s} \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} \frac{\partial v_1}{\partial e_1} & 0 \\ 0 & \frac{\partial v_2}{\partial e_2} \end{bmatrix}$$

Multiplied out we obtain:

$$\begin{bmatrix} \frac{ds}{de_1} \\ \frac{ds}{de_2} \end{bmatrix} = \begin{bmatrix} \frac{\partial v_1 / \partial e_1}{\partial v_1 / \partial s - \partial v_2 / \partial s} \\ \frac{\partial v_1 / \partial e_2}{\partial v_1 / \partial s - \partial v_2 / \partial s} \end{bmatrix}$$

Each entry can be scaled by multiplying both sides by  $e_1$  and dividing by  $s$ , for example, to scale  $\partial s / \partial e_1$ :

$$\frac{ds}{de_1} \frac{e_1}{s} = \frac{e_1}{v_1} \frac{\partial v_1}{\partial e_1} \frac{1}{\frac{\partial v_1}{\partial s} \frac{s}{v_1} - \frac{\partial v_2}{\partial s} \frac{s}{v_2}}$$

Note the trick for scaling the denominator terms on the right-side where top and bottom are multiplied by  $v_1$ , noting that  $v_1 = v_2$  at steady state. Rearranging yields:

$$C_{e_1}^s = \varepsilon_{e_1}^1 \frac{1}{\varepsilon_s^1 - \varepsilon_s^{v_2}}$$

As before,  $\varepsilon_{e_1}^1 = 1$  so that:

$$C_{e_1}^s = \frac{1}{\varepsilon_s^1 - \varepsilon_s^2}$$

This is the same equation as the one previously derived (8.19). The one limitation of this approach is that it cannot, without a minor modification, be used on pathways that include conserved moieties. This is because the inverse in equation (8.21) cannot be evaluated. This issue will be revisited in a later chapter when moiety cycles are discussed in some depth.

The flux control coefficients can be similarly derived. The system equation for fluxes is shown below:

$$\mathbf{J} = \mathbf{v}(\mathbf{s}(\mathbf{e}), \mathbf{e})$$

Differentiating with respect to  $\mathbf{e}$  gives:

$$\frac{d\mathbf{J}}{d\mathbf{e}} = \frac{\partial \mathbf{v}}{\partial \mathbf{s}} \frac{d\mathbf{s}}{d\mathbf{e}} + \frac{\partial \mathbf{v}}{\partial \mathbf{e}}$$

Scaling with  $\mathbf{J}$ ,  $\mathbf{v}$  and  $\mathbf{e}$  yields:

$$\mathbf{C}_e^J = \boldsymbol{\varepsilon}_s^v \mathbf{C}_e^s + \boldsymbol{\varepsilon}_e^v$$

If we assume  $v \propto e_i$  then  $\boldsymbol{\varepsilon}_e^v = \mathbf{I}$ , where  $\mathbf{I}$  is the identity matrix, so that:

$$\mathbf{C}_e^J = \boldsymbol{\varepsilon}_s^v \mathbf{C}_e^s + \mathbf{I}$$

The matrix method is useful because can be encoded using symbolic algebra tools such as Mathematica or Maxima.

## 8.6 Relationship to S-Systems

---

An alternative approach to MCA is Biochemical Systems Theory, or BST that was developed by Savageau and colleagues over many years. There is considerable overlap between BST and MCA. The main difference is the emphasis that BST places on developing approximate simulation models based on power laws:

$$\alpha x^g$$

A particular type of power law formalism, called S-systems, consists of two terms. The first term aggregates all inputs to a species as a single power law term. The second term aggregates all consumption terms as a single power law and is subtracted from the first term. A single S-systems equation that describes the rate of change of a species  $X$ , would look like:

$$\frac{dx}{dt} = \alpha \prod_{j=1}^n x_j^{g_{ij}} - \beta_i \prod_{j=1}^n x_j^{h_{ij}}$$

$\alpha$  and  $\beta$  are called the rate constants, and  $g$  and  $h$  the kinetic orders. The product terms include all species that might influence production in the first term, and all species that might influence consumption in the second term.

For example, consider a branch with a single reaction,  $v_1$ , producing a species,  $x$ , and two reactions,  $v_2$  and  $v_3$  consuming  $x$ . The differential equation for this system is:

$$\frac{dx}{dt} = v_1 - (v_2 + v_3)$$

The S-system equation for the rate of change in  $x$  is given by:

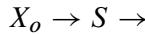
$$\frac{dx}{dt} = \alpha x^g - \beta x^h$$

Note that the two consumption steps have been merged into a single term  $\beta x^h$ . In the process of aggregating, some information is lost. For example, information on the relative magnitudes of the consumption rates,  $v_2$  and  $v_3$ , is missing and cannot be recovered. The chief advantage in aggregating is that the equation can be analytically solved for the steady state solution, something that is invariably impossible to do otherwise. The analysis of S-system models however leads to results that are identical to those of MCA, and many of the terms can be interchanged between the two approaches.

Instead of elasticities, BST uses the term kinetic orders. Instead of control coefficients, BST uses the term logarithmic gains. The main difference is that the analysis of branched systems or systems with moiety conserved cycles (Chapter 16) can be more difficult when using BST. The approaches used to derive the logarithmic gains are completely different in BST and MCA. BST differentiates the analytical steady state solutions from the S-system. MCA can derive the same results using operational proofs, implicit differentiation or a variety of matrix

methods. The stability of steady state solutions – see Chapter 13 – is automatically obtained from BST, whereas with MCA, stability must be determined by computing the Jacobian separately. Finally, the theorems that are highlighted by MCA are of secondary importance to BST and are rarely mentioned. Thus there are pros and cons to both approaches. In this textbook the focus is obviously MCA.

The approach BST takes can be illustrated with a simple two-step pathway:



The S-system equation can be written as:

$$\frac{dx}{dt} = \alpha x_o^{g_1} s^{g_2} - \beta s^h$$

Note that the first reaction  $X_o \rightarrow s$  has two terms raised to a power because the reaction is influenced by both  $X_o$  and  $S$ . To compute the steady state level of  $S$ ,  $dx/dt$  can be set to zero such that:

$$\alpha x_o^{g_1} s^{g_2} = \beta s^h$$

Further rearrangement leads to:

$$\frac{\beta}{\alpha} = x_o^{g_1} s^{g_2} s^{-h}$$

Taking logs on both sides:

$$\log\left(\frac{\beta}{\alpha}\right) = g_1 \log(x_o) + (g_2 - h) \log(s)$$

And solving for  $\log(s)$ :

$$\log(s) = \frac{\log\left(\frac{\beta}{\alpha}\right) - g_1 \log(x_o)}{(g_2 - h)} \quad (8.22)$$

Recall that the definition (7.1) of the control coefficients was given in terms of the logarithmic derivative:

$$C = \frac{\partial \log V}{\partial \log p}$$

where  $V$  is the variable and  $p$  the parameter. Given that in the previous example the steady state solution is in terms of  $\log(s)$ , it is straightforward to compute the control coefficients (or logarithmic gains) by differentiating equation (8.22) with respect to one of the parameters. For example, the logarithmic gain for the steady state concentration of  $S$  with respect to the rate constants  $\alpha$  and  $\beta$  can be computed as follows. To make things clearer, equation (8.22) is reexpressed as:

$$\log(s) = \frac{\log(\beta) - \log(\alpha) - g_1 \log(x_o)}{(g_2 - h)}$$

Therefore, differentiating yields:

$$\frac{\partial \log(s)}{\partial \log(\alpha)} = \frac{1}{h - g_2} \quad \frac{\partial \log(s)}{\partial \log(\beta)} = -\frac{1}{h - g_2} \quad (8.23)$$

Note that the signs have been swapped in the numerators and denominators. Similarly, the logarithmic gain with respect to  $x_o$  can also be derived:

$$\frac{\partial \log(s)}{\partial \log(x_o)} = \frac{g_1}{h - g_2}$$

The logarithmic gains  $\partial \log(s)/\partial \log(\alpha)$  and  $\partial \log(s)/\partial \log(\beta)$  correspond to the control coefficients,  $C_{e1}^s$  and  $C_{e2}^s$ . Since the kinetic orders  $g_1$  and  $h$  are equivalent to the elasticities  $\varepsilon_s^1$  and  $\varepsilon_s^2$  respectively, the analysis using BST and MCA yields identical results:

$$C_{e1}^s = \frac{\partial \log(s)}{\partial \log(\alpha)} = \frac{1}{\varepsilon_s^2 - \varepsilon_s^1} = \frac{1}{h - g_2}$$

$$C_{e2}^s = \frac{\partial \log(s)}{\partial \log(\beta)} = -\frac{1}{\varepsilon_s^2 - \varepsilon_s^1} = -\frac{1}{h - g_2}$$

The logarithmic gain  $\partial \log(s)/\partial \log(x_o)$  corresponds to the response coefficient,  $R_{x_o}^s$ . Recall that the response coefficient is equal to the elasticity,  $\varepsilon_{x_o}^1$  times the concentration control coefficient  $C_{e1}^s$ , that is:

$$R_{x_o}^s = \varepsilon_{x_o}^1 C_{e1}^s$$

In the BST formalism,  $\varepsilon_{x_o}^1 = g_1$ , therefore we can make the following equivalence:

$$R_{x_o}^s = \varepsilon_{x_o}^1 C_{e1}^s = \varepsilon_{x_o}^1 \frac{1}{\varepsilon_s^2 - \varepsilon_s^1} = \frac{\partial \log(s)}{\partial \log(x_o)} = g_1 \frac{1}{h - g_2}$$

Table 8.1 summarizes some of the equivalent terms between MCA and BST.

Using this small example it is evident that MCA and BST are very similar. Perhaps by historical accident MCA is more developed given its European origins where theory tends to get more attention. Both approaches ask some of the same questions and yield the same answers. The use of power laws in deriving the steady state solutions and logarithmic gains is certainly elegant, but the aggregation that S-systems require poses challenges for more complex pathways.

MCA derives theorems which BST does not emphasize. However the theorems allow additional insight to be gained. MCA also focuses much more on fluxes through pathways which BST tends to be more silent on. Steady state stability is dealt with more explicitly in BST and has resulted in a variety of very useful results (see later chapters), especially for oscillating systems. However with recent developments in the theory, MCA can accomplish the same thing.

MCA has been integrated much more fully into stoichiometric analysis and offers an elegant mathematical framework for dealing with conserved moieties and branched systems (see later chapters). The use of generalized power laws where aggregation is not required may offer a useful approach to building approximate models when data is limited.

MCA Symbol		BST Symbol	
Control Coefficient	$C_{e_1}^s$	Logarithmic Gain	$\frac{\partial \log(s)}{\partial \log(\alpha)}$
Control Coefficient	$C_{e_2}^s$	Logarithmic Gain	$\frac{\partial \log(s)}{\partial \log(\beta)}$
Response Coefficient	$R_{x_o}^s$	Logarithmic Gain	$\frac{\partial \log(s)}{\partial \log(x_o)}$
Elasticity	$\varepsilon_s^1$	Kinetic Order	$g_2$
Elasticity	$\varepsilon_s^2$	Kinetic Order	$h$
Elasticity	$\varepsilon_{x_o}^1$	Kinetic Order	$g_1$

**Table 8.1** Summary of equivalent terms used in MCA and BST.

---

## Further Reading

1. Kacser H and Burns JA, (1973). In: Davies, D.D (ed.), Rate Control of Biological Processes, vol. 27 of Symp. Soc. Exp. Biol. p. 65-104. Cambridge University Press.
2. Kacser H and Burns J (1979) Molecular Democracy: Who Shares the Controls? Biochem Soc Trans, 7, 1149-1160.
3. Heinrich R and TA Rapoport (1974) A linear steady-state treatment of enzymatic chains. General properties, control and effector strength. Eur. J. Biochem. 42:89-95.
4. Heinrich R, Rapoport SM, Rapoport TA (1977) Metabolic regulation and mathematical models. Prog Biophys Mol Biol. 1977;32(1):1-82.
5. Fell D A (1996) Understanding the Control of Metabolism, Frontiers in Metabolism, ISBN-10: 185578047X
6. Savageau M (1976) Biochemical systems analysis: a study of function and design in molecular biology, Addison-Wesley. Note: The text was republished in 2010 and is available on Amazon.

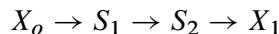
---

## Exercises

1. A given species  $S$  has a single production step and a single consumption step. The elasticity of the production step with respect to  $S$  was found to be -1.6, and for the

consumption step, 0.12.

- a) Explain why the production step elasticity is negative.
- b) From the information, what can you say about the flux control coefficients of the production and consumption steps?
2. Explain why examination of a single enzyme in a pathway will not necessarily give a good indication of how flux limiting the enzyme is.
3. The response coefficient relationship has two important lessons for those looking to develop new therapeutic drugs. What are they?
4. What is a canonical control coefficient?
5. Given the model described by the Tellurium script (8.1), ignore everything in the script after line 14. Find the flux control coefficient for each step and then confirm numerically that the flux summation theorem holds. Use perturbations to estimate the flux control coefficients.
6. Use BST to derive the logarithmic gains for  $S_1$  and  $S_2$  with respect to the rate constants in all three steps for the three-step pathway:



where  $X_o$  and  $X_1$  are fixed (boundary) species. Confirm that the results you get correspond to those obtained using MCA.

## 8.A Python/Tellurium Scripts

Calculating control coefficients and elasticities using Python and Tellurium.

```
import tellurium as te
import numpy as np
r = te.loada ('''
    J1: $Xo -> S1; E1*(k1*Xo - k2*S1);
    J2: S1 -> S2; E2*(k3*S1 - k4*S2);
    J3: S2 -> $X1; E3*(k5*S2);

    Xo = 5
    k1 = 0.2; k2 = 0.04;
    k2 = 0.7; k3 = 0.34;
    k4 = 0.65;
    E1 = 1; E2 = 1; E3 = 1;
''')

print r.steadyState()
```

```
print "Flux Control Coefficients:"  
print r.getCC ("J1", "E1"), r.getCC ("J1", "E2")  
print r.getCC ("J1", "E3")  
  
print "S1 Concentration Control Coefficients:"  
print r.getCC ("S1", "E1"), r.getCC ("S1", "E2")  
print r.getCC ("S1", "E3")  
  
print "S2 Concentration Control Coefficients:"  
print r.getCC ("S2", "E1"), r.getCC ("S2", "E2")  
print r.getCC ("S2", "E3")  
  
print "Elasticities:"  
print r.getEE ("J1", "Xo"), r.getEE ("J2", "S1")  
print r.getEE ("J2", "S1"), r.getEE ("J2", "S2")  
print r.getEE ("J3", "S2"),
```

**Listing 8.1** Computing Control Coefficients.

# 9

## ***Experimental Methods***

---

### **9.1 Introduction**

---

Up to this point, we've largely discussed the theory of metabolic control analysis. However, without the ability to make measurements and test predictions on real biological systems, the approach outlined in the previous chapters will find little practical use. In this chapter we will outline various methods that have been developed to measure control coefficients and some examples of applications used in real metabolic pathways.

#### **Measuring Control Coefficients**

Various approaches have been used in the past to measure control coefficients experimentally. All revolve around the need to change either the concentration or activity of an enzyme or protein. The different methods can be categorized into six general approaches:

- 
1. Use of classical genetics to manipulate gene expression(s)
  2. Titration of enzymes with specific inhibitors
  3. Double modulation method
  4. *in vitro* reconstitution and enzyme titration
  5. Gene engineering to change enzyme levels *in vivo*

---

  6. Computer modeling (not really an experimental method)
-

## 9.2 Using Classical Genetics

---

There is an interesting story related to the origins of metabolic control analysis from the Kacser lab<sup>1</sup>. They noticed that a mutation in one of the amino acid biosynthesis enzymes in the fungus *Neurospora crassa* yielded a loss of 95% enzyme activity, and yet this hardly affected the observed phenotype. How was this possible? Instinctively, one might try to answer this question by studying the enzyme in question, perhaps studying its kinetics in detail or understanding its catalytic activity by determining its protein structure. However this would not ultimately explain the effect of enzyme loss on the phenotype. In an insightful move, the Kacser lab decided that the answer must lie in the enzyme's cellular context. That is, the network within which the enzyme operated. It was this observation that began the study of how networks behaved and how phenotype was related to phenotype, not in terms of individual proteins or enzymes, but in terms of networks. Similar conclusions were made by two other groups who also developed MCA independently, namely Heinrich and Rapoport in Berlin and Savageau in Michigan, USA.

The Kacser lab investigated the amino acid biosynthesis pathway by changing the copy number of a gene using classical genetics [41]. They assumed that the activity or concentration of the enzyme or protein in question was proportional to the copy number. This work was carried out using the fungus *Neurospora crassa* where arginine biosynthesis was studied. This fungus forms multinucleated mycelia that generate polyploid spores. By mixing different ratios of spores containing genes encoding wild and mutant enzymes, it was possible to generate mycelia with different activities of the arginine pathway enzymes. From such experiments it was determined that four enzymes, acetyl-ornithine aminotransferase, ornithine transcarbamoylase, arginine succinate synthetase, and arginine-succinate lyase all had flux control in the range of 0.02 to 0.2. This indicated that none of the enzymes exerted significant control over arginine synthesis.

Another study by the same group investigated the flux control of alcohol dehydrogenase (ADH) in *Drosophila melanogaster*. ADH is present in three alleles that encode isoforms with different maximal activities. When mixing the various combinations (including mutations of the isoforms), it was possible to change the total activity of ADH and measure the ethanol production. From this study it was concluded that ADH has a flux control coefficient of zero.

## 9.3 Genetic Engineering

---

*Neurospora crassa* and *Drosophila melanogaster* were special cases where gene dosage could be engineered by classical genetic methods. However most systems are not so easily altered. As a result, alternative methods based on inhibitors were developed. In particular oxidative phosphorylation is susceptible to a large repertoire of inhibitors. Control coeffi-

---

<sup>1</sup>personal communication

cients were estimated by titrating in a given inhibitor and measuring the effect on a flux or species concentration [48]. By taking into account how the inhibitor acted, it is possible to obtain estimates for the control coefficient of the inhibited step. We will come back to this in the next section.

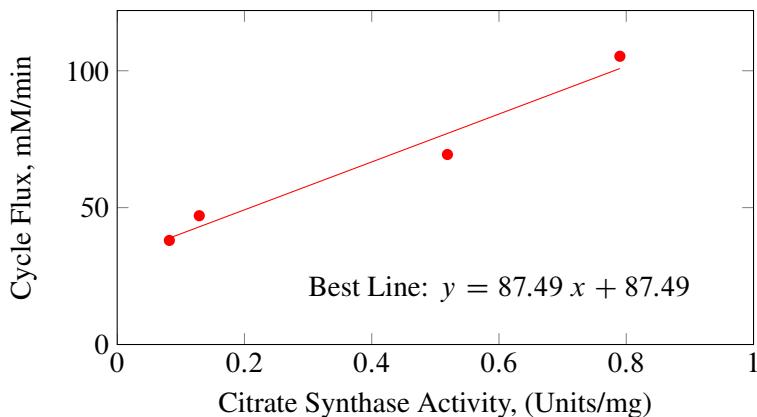
With modern developments in genetic engineering and molecular biology, it is now relatively straightforward to estimate control coefficients. For example, inducible or repressor operator sites can be added to a gene of interest and the effect of up regulating or down regulating the wild-type activity can be measured. In addition, protein levels can be knocked down by using a variety of method such as RNA antisense. This method was used to estimate the flux control coefficient for ribulose-bisphosphate-carboxylase (rubisco) which is responsible for fixing carbon dioxide in plants [129]. The traditional view has been that Rubisco catalyzes a rate-limiting step, that is a step with a high flux control coefficient. However, studies using DNA antisense in tobacco plants to reduce the level of Rubisco show that during high illumination, flux control was estimated in the range of 0.69 to 0.83. In moderate illumination or high carbon dioxide levels, the flux control fell dramatically to 0.05 to 0.2. This study reconfirms two important points that were highlighted in the theoretical analysis, the first is that control is not fixed but depends on external conditions. Secondly, rate limitingness cannot be determined by simple inspection of the pathway but instead must be actively measured.

One of the earliest uses of an inducible promoter to control gene expression and thereby estimate rate control was the work by Walsh and Koshland [147]. In their paper they describe using the inducer IPTG (isopropyl- $\beta$ -D-thiogalactopyranoside) to control the Krebs cycle enzyme citrate synthase in *E. coli* via a *tac* promoter<sup>2</sup>. The flux through citrate synthase was measured indirectly by measuring radiolabelled substrate incorporation into carbon dioxide, fatty acids and cellular mass. Figure 9.1 shows the data replotted from their paper (note errors bars were displayed in the original data).

Using this data, Walsh and Koshland suggested that under an acetate medium, citrate synthase was rate-controlling. However, the extent of control was not considered. Using their data, it is relatively easy to compute the flux control coefficient for citrate synthase. Given the straight line fit through the points, the slope of the line is 87.49 flux units per citrate synthase activity. In order to relate this to the flux control coefficient, the slope must be multiplied by the enzyme activity and divided by the flux. This is easily done using the regression parameters from the straight line fit. Figure 9.2 shows the flux control coefficient plotted as a function of citrate synthase activity. The first thing to note is that at the wild-type point, citrate synthase is not completely rate limiting, but has a flux control coefficient of 0.65. This means that a 1% increase in citrate synthase activity will lead to a 0.65% change in cycle flux. Because the coefficient is less than one, flux control must also reside elsewhere.

The second and somewhat unusual observation is that the flux control coefficient decreases as the citrate synthase activity decreases. Under normal circumstances the expectation is that

<sup>2</sup>The *tac* promoter is a synthetic promoter made from a combination of the lac and trp promoters.



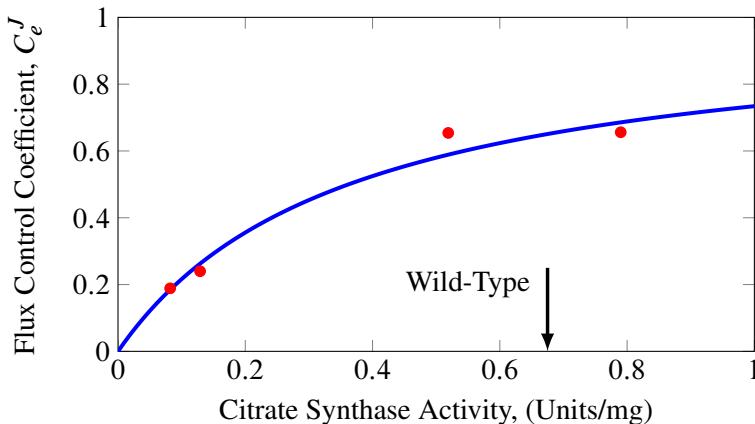
**Figure 9.1** Data from Walsh and Koshland [147]. The cycle flux represents the sum of the Krebs and Glyoxylate fluxes under acetate medium conditions. Wildtype level of citrate synthase is round at 0.675 units/mg. Original error bars are omitted from this plot.

flux control *increases* as enzyme activity decreases. There are a number of possible reasons for this: i) the data could be wrong; ii) the straight line fit through the data is inappropriate. If the flux had been measured at lower and higher levels of citrate synthase, the data would show more of a hyperbolic like response. Meaning that flux control would be high at low enzyme activities, and low at higher enzyme activities; iii) there are unrecorded changes to other enzymes in the pathway as a result of changes to the citrate synthase level. For example, it is possible that other enzymes are expressed less, such that any flux control is transferred from these other enzymatic steps to citrate synthase so that the flux control at citrate synthase appears to rise as citrate synthase is decreased.

In the original paper Walsh and Koshland did not compute the flux control coefficient. Instead they assumed that the linear relationship meant that flux control at citrate synthase was fixed and high. A closer investigation reveals the opposite, and also raises some additional and interesting questions.

## 9.4 Titration by Inhibitors

The first approach devised to estimate control coefficients was the use of classical genetic to manipulate gene copy number. This however is difficult to do and requires considerable expertise in genetic manipulation. In the late 1970s and early 1980s, researchers realized that inhibitors could also be used to change protein and enzyme activity. One of the most well studied pathways is oxidative phosphorylation, due in part to the wide variety of available inhibitors. Inhibitors include irreversible, for example cyanide can bind to cytochrome c oxidase (Site 3), noncompetitive inhibitors such as Rotenone that can bind to NADH-CoQ-oxidoreductase (Site I) or Antimycin that can bind to CoA-cytochrome c oxidoreductase



**Figure 9.2** Data from Walsh and Koshland [147]. Flux control coefficient computed using best line fit from Figure 9.1. At the wildtype level of citrate synthase, the flux control coefficient has a value of 0.65 indicated by the arrow. The response is completely counter intuitive to what one would expect and suggests that either the data is incomplete or that there are other physiological changes taking place.

and a competitive inhibitor such as malate that binds to dicarboxylate transporter.

By titrating the inhibitor and measuring the steady state response, it is possible to estimate the control coefficient at the inhibited site by extrapolating the response curve back to zero inhibitor. Each inhibitor type however must be treated differently [48]. In a previous chapter the response coefficient was introduced:

$$R_x^J = \frac{dJ}{dx} \frac{x}{J}$$

The response coefficient describes the effect of an external signal,  $X$ , on a pathway. The response coefficient can be expressed in terms of the control coefficient and the external signal elasticity using the relationship:

$$R_x^J = C_e^J \varepsilon_x^v$$

Expanding this relationship yields:

$$\frac{dJ}{dx} \frac{x}{J} = C_e^J \frac{\partial v}{\partial x} \frac{x}{v} \quad (9.1)$$

We can cancel the  $x$  terms on both sides and rearranging so that  $C_e^J$  is on one side we get:

$$C_e^J = \frac{dJ}{dx} \frac{1}{J} / \frac{\partial v}{\partial x} \frac{1}{v}$$

The term  $\frac{dJ}{dxJ}$  can be derived from the initial slope of the inhibition curve, and  $v\partial v/\partial x$  from the inhibition characteristics of the inhibited enzyme. Since we seek the value of  $C_e^J$  when there is no inhibitor present, we should measure both terms when  $x = 0$ . We can illustrate this approach with a number of examples. Consider first a non-competitive inhibitor. In the simplest case the rate of an enzymatic reaction subject to non-competitive inhibition is given by the well known equation:

$$v = \frac{V_m s}{(K_m + s)(1 + i/K_i)}$$

The elasticity for this rate law with respect to inhibitor concentration,  $I$  can be shown to be equal to:

$$\varepsilon_i^v = -\frac{i}{K_i + i}$$

Substituting this into equation (9.1) and setting the inhibitor concentration to zero, ( $i = 0$ ) we arrive at an expression for the control coefficient when the inhibitor concentration is extrapolated to zero:

$$C_x^J = -\frac{K_i}{J} \frac{dJ}{dx}$$

This approach requires an estimate for the inhibition constant,  $K_i$  and the slope of the response flux versus inhibitor at zero inhibitor. The same analysis can be done for competitive inhibitors. The irreversible rate law for a competitive inhibitor is given by the well known equation:

$$v = \frac{V_m s}{s + K_m \left(1 + \frac{i}{K_i}\right)}$$

The elasticity for this rate law with respect to the competitive inhibitor,  $I$ , can be shown to be equal to:

$$\varepsilon_i^v = -\frac{i/K_i}{1 + s/K_m}$$

Inserting this into equation (9.1) and setting the inhibitor concentration to zero ( $i = 0$ ) we arrive at an expression for the control coefficient based on titrating an enzymatic step with a competitive inhibitor:

$$C_e^J = -\frac{K_i}{J} \left(1 + \frac{s}{K_m}\right) \frac{dJ}{di}$$

This can be generalized if need be to the reversible case:

$$C_e^J = -\frac{K_i}{J} \left(1 + \frac{s}{K_s} + \frac{p}{K_p}\right) \frac{dJ}{di}$$

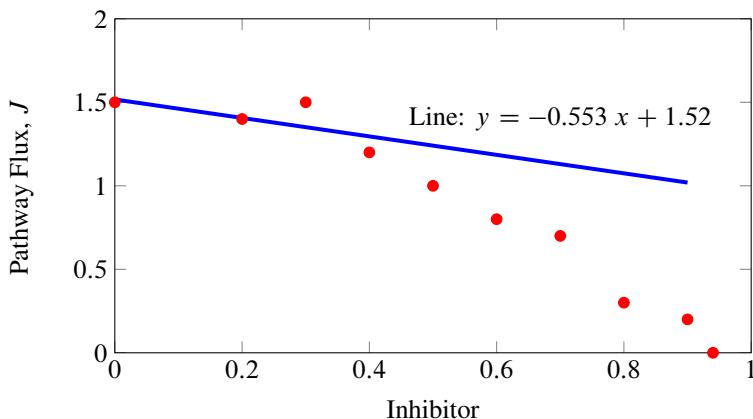
Where  $s$  is the concentration of substrate and  $p$  the concentration of product.  $K_s$  and  $K_p$  are the  $K_m$  constants for the substrate and product respectively.

**Example 9.1**

The following data (constructed from a simulated pathway with added noise) was collected from a pathway and measures the flux through the pathway at various concentrations of an irreversible inhibitor. Use the data to estimate the flux control coefficient through the pathway.

Inhibitor Concentration	Pathway Flux
0	1.5
0.2	1.4
0.3	1.5
0.4	1.2
0.5	1.0
0.6	0.8
0.7	0.7
0.8	0.3
0.9	0.2
0.94	0.0

A plot of the data is shown below:



The curve that follows the points is not known, which means it is difficult to select a suitable nonlinear function to fit the points. Instead we will take the first four data points and plot a straight line through them. This is shown by the blue continuous line on the plot. The slope of the fitted line was found to be  $-0.553$ . This corresponds to the  $dJ/dx$  term in equation (9.2). The flux at zero inhibitor is 1.5 and the  $I_{max}$  is 0.94, the concentration of inhibitor that yields zero flux. Taking these together and inserting the values into equation (9.2), we obtain the flux control coefficient:

$$C_e^J = -0.553 \frac{0.94}{1.5} = 0.35$$

The simulation model that was used to obtain the data gave a control coefficient of 0.367 which is close to the estimated value. The key to obtaining a reasonable estimate is to secure sufficient points at low inhibitor concentration in order to compute a best line fit through the first few points. Attempts to fit polynomials, logistic curves, or hyperbolic curves will likely yield poor estimates.

## Irreversible Inhibitors

Irreversible inhibitors are a special case; the amount of inhibitor required to completely inhibit the enzyme should be equal to the amount of enzyme. This assumes that one molecule of inhibitor binds to a single enzyme and completely inhibits the enzyme's activity. That is,  $x_{max} = e$ . The equation to compute the control coefficient using an irreversible enzyme is given by:

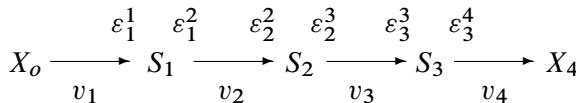
$$C_e^J = -\frac{x_{max}}{J} \frac{dJ}{dx} \quad (9.2)$$

The negative sign is included because the slope of  $dJ/dx$  is negative.

## 9.5 Double Modulation Technique

---

The double modulation method, first proposed by Kacser and Burns in 1979 is an elegant method for estimating elasticities *in vivo* but is difficult to execute experimentally. Consider the pathway:



Let us focus on reaction  $v_3$  flanked by species  $S_2$  and  $S_3$ , respectively. Let us make a perturbation in the upstream source metabolite,  $X_o$ . Changes will propagate through the pathway resulting in changes in  $S_2$  and  $S_3$ . If the changes are sufficiently small, we can write down the change in flux using the following relation:

$$\frac{\delta J^\dagger}{J^\dagger} = \varepsilon_2^3 \frac{\delta s_2^\dagger}{s_2^\dagger} + \varepsilon_3^3 \frac{\delta s_3^\dagger}{s_3^\dagger} \quad (9.3)$$

The quantities have a superscript  $\dagger$  to indicate this is the set of changes that accompanied the changes are a result of a perturbation to  $X_o$ . We can now carry out a separate experiment where we perturb the downstream sink pool  $X_4$ . This time using the superscript  $\ddagger$ . Again, we will observe propagations in the pathway resulting in changes to  $S_2$  and  $S_3$ . Note that because the disturbance is from another source, the changes in  $S_2$ ,  $S_3$  and flux  $J$  will be different. That is:

$$\frac{\delta J^\ddagger}{J^\ddagger} = \varepsilon_2^3 \frac{\delta s_2^\ddagger}{s_2^\ddagger} + \varepsilon_3^3 \frac{\delta s_3^\ddagger}{s_3^\ddagger}$$

We now have two equations in two unknowns,  $\varepsilon_2^3$  and  $\varepsilon_3^3$ . Assuming changes in metabolite concentration and fluxes can be measured, the two equations can be used to solve for the

elasticities. In principle, if all the metabolite changes were measured in the entire pathway, we could estimate all the elasticities. Once we have the elasticities the control coefficients can be estimated using the methods described in the last chapter. Another point worth making, it doesn't matter what changes are made to illicite the perturbation. In this case we used  $X_o$  and  $X_4$  but changes in enzyme levels or addition of inhibitors are equally valid ways to perturb the system. The method has been generalized by Acerenza and Cornish-Bowden [1] and an experimental application was published by Giersch [43].

### **Example 9.2**

---

Using the same system from the previous example, the following perturbation data were obtained by carrying out two perturbations, one upstream and one downstream of the reaction under observation. Estimate the values for the elasticities  $\varepsilon_1^2$  and  $\varepsilon_2^2$ .

Before any perturbations were made, the following reference flux and concentrations were recorded:

$$\begin{aligned} J &= 1.5 \\ s_1 &= 0.74 \\ s_2 &= 0.92 \end{aligned}$$

A perturbation that involved increasing the input pool by 30% was applied and a new steady state flux and concentrations were recorded:

$$\begin{aligned} J &= 1.6 \\ s_1 &= 0.91 \\ s_2 &= 1.15 \end{aligned}$$

Note that it was not possible to perturb the output metabolite from the pathway, instead an inhibitor was applied to one of the downstream steps. The degree of inhibition at the inhibited site is not known (and is irrelevant), but the new steady state flux and concentrations were recorded:

$$\begin{aligned} J &= 1.28 \\ s_1 &= 0.905 \\ s_2 &= 1.32 \end{aligned}$$

From this data, estimate the two elasticities,  $\varepsilon_1^2$  and  $\varepsilon_2^2$  with respect to  $S_1$  and  $S_2$ . Formulate two equations, corresponding to each perturbation, of the form:

$$\frac{\delta J}{J} = \varepsilon_1^2 \frac{\delta s_1}{s_1} + \varepsilon_2^2 \frac{\delta s_2}{s_2}$$

From the data these two equations are:

$$\begin{aligned} \frac{0.1}{1.5} &= \varepsilon_1^2 \frac{0.17}{0.74} + \varepsilon_2^2 \frac{0.23}{0.92} \\ \frac{-0.22}{1.5} &= \varepsilon_1^2 \frac{-0.165}{0.74} + \varepsilon_2^2 \frac{0.4}{0.92} \end{aligned}$$

Evaluating the ratios we obtain:

$$\begin{aligned} 0.066 &= \varepsilon_1^2 0.23 + \varepsilon_2^2 0.25 \\ -0.14 &= -\varepsilon_1^2 0.23 + \varepsilon_2^2 0.43 \end{aligned}$$

Solving for  $\varepsilon_1^2$  and  $\varepsilon_2^2$  yields:

$$\varepsilon_1^2 = 1.531 \quad \varepsilon_2^2 = -1.144$$

The values from the model are  $\varepsilon_1^2 = 1.69$   $\varepsilon_2^2 = -1.3$ . The discrepancy in the values is a result of two factors. The first is that all values were rounded down to two decimal places, secondly and more importantly, we made relatively large perturbations. The double modulation method depends on making small enough changes such that the relationship described by equation (9.3) remains true. If the perturbations are too high, equation (9.3) is only an approximation. Ideally one might make multiple perturbation at different strengths, plotting the resulting changes and extrapolating the plotted response back to the zero axis.

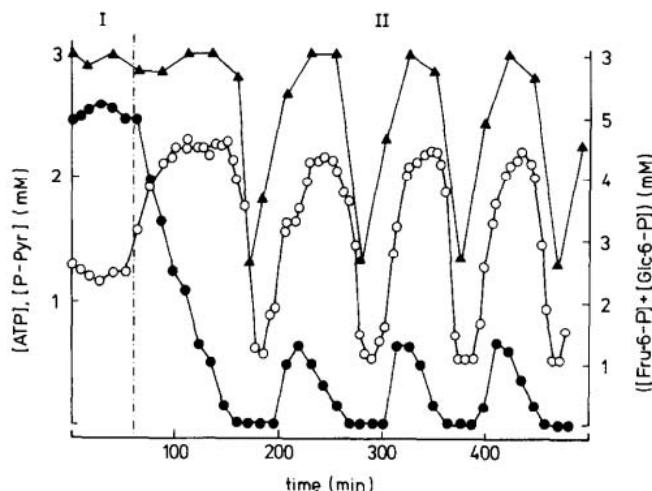
---

## 9.6 Reconstituted Systems

---

There has been a long history of using reconstituted systems as a way to understand cellular function. In the 1940s, actomyosin threads from muscle were reconstituted by Szent-Gyoergyi and later Straub, which led to the understanding of how muscles contract [132]. Perhaps more well known is the work by Arthur Kornberg [72] in the late 1950s who studied DNA replication. Liu and Fletcher [78] provide a more detailed view of reconstituted systems such as mitotic spindles, cell motility and membrane dynamics. The Liu and Fletcher article makes no mention of reconstituting metabolic or signaling pathways however. It is therefore worth mentioning a remarkable set of papers published in the late 1970s and early 1980s by a group of East German scientists. Their work involved the partial reconstitution of glycolysis using purified enzymes: Pyruvate Kinase, Adenylate Kinase, Phosphofructokinase, Fructose 1,6-bisphosphatase and Glucose-6-phosphatase. The five enzymes were studied in a continuously fed stirred tank reactor where enzymes and metabolites were pumped into a 1 ml chamber. Effluent was collected and metabolite levels measured. A number of variants of this setup were also explored including trapping the enzymes in a polyacrylamide gel and varying the enzymes in the mix. In a series of papers, this system (including variants) was used to study bistability, oscillatory and other behaviors [117, 33, 118, 34, 35].

Very little additional work has been done in reconstituted systems until recently. It is worth mentioning the work by Chassagnole et al [21] on reconstituting the threonine pathway, branch-point studies by Curien [26], and Kouril [74] on reconstituting gluconeogenesis in *sulfolobus solfataricus*. More interesting from the perspective of this book, was the work carried out by Torres et al [136], who applied metabolic control analysis to the reconstituted system using rat liver extract containing six enzymes. By titrating the extract with additional enzyme, the authors were able to calculate the flux control coefficients. More recently,



**Figure 9.3** This plot [34] shows sustained oscillations in a reconstituted system containing five enzymes Pyruvate Kinase, Adenylate Kinase, Phosphofructokinase, Fructose 1,6-bisphosphatase and Glucose-6-phosphatase. Oscillations are shown for ATP, Phosphoenolpyruvate and Fructose-6-P+Glucose-6-P.

Panke and his group [16] have put together a remarkable experimental setup for studying a reconstituted system of glycolysis. What was most interesting about this study was the use of online mass-spectrometry to obtain real-time measurements of fourteen metabolites. In addition, they titrated the system with enzymes to determine the flux control coefficients. Given experimental setups such as those developed by Panke, it is possible to obtain highly detailed measurements of the dynamics of metabolic pathways. Of all the experimental approaches, this probably offers the most exciting potential to understand the dynamics of metabolic pathways.

## 9.7 By Computer Simulation

There are many instances in the literature where control coefficients have been computed by simulation. This is achieved by building a kinetic model of the pathway and using the computer to carry out the perturbations and thus estimate the control coefficients. Computer models can be developed to investigate the properties of particular pathway motifs or as a means to build predictive models of actual pathways. Some of the earliest uses of computer simulation to compute control coefficients was work done by Reinhart Heinrich and colleagues in developing a series of models of red blood cell metabolism [101]. In doing so they used the computer simulation to estimate the flux and concentration control coefficients. However one of the impediments to using computer simulation has been the availability of suitable data and more significantly, the lack of validation of the models once

built. In recent years this has begun to change and there now exist a number of well validated and hence reliable computer models of metabolism.

Of particular note is the work by Smallbone, Mendes and coworkers [124] who developed a validated computer model of yeast glycolysis. The authors developed an initial model based on existing literature data. From this model estimates for the control coefficients were calculated using the simulation model. Those enzymatic steps with the largest flux control where isolated and their kinetics remeasured experimentally under physiological conditions. The new kinetic data was used to improve the original simulation model; the method was then repeated until all the enzymes had been experimentally characterized. This resulted in the development of a much more reliable model of glycolysis than was hitherto possible. A similar approach was very successfully applied to developing a reliable kinetic model of glycolysis in the malaria parasite, *Plasmodium falciparum* [91, 97] by Penlker, Snoep and coworkers.

The experience gained by these studies has highlighted the fact that one of the most important aspects when building a kinetic model is to have the kinetic properties of the individual enzyme measured under physiological conditions. This may seem like an obvious thing to do but until recently it was an aspect that was largely ignored. One possible reason for this is that previously, model building relied on global fitting of the model parameters to experimental data. That is, all model parameters were fitted simultaneously. This appears to have resulted in models that fitted the experimental data but failed when an attempt was made to use the model to predict behaviour beyond the scope of the fitted model, indicating that the model had failed to generalize. The approaches used by Mendes and Snoep indicate that it is possible to build reliable computer models of metabolism and to use computer simulation to estimate the control coefficients. Further commentary on this approach to modeling can be found in the article by van Eunen and Bakker [142].

---

## 9.8 By Calculation - Serine Pathway

---

Fell and Shell [38] (see also [126]) considered the serine biosynthesis pathway in rabbit liver. Serine is made by a linear pathway of three steps that branches off glycolysis at 3-phosphoglycerate.



The first step, catalyzed by 3-phosphoglycerate dehydrogenase, requires the reduction of NAD to NADH. The second step is catalyzed by a 3-phosphoserine aminotransferase and requires glutamate as a cofactor (producing 2-oxoglutarate), and the third step catalyzed by 3-phosphoserine phosphatase that results in the release of free phosphate. The question that the authors set out to answer was what determines the flux control coefficients for the three steps? Based on traditional views of metabolic regulation, it would be expected that the committed step (3-phosphoglycerate dehydrogenase) would be the rate-limiting step in

the pathway. The question is whether this can be shown by computing the flux control coefficients.

In order to make the analysis simpler, the authors assumed that the flux through the serine pathway was small compared to the main glycolytic pathway. This means that whatever changes occurred in the serine pathway, had a negligible effect on glycolytic intermediate 3-phosphoglycerate. Likewise, it was assumed that the cofactor concentrations for NAD/NADH and glutamate/2-oxoglutarate would be held relatively constant by other cellular processes. An additional assumption had to be made regarding the first two enzymes. Because the concentration of phosphohydroxypyruvate has been found to be too low to measure, the authors decided to merge the two first steps into a single grouped step. As a result, the final calculation would yield a single flux control coefficient for the group, and another flux control coefficient for the last step.

In order to compute the flux control coefficients, the authors needed to estimate the elasticities. The first elasticity to compute is the elasticity for the first two grouped enzymes with respect to the intermediate 3-phosphoserine. The maximal activities for the first two enzymes are high in rabbit liver which allowed the authors to assume that the degree of saturation of the enzymes by substrate was low. Note that the intermediate 3-phosphohydroxypyruvate is already very low, and is an additional factor that justifies this assumption. If the saturation levels are very low then the elasticities can be computed using equation (6.8) and repeated here:

$$\begin{aligned}\varepsilon_s^v &= \frac{1}{1 - \Gamma/K_{eq}} = \frac{1}{1 - \rho} \\ \varepsilon_p^v &= -\frac{\Gamma/K_{eq}}{1 - \Gamma/K_{eq}} = -\frac{\rho}{1 - \rho}\end{aligned}$$

To use these equations the mass-action ratio,  $\Gamma$ , is required together with the equilibrium constant for the two step group. The concentrations for 3-phosphoglycerate, 3-phosphoserine, NAD/NADH and glutamate/2-oxoglutarate were obtained from the literature and used to compute the mass-action ratio. The joint equilibrium constant was obtained by multiplying the two individual equilibrium constants, again obtained from the literature. From this information; the elasticity for 3-phosphoserine with respect to the group was computed as -1.43.

The next step was to compute the elasticity of 3-phosphoserine phosphatase with respect to 3-phosphoserine (PSer). It was determined by kinetic fitting that 3-phosphoserine phosphatase is uncompetitively inhibited by serine. This resulted in formulating the following rate law for 3-phosphoserine phosphatase:

$$v = \frac{V_m \text{PSer } a}{\text{PSer} + K_m a} \quad \text{where } a = \frac{1 + \text{Ser}/K_1}{1 + \text{Ser}/K_2}$$

where the various kinetic constants were determined from non-linear fitting of literature data -  $K_m = 0.089\text{mM}$ ,  $K_1 = 16.5\mu\text{M}$  and  $K_2 = 0.6\text{ mM}$ . The elasticity computed in this way

was determined to be 0.041. The low value is attributed to the significant inhibition from serine. Combining the summation and connectivity theorems as given by equation (8.3) and (8.4) allows the flux control coefficients to be computed:

$$C_{1,2}^J = 0.03 \quad C_3^J = 0.97 \quad (9.4)$$

What we see is that the last step has most of the control, the first two steps have very little control. This tells us that the committed step is not necessarily the rate determining step. The authors go on to make further observations depending on the feeding state of the animals. For example, under a glucose/ethanol feed regime, the control coefficients are computed to be 0.46 and 0.54, respectively. That is, neither step dominates. What is more interesting is that the first two steps under these conditions are closer to equilibrium and yet have a flux control of 0.46. This shows that it is possible for near-equilibrium steps to acquire significant flux control. The paper includes much more detail on other aspects of the study and the reader should refer to the original paper [38].

## 9.9 Leveraging CRISPR Technology

---

CRISPR [148] (Clustered Regularly Interspaced Short Palindromic Repeats) systems and the associated cas genes are a diverse set of immunity mechanisms found in many bacteria and archaea used to protect themselves from viruses and other foreign nucleic acids. The cas9 protein (CRISPR-associated protein 9) derived from type II CRISPR is a powerful tool for engineering genomes. The protein has the ability to cut DNA at specific sites depending on the presence of a guide RNA. This makes it easy to use cas9 to target any DNA site simply by providing a suitable guide RNA. The nuclease activity of cas9 can be deactivated by mutating the two nuclease domains on the protein. The deactivated cas9 is referred to as dCas9. dCas9 still has the RNA-guided binding properties of the ordinal cas9 except it no longer cuts the DNA. If dCas9 is targeted to bind the promoter regions on bacterial genomes, steric hindrance will repress gene transcription. This mode of action is sometimes called CRISPR interference or CRISPRi. In mammalian systems dCas9 can be fused with repressor (KRAB) or activator domains such as VP64, and thus down or up regulate specific genes of interest. For example, repression by dCas9 of bacterial transcription can be up to a 1000-fold, and by varying the guide RNA, the degree of repression can be modulated.

CRISPR technology offers a tantalizing possibility to measure control coefficients across an entire pathway or even the entire organism. For example, using pooled libraries of guide RNAs, cells can be transfected such that on average, each cell acquires a single guide RNA. At this point cells are separated and cultured. For each culture the dCas9 system can be activated by inducing expression of the guide RNA. Once the cells are at a new steady state, the metabolite concentrations can be measured using standard metabolomic methods. The guide RNAs can be barcoded so that for each cell culture the particular gene that was perturbed can be identified. Metabolite measurements as well as the degree to which RNA expression was changed can then be used to compute the control coefficients.

## Further Reading

---

1. Kacser H and Burns J. (1979) Molecular Democracy: Who Shares the Controls? *Biochem Soc Trans*, 7, 1149-1160.
2. Rafael Moreno-Sanchez, Emma Saavedra, Sara Rodrguez-Enriquez, and Viridiana Olin-Sandoval (2008) Metabolic Control Analysis: A Tool for Designing Strategies to Manipulate Metabolic Pathways, *Journal of Biomedicine and Biotechnology*, Volume 2008, Article ID 597913, doi:10.1155/2008/597913

## Exercises

---

1. Write an essay on the use of physiologically measured enzyme kinetics to build reliable simulation models of metabolism.



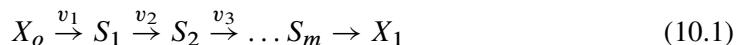
# 10

## *Linear Pathways*

### 10.1 Basic Properties

---

Linear pathways represent the simplest network motif and are a good starting point to begin to gain insight into how cellular networks operate. The simplest linear pathway is one where the kinetics of each step follows simple mass-action. We will also assume throughout this chapter that there are no feedback loops, each step is only affected by its immediate reactant or product. Consider the following linear pathway:



This pathway has  $m$  floating species and  $n$  reactions ( $n = m + 1$ ).  $X_o$  and  $X_1$  are fixed species representing the source and sink pools, respectively. We can assume that each reaction obeys the following simple reversible mass-action kinetic law:

$$v_i = k_i s_{i-1} - k_{-i} s_i \quad (10.2)$$

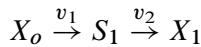
where  $k_i$  and  $k_{-i}$  are the forward and reverse rate constants, respectively.  $S_{i-1}$  is the substrate and  $S_i$  the product. Recall that the equilibrium constant for such a simple reaction is given by:

$$K_{eq} = q = \frac{k_i}{k_{-i}} = \frac{s_i}{s_{i-1}}$$

which means we can replace the reverse rate constant and rewrite the rate law as:

$$v_i = k_i \left( s_{i-1} - \frac{s_i}{q_i} \right) \quad (10.3)$$

This model is simple enough that we can derive the analytical equation for the steady state flux through the pathway. One way to do this is to first start with a two-step pathway:



The rates for the two steps are given by:

$$v_1 = k_1 \left( x_0 - \frac{s_1}{q_1} \right) \quad v_2 = k_2 \left( s_1 - \frac{x_1}{q_2} \right)$$

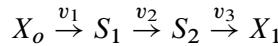
By setting  $v_1 = v_2$ , we can solve for the steady state concentration of  $S_1$ :

$$s_1 = \frac{q_1}{q_2} \frac{k_2 x_1 + k_1 q_2 x_0}{k_1 + k_2 q_1}$$

Inserting this solution into one of the rate laws leads to the steady state flux:

$$J = \frac{\frac{x_0 q_1 q_2 - x_1}{1}}{\frac{1}{k_2} q_1 q_2 + \frac{1}{k_1} q_2} \quad (10.4)$$

We can also derive the flux equation for a three-step pathway



Setting  $v_1 = v_2 = v_3$ , we can solve for steady state concentration of  $S_1$  and  $S_2$ :

$$s_1 = \frac{q_1}{q_2} \frac{k_2 k_3 x_1 + k_1 k_2 x_0 + k_1 k_3 q_2 q_3 x_0}{k_1 k_2 + k_1 k_3 q_2 + k_2 k_3 q_1 q_2}$$

$$s_2 = \frac{q_2}{q_3} \frac{k_1 k_3 x_1 + k_2 k_3 q_1 x_1 + k_1 k_2 q_1 q_3 x_0}{k_1 k_2 + k_1 k_3 q_2 + k_2 k_3 q_1 q_2}$$

Inserting either  $s_1$  or  $s_2$  into one of the reaction rate laws will produce the pathway steady state flux:

$$J = \frac{\frac{x_0 q_1 q_2 q_3 - x_1}{1}}{\frac{1}{k_1} q_1 q_2 q_3 + \frac{1}{k_2} q_2 q_3 + \frac{1}{k_3} q_3}$$

If we compare the solutions to the two step and three step pathway we see a pattern forming from which we can deduce that the flux for a pathway of arbitrary length will be given by:

$$J = \frac{\frac{x_0}{1} \prod_{i=1}^n q_i - x_1}{\sum_{i=1}^n \frac{1}{k_i} \left( \prod_{j=i}^n q_j \right)} \quad (10.5)$$

where  $n$  is the number of steps in the linear chain of reactions. For example, if the pathway has four steps then the steady state flux is given by:

$$J = \frac{x_o q_1 q_2 q_3 q_4 - x_1}{\frac{1}{k_1} q_1 q_2 q_3 q_4 + \frac{1}{k_2} q_2 q_3 q_4 + \frac{1}{k_3} q_3 q_4 + \frac{1}{k_4} q_4}$$

and so on. A pattern in the steady state concentrations equations is also evident but it is more subtle. The Chapter appendix gives the solutions to steady state species concentrations for a four-step pathway where the patterns begins to be more obvious.

The first thing to note about the flux relationship is that the flux is a function of all kinetic and thermodynamic parameters. There is no single parameter that determines the flux completely. This means that for a pathway with randomly assigned parameters, it is extremely unlikely to have the first step as the rate limiting step. It would require a very unlikely set of parameter values for that to occur.

From the flux expression we can compute the corresponding flux control coefficients. For this we need to differentiate the flux equation with respect to an enzyme activity-like parameter. We can use the  $k_i$  parameter as a proxy for the enzyme activity. The result of this yields the following expression for the flux control coefficient of the  $i$ th step:

$$C_i^J = \frac{\frac{1}{k_i} \prod_{j=i}^n q_j}{\sum_{j=1}^n \frac{1}{k_j} \prod_{k=j}^n q_k} \quad (10.6)$$

Note that the sum,  $\sum C_i^J = 1$ , in accordance to the flux summation theorem. The equation also indicates that, at least in this case, the control coefficients are less than one but greater than zero,  $0 \leq C_i^J \leq 1$ .

For a three-step pathway the flux control coefficients for each step will be given by:

$$C_1^J = \frac{1}{k_1} q_1 q_2 q_3 / D$$

$$C_2^J = \frac{1}{k_2} q_2 q_3 / D \quad \text{where } D = \frac{1}{k_1} q_1 q_2 q_3 + \frac{1}{k_2} q_2 q_3 + \frac{1}{k_3} q_3$$

$$C_3^J = \frac{1}{k_3} q_3 / D$$

Note that each term in a numerator can be found in the denominator.

## 10.2 Product Insensitive Steps and Fast Reactions

From the flux control coefficient equation (10.6) we can make some general statements. Let us assume for example that each equilibrium constant,  $q_i$ , is greater than one,  $q_i > 1$ , and that all forward rate constants are equal to each other and all reverse rate constants are equal to each other. This also means that all equilibrium constants are equal. If we now take the ratio of two adjacent steps, for example the  $i^{\text{th}}$  and  $i + 1^{\text{th}}$  step, we find:

$$\frac{C_i^J}{C_{i+1}^J} = \frac{1/k_i \prod_{j=i}^n q_j}{1/k_{i+1} \prod_{j=i+1}^n q_j} = \frac{k_{i+1}}{k_i} q_i$$

Since  $q_i = k_i/k_{-i}$ :

$$\frac{C_i^J}{C_{i+1}^J} = \frac{k_{i+1}}{k_i} \frac{k_i}{k_{-i}} = \frac{k_{i+1}}{k_{-i}}$$

Given that we set all the forward rate constants to equal to each other and all the reverse rate constants equal to each other, the ratio  $k_{i+1}/k_{-i}$  must equal the equilibrium constant,  $q$ , therefore:

$$\frac{C_i^J}{C_{i+1}^J} = q \quad (10.7)$$

That is, the ratio of two adjacent control coefficients is equal to the equilibrium constant. Because we assumed that  $q > 1$ , it must be true that  $C_i^J > C_{i+1}^J$ , that is **earlier steps** will have **more** flux control. This pattern applies across the entire pathway such that steps near the beginning of a pathway will have more control than steps near the end. We will call this effect **front loading** and gives some credence to the traditional idea that the first or committed step is the most important step in a pathway. However, front loading only applies to unregulated pathways; the moment we add regulation to the pathway, this picture changes. We will consider front loading again in a later section.

As an illustration, consider a five step linear pathway and assume the equilibrium constant for each step is,  $q = 2.0$ . This means the ratio of adjacent flux control coefficients will be two. Taking into account the summation theorem we arrive at the following flux control coefficient values across the pathway:

Step	1	2	3	4	5
$C_i^J$	0.52	0.26	0.13	0.06	0.03

Another way to look at a linear pathway is via the mass-action ratio:

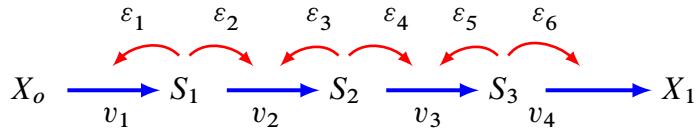
$$\Gamma = \frac{s_i}{s_{i-1}}$$

where the species concentrations are measured at steady state. We define the disequilibrium ratio,  $\rho$ , to be equal to:

$$\rho = \frac{\Gamma}{K_{eq}}$$

If a step is near equilibrium, then  $\rho \simeq 1$ , whereas if a step is far from equilibrium, then  $\rho \ll 1$ .

Consider the following linear pathway where  $X_o$  and  $X_1$  are fixed species:



The elasticities have been labeled 1 to 6, for example  $\varepsilon_1$  represents  $\varepsilon_1^1$ ,  $\varepsilon_2$  represents  $\varepsilon_1^2$  etc. Considering the connectivity theorem for each metabolite, the ratios of all the flux control coefficients can be shown to be:

$$C_1^J : C_2^J : C_3^J : C_4^J = \\ 1 : -\frac{\varepsilon_1}{\varepsilon_2} : -\frac{\varepsilon_1}{\varepsilon_2} \left( -\frac{\varepsilon_3}{\varepsilon_4} \right) : -\frac{\varepsilon_1}{\varepsilon_2} \left( -\frac{\varepsilon_3}{\varepsilon_4} \right) \left( -\frac{\varepsilon_5}{\varepsilon_6} \right)$$

or for a pathway of arbitrary length, the  $n^{\text{th}}$  term will equal:

$$\prod_{i=1}^{n-1} \left( -\frac{\varepsilon_i}{\varepsilon_{i+1}} \right)$$

If we assume that the enzymes are operating below saturation so that they are governed by the rate law,  $v_i = Vm_i / Km_i (S_{i-1} - S_i / K_{eq_i})$ , then we can replace the substrate elasticities by  $1/(1 - \rho_i)$  and the product elasticities by  $-\rho_i/(1 - \rho_i)$  (See (6.8)). If we apply these substitutions, the ratios of flux control coefficients become:

$$C_1^J : C_2^J : C_3^J : C_4^J = \\ (1 - \rho_1) : \rho_1(1 - \rho_2) : \rho_1\rho_2(1 - \rho_3) : \rho_1\rho_2\rho_3(1 - \rho_4) \quad (10.8)$$

or for an arbitrary length pathway, the  $n^{\text{th}}$  term is equal to:

$$\left( \prod_{i=1}^{n-1} \rho_i \right) (1 - \rho_n) \quad (10.9)$$

This is an important result, because by knowing just the equilibrium constants and the concentrations of the intermediate pools, it is possible to obtain an idea of the relative strengths of the flux control coefficients across the pathway.

For a linear pathway without regulation and where each catalyzed reaction is operating below its substrate and product  $K_m$ , the relative distribution of flux control coefficients can be determined from the metabolite concentrations and equilibrium constants.

**Irreversible Steps:** We can draw some interesting conclusions from relation (10.9). Let us make one of the steps irreversible, say step  $i$ , so that the disequilibrium ratio for that step is zero, ( $\rho_i = 0$ ). Since  $\rho_i$  appears as a multiplier in the terms downstream of the irreversible step, all the flux control coefficients for steps beyond will be zero. Thus steps beyond an irreversible reaction have no control over the flux (This also assumes no product inhibition). However, steps upstream of the irreversible step may still have control. Therefore, providing that the irreversible step is not the first step of the pathway, an irreversible step will not necessarily carry a control coefficient of one.

In a linear pathway governed by **linear kinetics** and without the presence of regulatory interactions, all steps downstream of an irreversible step, ( $\rho_i = 0$ ), have no flux control.

Although this result was derived assuming linear kinetics, the result is more general and applies equally to steps governed by non-linear Michaelis-Menten kinetic laws or steps that show cooperativity. The more general result will be shown in a later section.

It is easy to understand why steps beyond an irreversible step have no control. Imagine a perturbation in an enzyme activity at a step downstream of an irreversible step. This perturbation will result in changes in metabolite concentration upstream of the perturbed step. However, a perturbation in the concentration of the product of the irreversible step will, by definition, have no effect on the reaction rate of the irreversible step. This also means that reaction rates of all reaction steps upstream of the irreversible remain unchanged. It is impossible, therefore, for downstream perturbations to change the overall flux through the pathway. In the extreme case where the first step is irreversible, the only step that has any influence on the pathway flux is the first step. All other steps have no influence. This means that the flux control coefficient for the first step will be one, and all downstream steps zero.

**Steps close to Equilibrium:** If any of the steps are near equilibrium, then the disequilibrium ratio for that step will be nearly equal to one (i.e. for step  $i$  close to equilibrium,  $\rho_i \approx 1$ ). Under these conditions, the term  $(1 - \rho_i)$  will equal approximately zero, and therefore the flux control coefficient for that step will also be near zero. In addition, steps other than step  $i$  act as if step  $i$  is not part of the pathway and the pathway appears shortened.

In a linear pathway governed by linear kinetics and without regulation, any step that is very close to equilibrium will have a flux control coefficient close to zero.

It is possible to show that the disequilibrium ratio,  $\rho$ , is equal to the ratio of the reverse and forward rates for a given reaction:

$$\rho = \frac{v_r}{v_f}$$

Since the forward rate will always be greater than the reverse rate for a pathway showing a

positive net rate, the disequilibrium ratio will always be less than one:

$$\rho \leq 1$$

Because  $\rho$  is always less than one, the tendency is for flux control to be higher near the front of the pathway since downstream steps have greater multiples of  $\rho$  values that are less than one (see later section on front loading 10.4).

### Relaxation Times

For a simple decay process, the half-life is given by  $\ln 2 / k_1$  where  $k_1$  is the rate constant for the process. The term  $1/k_1$  is often called the **relaxation time** and gives an idea of how fast the process changes. For a reversible system such as:



where the initial concentration of  $A = A_o$  and for  $B$  is zero, the change in the concentration of  $A$  as a function of time is given by (See (4.4)):

$$a(t) = a_o e^{-(k_1+k_2)t} + \frac{k_2 T}{k_1 + k_2} \left( 1 - e^{-(k_1+k_2)t} \right)$$

If  $k_1 + k_2$  is large then the dynamics will be dominated by the first term in the equation and we can approximate  $a(t)$  by:

$$a = a_o e^{-t(k_1+k_{-1})}$$

where  $k_1$  and  $k_{-1}$  are the forward and reverse rate constants, respectively. The term  $(k_1+k_{-1})$  is analogous to the half-life for the simple decay process and by analogy, the reciprocal of  $(k_1 + k_{-1})$  is called the **relaxation time**, usually denoted by  $\tau$ :

$$\tau = \frac{1}{(k_1 + k_{-1})}$$

Returning to the linear pathway in (10.1), let us assume that all the equilibrium constants are equal to one,  $q_i = 1$ . This means that the forward and reverse rate constants for each reaction are equal. Applying these assumptions to the flux control equation (10.6) we find that:

$$C_i^J = \frac{\frac{1}{k_i}}{\sum_{j=1}^n \frac{1}{k_j}}$$

and noting that since  $k_i = k_{-1}$ , then  $\tau = 1/(k_i + k_{-1}) = 1/(2k_i)$ , we finally obtain:

$$C_i^J = \frac{\tau_i}{\tau_1 + \dots + \tau_n}$$

This relation shows how a given flux control coefficient depends on the relaxation time of the particular step relative to the sum of all the relaxation times. That is, the higher the relaxation time, the larger the flux control. This result relates to the previous section where steps close to equilibrium tend to have small flux control coefficients. Steps close to equilibrium will necessarily have small relaxation times.

Although the results shown in this section and the previous section tell us that steps close to equilibrium will tend to have small flux control coefficients, we must be very careful in this assertion. In all the equations that predict the values for the flux control coefficients, the one common theme is that no step can be considered in isolation. Thus although a step may be close to equilibrium, this observation must be considered in the context of all the others.

Although a step may be close to equilibrium, the step must be considered in the context of all the others in order to determine the absolute flux control coefficient.

### 10.3 Saturable Enzyme Kinetics

The previous examples used linear mass-action kinetics for the individual steps. What happens if we replace linear mass-action kinetics with saturable enzymatic rate laws? In such situations we are unable to generate analytical solutions for the flux, as in equation (10.5), and since we cannot derive a flux expression, we also cannot generate control coefficient equations such as (10.6). Instead, we must use the method described in section (8.5) and derive the control coefficients in terms of the elasticities. One way to derive the control coefficients is to use the summation and connectivity theorems. Section (8.5) derived the equations for a two-step pathway:

$$\begin{aligned} C_{e_1}^J &= \frac{\varepsilon_1^2}{\varepsilon_1^2 - \varepsilon_1^1} & C_{e_2}^J &= -\frac{\varepsilon_1^1}{\varepsilon_1^2 - \varepsilon_1^1} \\ C_{e_1}^s &= \frac{1}{\varepsilon_1^2 - \varepsilon_1^1} & C_{e_2}^s &= -\frac{1}{\varepsilon_1^2 - \varepsilon_1^1} \end{aligned}$$

Using these equations we can look at some simple extreme behaviors. For example, let us assume that the first step is *completely* insensitive to its product, S, then  $\varepsilon_1^1 = 0$ . In this case, the control coefficients reduce to:

$$C_{e_1}^J = 1 \quad C_{e_2}^J = 0$$

That is, all the control (or sensitivity) is on the first step. This situation represents the classic rate-limiting step. The flux through the pathway is completely dependent on the first step. Under these conditions, no other step in the pathway can affect the flux. The effect

State of $v_1$	$C_{e_1}^J$	$C_{e_2}^J$
Product Insensitive	1	0
At equilibrium	0	1

**Table 10.1** Values of flux control coefficients given the state of  $v_1$ .

is however dependent on the complete insensitivity of the first step to its product. Such a situation is likely to be rare in real pathways. In fact, the classic rate limiting step has almost never been observed experimentally. Instead, a range of “limitingness” is observed, with some steps having more “limitingness” (control) than others. To shift control off the first step, the strength of product inhibition must be increased.

What happens if the first step is near equilibrium? In this situation,  $\varepsilon_1^1$  will approach  $-\infty$  (see Figure 6.4) so that the first step hardly has any flux control and all the control is on the second step.

### Control Coefficients for a Three-Step Pathway

What about a three-step pathway:



The flux control coefficient summation theorem is given by:

$$C_{e_1}^J + C_{e_2}^J + C_{e_3}^J = 1$$

Given that we have two species concentrations,  $S_1$  and  $S_2$ , we have two connectivity theorems:

$$C_{e_1}^J \varepsilon_1^1 + C_{e_2}^J \varepsilon_1^2 = 0$$

$$C_{e_2}^J \varepsilon_2^2 + C_{e_3}^J \varepsilon_2^3 = 0$$

These three equations can be combined to give expressions that relate the flux control coefficients in terms of the elasticities, thus:

$$\begin{aligned} C_{e_1}^J &= \frac{\varepsilon_1^2 \varepsilon_2^3}{\varepsilon_1^2 \varepsilon_2^3 - \varepsilon_1^1 \varepsilon_2^3 + \varepsilon_1^1 \varepsilon_2^2} \\ C_{e_2}^J &= -\frac{\varepsilon_1^1 \varepsilon_2^3}{\varepsilon_1^2 \varepsilon_2^3 - \varepsilon_1^1 \varepsilon_2^3 + \varepsilon_1^1 \varepsilon_2^2} \\ C_{e_3}^J &= \frac{\varepsilon_1^1 \varepsilon_2^2}{\varepsilon_1^2 \varepsilon_2^3 - \varepsilon_1^1 \varepsilon_2^3 + \varepsilon_1^1 \varepsilon_2^2} \end{aligned} \quad (10.11)$$

The first thing to note from these equations is that if the first step is product insensitive, that is  $\varepsilon_1^1 = 0$ , then  $C_{e_1}^J = 1$  and  $C_{e_2}^J$  and  $C_{e_3}^J$  are zero (confirm this yourself). As with the two-step example, if any of the steps is close to equilibrium (compared to the other two), its flux control coefficient will be close to zero. For example, if the second step is close to equilibrium, that is,  $\varepsilon_1^2 \rightarrow \infty$  and  $\varepsilon_2^2 \rightarrow -\infty$ , then  $C_{e_2}^J \rightarrow 0$ .

If the first step of the pathway is product insensitive, then flux control resides exclusively on the first step. All other steps have no influence on the flux through the pathway.

Now assume that each enzyme experiences a small amount of product inhibition, for example each product elasticities,  $\varepsilon_1^1$  and  $\varepsilon_2^2$  equal -0.1. Also assume that the substrate levels are roughly at the  $K_m$  for each enzyme. This means that each substrate elasticity will be 0.5, this includes  $\varepsilon_1^2$  and  $\varepsilon_2^3$ . With these values, the flux control coefficients can be estimated, as shown in Table 10.2. Flux control is clearly biased towards the start of the pathway but some control is found in steps downstream of the first reaction. As previously mentioned, flux control that is biased towards the front of the pathway is called front loading, a topic we will discuss in more detail later in the chapter.

Step	Flux Control Coefficient (-0.1)	(-0.2)
$J_1$	0.806	0.64
$J_2$	0.161	0.26
$J_3$	0.032	0.1

**Table 10.2** Distribution of flux control assuming weak (-0.1) and moderately weak (-0.2) product inhibition and substrate levels at the enzyme's  $K_m$ . Note that when the product inhibition is strengthened to -0.2, there is a significant shift in flux control.

What happens if all three steps are close to equilibrium? At first glance it might seem that no step has flux control. We know from the previous results, steps close to equilibrium have little ability to control flux. However, every system must obey the flux summation theorem where all flux control coefficients sum to one. The division of control in a pathway where all steps are close to equilibrium is instead decided by the relative degree of equilibrium between each step.

It is possible for steps close to equilibrium to have significant flux control depending on the context of the reaction.

## Concentration Control Coefficients

To compute the concentration control coefficients we need a different set of theorems. There are two sets of concentration control coefficients, one with respect to  $S_1$ , and another with

respect to  $S_2$ . For example, if we were to consider the control coefficients with respect to  $S_2$ , we would use the following summation theorem:

$$C_{e_1}^{s_2} + C_{e_2}^{s_2} + C_{e_3}^{s_2} = 0$$

and the two connectivity theorems:

$$C_{e_2}^{s_2} \varepsilon_2^2 + C_{e_3}^{s_2} \varepsilon_2^3 = -1$$

$$C_{e_1}^{s_2} \varepsilon_1^1 + C_{e_2}^{s_2} \varepsilon_1^3 = 0$$

Solving for  $C_{e_1}^{s_2}$ ,  $C_{e_2}^{s_2}$  and  $C_{e_3}^{s_2}$  yields:

$$C_{e_1}^{s_2} = \frac{\varepsilon_1^2}{\varepsilon_1^2 \varepsilon_2^3 - \varepsilon_1^1 \varepsilon_2^3 + \varepsilon_1^1 \varepsilon_2^2}$$

$$C_{e_2}^{s_2} = \frac{-\varepsilon_1^1}{\varepsilon_1^2 \varepsilon_2^3 - \varepsilon_1^1 \varepsilon_2^3 + \varepsilon_1^1 \varepsilon_2^2}$$

$$C_{e_3}^{s_2} = \frac{\varepsilon_1^1 - \varepsilon_1^2}{\varepsilon_1^2 \varepsilon_2^3 - \varepsilon_1^1 \varepsilon_2^3 + \varepsilon_1^1 \varepsilon_2^2}$$

Note that the denominator is positive and the numerators for  $C_{e_1}^{s_2}$  and  $C_{e_2}^{s_2}$  are positive, indicating that increases in  $e_1$  or  $e_2$  result in increases in  $s_2$ . In contrast, the numerator for  $C_{e_3}^{s_2}$  is net negative indicating that increases in  $e_3$  result in decreases in  $s_2$ . We can apply similar reasoning to derive the concentration control coefficients with respect to  $s_1$ :

$$C_{e_1}^{s_1} = \frac{\varepsilon_2^3 - \varepsilon_2^2}{\varepsilon_1^2 \varepsilon_2^3 - \varepsilon_1^1 \varepsilon_2^3 + \varepsilon_1^1 \varepsilon_2^2}$$

$$C_{e_2}^{s_1} = \frac{-\varepsilon_2^3}{\varepsilon_1^2 \varepsilon_2^3 - \varepsilon_1^1 \varepsilon_2^3 + \varepsilon_1^1 \varepsilon_2^2}$$

$$C_{e_3}^{s_1} = \frac{\varepsilon_2^2}{\varepsilon_1^2 \varepsilon_2^3 - \varepsilon_1^1 \varepsilon_2^3 + \varepsilon_1^1 \varepsilon_2^2}$$

The concentration control coefficients can be estimated using some realistic values for the elasticities. Assume that each enzyme experiences a small amount of product inhibition, such that each product elasticity,  $\varepsilon_1^1$  and  $\varepsilon_2^2$  are equal to -0.1. Let us also assume that the substrate levels are roughly at the  $K_m$  for each enzyme. This means that each substrate elasticity will be 0.5, this includes  $\varepsilon_1^2$  and  $\varepsilon_2^3$ . Table 10.3 shows the results of the calculations.

Note how increases in enzymes downstream of a metabolite result in the metabolite decreasing in concentration (negative coefficient), while changes in enzymes upstream of a metabolite result in increases in the metabolite.

Step	$C_i^{s_1}$	$C_i^{s_2}$
$e_1$	1.982	1.802
$e_2$	-1.802	0.180
$e_3$	-0.180	-1.982

**Table 10.3** Distribution of concentration control assuming weak product inhibition and substrate levels at the enzyme's  $K_m$ .

Consider a linear pathway where increases in substrate increase the reaction rate and increases in product decrease the reaction rate. An increase in activity of a given enzyme will result in all species downstream of the enzyme to increase and all species upstream of the enzyme to decrease.

Another important observation is that reactions close to equilibrium have little influence over the species concentrations. Consider the middle reaction,  $v_2$ . If  $v_2$  is close to equilibrium, then its substrate elasticity,  $\varepsilon_1^2 \gg 0$ , and the product elasticity,  $\varepsilon_2^2 \ll 0$ . Under these conditions  $C_{e_2}^{s_1}$  and  $C_{e_2}^{s_2}$  tend to zero because the terms  $\varepsilon_1^2$  or  $\varepsilon_2^2$  only appears in the denominator.

Reaction steps which are close to equilibrium have little influence over species concentrations in a linear pathway.

Heinrich and Shuster in their book The Regulation of Cellular System [53] also showed it is possible to express the concentration control coefficients in terms of the flux control coefficients in a linear pathway. They showed by application of the connectivity and summation theorems that:

For steps at or before  $i$ , that is  $1 \leq j \leq i$ :

$$C_j^{s_i} = \frac{C_j^J}{C_{i+1}^J \varepsilon_i^{i+1}} \sum_{k=i+1}^{n+1} C_k^J$$

For steps downstream of  $i$ , that is  $i + 1 \leq j \leq n + 1$ :

$$C_j^{s_i} = \frac{C_j^J}{C_i^J \varepsilon_i^i} \sum_{k=1}^i C_k^J$$

What both equations tell us is that the value for a concentration control coefficient at step  $j$  is proportional to the flux control coefficient at step  $j$ . Therefore if flux control at a particular step is small, then the ability of the same step to control concentration is also diminished. Given that the denominator contains elasticity terms, a low flux control coefficient isn't however a sufficient criterion for low concentration control.

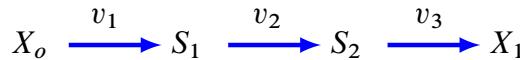
## 10.4 Front Loading

Consider a linear pathway with linear reversible kinetics on each step where the  $q$  is the equilibrium constant  $q > 1$ . Then given two adjacent flux control coefficients, the upstream coefficient will always be equal or larger than the downstream coefficient, that is for the  $i^{\text{th}}$  step the following is true (see equation (10.7)):

$$C_i^J \geq C_{i+1}^J$$

This means that in a linear pathway, control will tend to be concentrated upstream. To understand why this is the case, consider the elasticities and control equations for a linear pathway.

Using the flux summation and connectivity theorems it is straightforward to derive the flux control equations. For example, for the three-step pathway:



one can derive the following flux control coefficient equations – see (10.11):

$$C_{e1}^J = \varepsilon_1^2 \varepsilon_2^3 / D$$

$$C_{e2}^J = -\varepsilon_1^1 \varepsilon_2^3 / D$$

$$C_{e3}^J = \varepsilon_1^1 \varepsilon_2^2 / D$$

where  $D$ , the denominator, is given by:

$$D = \varepsilon_1^2 \varepsilon_2^3 - \varepsilon_1^1 \varepsilon_2^3 + \varepsilon_1^1 \varepsilon_2^2$$

It is possible to do this for pathways with additional steps, where a clear pattern emerges. For a pathway with  $n$  steps where  $n$  is even, we have the following equations:

$$C_1^J = \varepsilon_1^2 \varepsilon_2^3 \varepsilon_3^4 \varepsilon_4^5 \cdots \varepsilon_n^{n+1} / D$$

$\vdots$

$$C_m^J = \prod_{k=m}^n \varepsilon_k^{k+1} \prod_{k=m-1}^1 \varepsilon_k^k / D$$

$\vdots$

$$C_n^J = \varepsilon_1^1 \varepsilon_2^2 \varepsilon_3^3 \varepsilon_4^4 \cdots \varepsilon_{n+1}^{n+1} / D$$

A careful examination of  $C_1^J$  reveals that the numerator is the product of all the substrate elasticities. This implies that a perturbation in  $e_1$  ‘hops’ from one enzyme to the next until it reaches the end of the pathway. Conversely, the control coefficient of the last enzyme includes all the product elasticities, that is the perturbation ‘hops’ from one enzyme to the next until it reaches the beginning of the pathway.

Looking at any intermediate enzyme step we find two groups of elasticities, one group representing the perturbation traveling downstream via the substrate elasticities, and the other representing the perturbation traveling upstream via product elasticities. This implies that the pattern of elasticities in the numerator reflects the path the perturbation takes as it ripples out from the source of the perturbation.

Recall that given a reversible mass-action rate law such as  $k_1 s - k_2 p$ , the elasticities are given by:

$$\varepsilon_s^v = \frac{1}{1 - \rho}$$

$$\varepsilon_p^v = -\frac{\rho}{1 - \rho}$$

From these equations it follows that  $\varepsilon_s^v + \varepsilon_p^v = 1$ , that is:

$$\| \varepsilon_s^v \| \geq \| \varepsilon_p^v \|$$

That is, the absolute value of the substrate elasticity is always greater than the product elasticity. Given that an upstream enzyme will have more substrate elasticities than product elasticities in the numerator of its control equation, it follows that the numerator will be larger when compared to an enzyme further downstream, which will have more of the smaller value product elasticities. This means that perturbations at a downstream enzyme will be attenuated compared to a similar perturbation at an upstream step. Hence the control coefficients upstream will, on average be, larger.

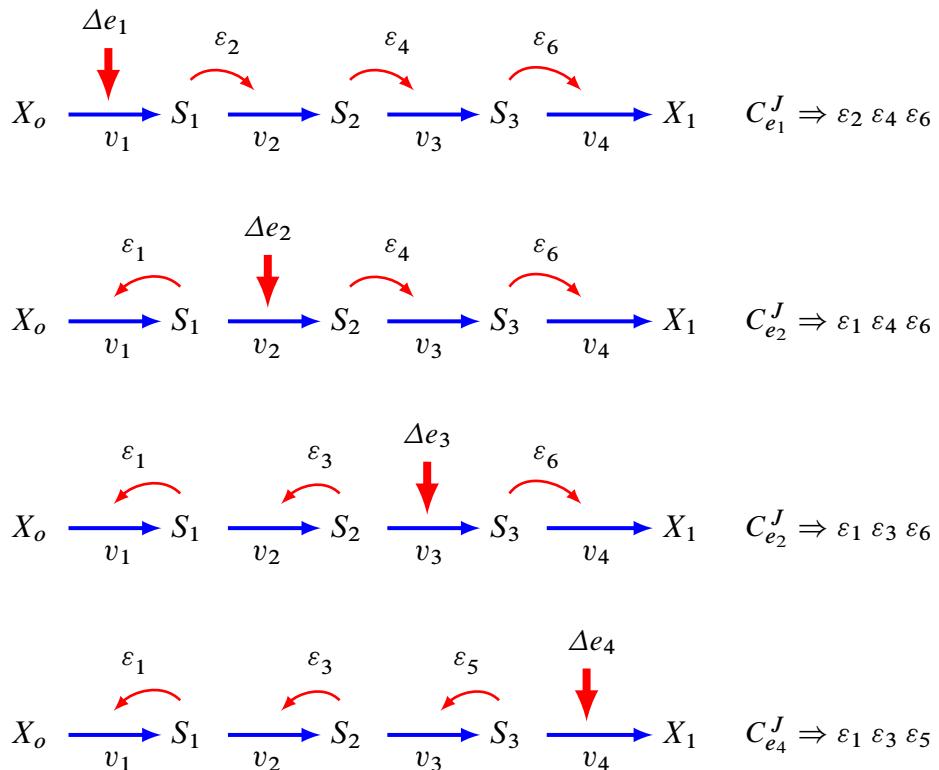
The origins of the asymmetry between the substrate and product elasticities is a thermodynamic one. If the thermodynamic gradient were to be reversed so that the pathway flux traveled ‘upstream’, the elasticity values exchange so that now the front loading occurs downstream, although ‘downstream’ is now ‘upstream’ because the flux has reversed.

The set of perturbations and the signal flow is shown in Figure 10.1.

In a linear pathway governed by linear kinetics and without regulation, flux control is biased towards the start of the pathway, an effect called **front loading**.

## Distribution of Concentration Control

In the previous section it was shown that in an unregulated pathway, flux control tends, on average, to be concentrated on the upstream steps. What about the control of concentra-



**Figure 10.1** Set of four perturbation patterns. In each perturbation, the signal travels outwards up and downstream from the point of action. In each case the elasticities in the numerator of the control equation indicate the pathway of the signal transmission.

tions? The equations that describe the concentration control coefficients are more complex. Consider a four-step linear pathway. A pattern in the concentration control coefficients can be discerned if we look at the three  $C_{e_1}^{s_1}$ ,  $C_{e_1}^{s_2}$ , and  $C_{e_1}^{s_3}$ . For reference, these are given below:

$$\begin{aligned} C_{e_1}^{s_1} &= \frac{-\varepsilon_2^2 \varepsilon_3^3 + \varepsilon_2^2 \varepsilon_3^4 - \varepsilon_2^2 \varepsilon_3^4}{D} \\ C_{e_1}^{s_2} &= \frac{\varepsilon_1^2 \varepsilon_3^3 - \varepsilon_1^2 \varepsilon_3^4}{D} \\ C_{e_1}^{s_3} &= -\frac{\varepsilon_1^2 \varepsilon_2^3}{D} \end{aligned}$$

$D$  is the common denominator. The number of terms in the numerator reduces by one for each species  $s_i$ , as we move downstream. Given that the denominator doesn't change, fewer terms in the numerator means smaller control coefficients. In other words, the further away a species is from a perturbation, the smaller the effect. This is intuitively reasonable because a signal, unless subject to other mechanisms, will attenuate as it propagates from the source of the disturbance. On average, for a linear unregulated pathway, we can state that:

$$C_{e_1}^{s_1} > C_{e_1}^{s_2} > C_{e_1}^{s_3} \dots > C_{e_1}^{s_n}$$

The pattern also applies to those perturbations at the center of a pathway. Intermediates near to the disturbance will be affected more than distant ones.

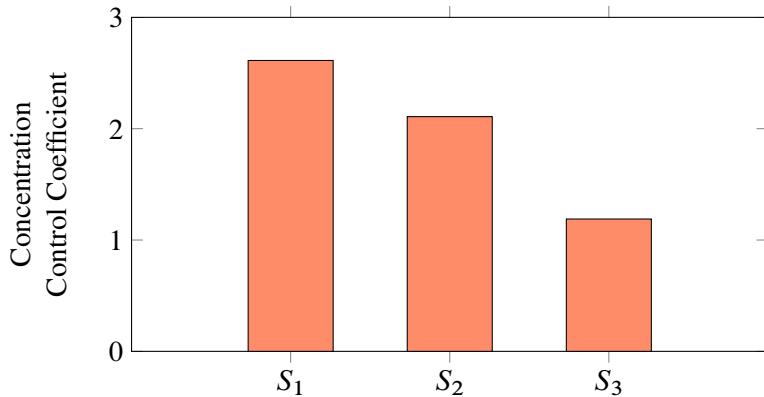
This result also affects the response coefficient,  $R_{x_o}^{s_i}$ . Because the response coefficient is the product of the  $x_o$  elasticity on  $v_1$ , and the corresponding concentration control coefficient,  $C_{e_1}^{s_i}$ , we see that the response coefficient of a species  $s_i$  with respect to  $x_o$  will progressively decrease as we observe species further and further away from  $x_o$ .

This effect can easily be illustrated via simulation. The elasticity values can be given random values and the corresponding control coefficients computed for a four-step pathway. This can be repeated 1000 times, each time with a different sample for the elasticities. The product elasticities can be uniformly sampled between 0 and -1, and the substrate elasticities sampled between 0 and 1. The result of the simulation is 1000 concentration control coefficients for  $C_{e_1}^{s_1}$ ,  $C_{e_1}^{s_2}$  and  $C_{e_1}^{s_3}$ . Taking the mean for each group of control coefficients the following mean values are obtained:

$$C_{e_1}^{s_1} = 2.6128 \quad C_{e_1}^{s_2} = 2.1083 \quad C_{e_1}^{s_3} = 1.1886$$

The results show the influence of  $e_1$  on a species diminishing as the species is further away from  $e_1$ .

In a linear pathway without regulation, concentration control diminishes the further away the species is from the disturbance.



**Figure 10.2** Plot showing the sensitivity of changes to downstream species as a function of a perturbation at the first step.

There is one caveat to the above result when considering a linear pathway with non-saturable kinetics. For example, where each reaction is governed by the rate law:

$$v_i = k_i \left( s_{i-1} - \frac{s_i}{q_i} \right)$$

In this situation, the elasticities for  $s_{i-1}$  and  $s_i$  are not free to take on any value, see (6.9) but are constrained by the relation:

$$\varepsilon_{i-1}^v + \varepsilon_i^v = 1$$

Consider the three-step pathway with concentration control coefficients:

$$C_{e_1}^{s_1} = \frac{\varepsilon_2^3 - \varepsilon_2^2}{\varepsilon_1^2 \varepsilon_2^3 - \varepsilon_1^1 \varepsilon_2^3 + \varepsilon_1^1 \varepsilon_2^2}$$

$$C_{e_1}^{s_2} = \frac{\varepsilon_1^2}{\varepsilon_1^2 \varepsilon_2^3 - \varepsilon_1^1 \varepsilon_2^3 + \varepsilon_1^1 \varepsilon_2^2}$$

Assume that the last step in the pathway is irreversible such that  $\varepsilon_3^3 = 0$ . Given the elasticity constraint, this would mean that  $\varepsilon_2^3 = 1$ . Finally, note that for the middle reaction,  $\varepsilon_1^2 = 1 - \varepsilon_2^2$ . Inserting these values for the elasticities into  $C_{e_1}^{s_1}$  and  $C_{e_1}^{s_2}$  we find they are reduced to:

$$C_{e_1}^{s_1} = \frac{1}{1 - \varepsilon_1^1}$$

$$C_{e_1}^{s_2} = \frac{1}{1 - \varepsilon_1^1}$$

Note the two control coefficients are identical. This generalizes to any length linear pathway that uses linear kinetics, where all  $C_{e_1}^{s_j}$  will be equal to each other. That is the sensitivity for every species with respect to  $e_1$  is the same. In other words no matter how far the species is from the disturbance, the species will respond in the same way as the species closest to the disturbance. This counter intuitive behavior is because the transmission of a signal through a set of linear kinetic laws is not attenuated.

## 10.5 Optimal Allocation of Protein

---

Protein synthesis constitutes a significant drain on resources in a cell [11, 70]. For example, protein synthesis consumes approximately 7.5 ATP equivalents per peptide bond. One glucose molecule yields roughly 36 molecules of ATP. Thus if the average number of peptide bonds in a protein is 300, it takes roughly 62 molecules of glucose to make just one protein molecule, not including the cost of the amino acids. In some cultured mammalian cells, protein synthesis consumes 35% to 50% of all ATP production. In addition to the energetic cost, proteins also occupy a significant proportion of cell volume at around 20 to 30% of the cell. This high level approaches the solubility limit of proteins and also limits the diffusion of other small molecules. These and other issues effectively put an upper limit on the total amount of protein in a cell. It would seem logical to assume that the distribution of a fixed amount of protein is not evenly distributed because some processes may require higher levels of protein compared to others, suggesting competition for protein between different processes. Such distributions are likely to be under evolutionary selection so that there exists an optimal allocation of the fixed amount of protein to all process in the cell. The optimal allocation is also likely to shift as environmental conditions change.

In this section we will consider what the optimal allocation of a fixed amount of protein in a metabolic pathway such that the steady state pathway flux is maximized.

Consider a very simple two-step metabolic scheme shown below:



Assume that the first step is catalyzed by an enzyme  $E_1$ , and the second step by enzyme,  $E_2$ . Let us reduce the amount of enzyme  $E_1$  by a small amount,  $\delta e_1$ , such that the pathway flux is reduced by an amount,  $\delta J$ . We can now increase the level of  $E_2$  by  $\delta e_2$  so that the pathway flux is returned to its original state. The net change in protein is therefore  $\delta e_1 + \delta e_2$ .

Let us also assume that the levels of  $E_1$  and  $E_2$  had previously been adjusted so that for a given flux, the total  $e_1 + e_2$  was at a minimum, that is the distribution of protein was optimal. In other words, it would not be possible to reduce the total amount of protein and at the same time adjust the protein distribution such that the flux is unchanged. Given this, it must be true that:

$$\delta e_1 + \delta e_2 = 0$$

With these changes in  $E_i$  and the fact that the flux does not change, we can write the

following:

$$C_{e_1}^J \frac{\delta e_1}{e_1} + C_{e_2}^J \frac{\delta e_2}{e_2} = \frac{\delta J}{J} = 0$$

Substituting  $\delta e_1 + \delta e_2 = 0$  into the above relation yields:

$$C_{e_1}^J \frac{1}{e_1} = C_{e_2}^J \frac{1}{e_2}$$

We can now invoke the flux summation theorem to eliminate one of the control coefficients to yield:

$$C_{e_1}^J \frac{1}{e_1} = \left(1 - C_{e_1}^J\right) \frac{1}{e_2}$$

Rearranging this to solve for  $C_{e_1}^J$  yields:

$$C_{e_1}^J = \frac{e_1}{e_1 + e_2}$$

This result can be generalized to any length pathway so that for a given total amount of protein and a given flux, the optimal allocation of protein at a particular step,  $i$ , is given by:

$$C_{e_i}^J = \frac{e_i}{\sum e_i}$$

## Further Reading

1. Heinrich R and Rapoport TA (1974) A linear steady-state treatment of enzymatic chains. General properties, control and effector strength. *Eur J Biochem.* 1974 Feb 15;42(1):89-95.
2. Schuster S, Heinrich R (1996) *The Regulation of Cellular Systems*, Springer, ISBN 978-1-4613-1161-4 (Unfortunately this book is out of print and second-hand editions can cost as much as 300 to 400 dollars).
3. Brown, GC (1991) Total cell protein concentration as an evolutionary constraint on the metabolic control distribution in cells. *J Theor Biol.*, 153(2), 195-203
4. Klipp E and Heinrich R (1999) Competition for enzymes in metabolic pathways:: Implications for optimal distributions of enzyme concentrations and for the distribution of flux control. *Biosystems*, 54, 1-14

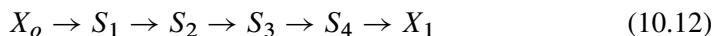
## Exercises

---

1. Show that summing all the  $C_i^J$  coefficients in equation (10.6) equals one.
2. Prove equation (10.8) in the main text.
3. In general, if a given enzymatic step is very close to equilibrium, what can be said about the flux control coefficient of that step?
4. Given a six-step linear pathway where the equilibrium constant for each step is 1.5 and the forward rate constants are equal to each other and reverse rate constants are equal to each other, compute the values for the flux control coefficients for each step. Hint: See (10.7).
5. In a four-step linear pathway each step is catalyzed by a reversible Michaelis-Menten rate law. In addition, each step is close to equilibrium. Does this mean that no step in the pathway can control flux? Explain your answer.
6. An unregulated linear pathway is made up of eight enzymatic reaction steps, all steps are product sensitive except for the fifth step. What can you say about the distribution of flux control in this pathway?
7. Show that the ratio of flux control coefficients in a linear pathway, such as in (10.1), where each reaction is governed by equation (10.3), is given by:

$$C_1^J : C_2^J : C_3^J : \dots = \\ (x_o - s_1/q_1) : (s_1 - s_2/q_2)/q_1 : (s_2 - s_3/q_3)/(q_1 q_2) : \dots$$

8. What is front loading?
9. Metabolic engineers wish to increase the production of an important commodity that is synthesized by a five-step metabolic pathway (10.12).



The pathway has no known negative feedback loops. In order to obtain a rough idea of the distribution of control in this pathway, the engineers obtain values for all the standard  $\Delta G^\circ$ 's and estimates for the steady state concentrations of all the metabolite pools in the pathway.  $\Delta G^\circ$ 's were obtained at 25°C. The table below shows the data they collected:

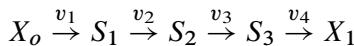
Step	$\Delta G^o$	Metabolite	Concentration
1	-12 kJ mol <sup>-1</sup>	$X_o$	0.9 mM
2	-2 kJ mol <sup>-1</sup>	$S_1$	0.2 mM
3	+1 kJ mol <sup>-1</sup>	$S_2$	0.05 mM
4	-5 kJ mol <sup>-1</sup>	$S_3$	0.45 mM
5	-4 kJ mol <sup>-1</sup>	$S_4$	0.15 mM
		$X_1$	0.01 mM

From the data they collected, what advice would you give concerning which step(s) are worth increasing in activity in order to increase the flux through the pathway?

10. Using a four-step linear pathway where each reaction uses reversible mass-action kinetics (Equation (10.3)), generate 10,000 variations of this pathway. Do this by setting the equilibrium constants to fixed values of  $q_1 = 2; q_2 = 4; q_3 = 8; q_4 = 16$  and then randomizing the forward rate constant between 0 and 1.0. For each pathway variant, compute the flux control coefficients. This can be done by modulating the rate constant for each step and observing the effect on the pathway flux, or by inserting the relevant values into equation (10.6). From the 10,000 variants, compute the distribution of flux control coefficients in the pathway. Explain the distribution of control coefficients you observe.
11. Derive the concentration control coefficient equations for a three-step pathway.
12. It is known that in a given linear pathway the distribution of protein across the enzymes is optimized for flux. In this situation, what is the easiest way to estimate all the flux control coefficients?
13. Prove equation (10.4).

## 10.A Appendix

Steady state concentrations for the four step pathway:



$$s_1 = \frac{q_1}{q_4} \frac{k_1 k_2 k_3 q_4 x_o + k_1 k_2 k_4 q_3 q_4 x_o + k_1 k_3 k_4 q_2 q_3 q_4 x_o + k_2 k_3 k_4 x_o}{k_1 k_2 k_3 + k_1 k_2 k_4 q_3 + k_1 k_3 k_4 q_2 q_3 + k_2 k_3 k_4 q_1 q_2 q_3}$$

$$s_2 = \frac{q_2}{q_4} \frac{k_1 k_2 k_3 q_1 q_4 x_o + k_1 k_2 k_4 q_1 q_3 q_4 x_o + k_1 k_3 k_4 x_1 + k_2 k_3 k_4 q_1 x_o}{k_1 k_2 k_3 + k_1 k_2 k_4 q_3 + k_1 k_3 k_4 q_2 q_3 + k_2 k_3 k_4 q_1 q_2 q_3}$$

$$s_3 = \frac{q_3}{q_4} \frac{k_1 k_2 k_3 q_1 q_2 q_4 x_o + k_1 k_2 k_4 x_1 + k_1 k_3 k_4 q_2 x_1 + k_2 k_3 k_4 q_1 q_2 x_1}{k_1 k_2 k_3 + k_1 k_2 k_4 q_3 + k_1 k_3 k_4 q_2 q_3 + k_2 k_3 k_4 q_1 q_2 q_3}$$



# 11

## ***Branched and Cyclic Systems***

In this chapter we will review branched and cyclic systems. Moiety conserved cycles will be treated as a separate chapter.

### **11.1 Branched Pathways**

---

Branching structures are one of the most common patterns in biochemical networks. Even a pathway such as glycolysis, often depicted as a straight chain in textbooks, is in fact a highly branched pathway.

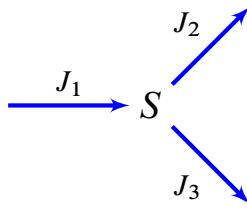
At any given branch node, where a node is a molecular species, there will be conservation of mass. Given a node species,  $s_i$ , with  $b$  branches entering the node and  $d$  branches leaving, the net rate of change in concentration of  $s_i$  is:

$$\sum_{i=1}^b v_i - \sum_{j=1}^d v_j = \frac{ds_i}{dt}$$

At steady state when  $ds_i/dt = 0$ , it must therefore be true that:

$$\sum_{i=1}^b v_i = \sum_{j=1}^d v_j$$

In this section we will investigate the control of flux through a branched system in response to changes in enzyme activity. Let us consider the simple branched pathway depicted in Figure 11.1.



**Figure 11.1** Simple branched pathway. This pathway has three different fluxes,  $J_1$ ,  $J_2$ , and  $J_3$ , which at steady state are constrained by  $J_1 = J_2 + J_3$ .

$C_{e_1}^{J_1}$	$C_{e_1}^{J_2}$	$C_{e_1}^{J_3}$	$C_{e_1}^s$
$C_{e_2}^{J_1}$	$C_{e_2}^{J_2}$	$C_{e_2}^{J_3}$	$C_{e_2}^s$
$C_{e_3}^{J_1}$	$C_{e_3}^{J_2}$	$C_{e_3}^{J_3}$	$C_{e_3}^s$

**Table 11.1** Set of control coefficients for a simple branch pathway.

In Figure 11.1,  $J_1$ ,  $J_2$  and  $J_3$  are the steady state fluxes. By the law of conservation of mass, at steady state, the fluxes in each limb will be governed by the relationship:

$$J_1 = J_2 + J_3$$

Given three different fluxes and one intermediate, there will be four sets of control coefficients, one set concerned with changes in the intermediate,  $S$ , and three sets corresponding to each of the three fluxes (Table 11.1).

For the branched system we can write a summation and a connectivity theorem with respect to each flux. For example, with respect to  $J_1$  we can write:

$$C_{e_1}^{J_1} + C_{e_2}^{J_1} + C_{e_3}^{J_1} = 1$$

and:

$$C_{e_1}^{J_1} \varepsilon_s^{v_1} + C_{e_2}^{J_1} \varepsilon_s^{v_2} + C_{e_3}^{J_1} \varepsilon_s^{v_3} = 0$$

This gives us two equations but three unknown flux control coefficients. To solve for the  $C_{e_i}^{J_1}$ , we need another equation.

Let the fraction of flux through  $J_2$  be given by  $\alpha = J_2/J_1$ , and the fraction of flux through  $J_3$  be given by  $1 - \alpha = J_3/J_1$ . Let us carry out the following thought experiment:

1. Increase the concentration of  $E_2$  by  $\delta e_2$ . This will cause a decrease in  $S$ , an increase in  $J_1$  (relief of product inhibition) and a decrease in  $J_3$ .
2. Restore the change in  $J_1$  by decreasing  $E_3$  such that  $S$  is restored to its pre-perturbation state. At the end of the thought experiment,  $\delta s = 0$ .

3. Since we have not changed  $E$  and  $\delta s = 0$ , it must be the case that  $\delta J_1 = 0$ .

From this experiment we can write down the system and local equations. The system equation is given by:

$$C_{e_2}^{J_1} \frac{\delta e_2}{e_2} + C_{e_3}^{J_1} \frac{\delta e_3}{e_3} = \frac{\delta J_1}{J_1} = 0$$

Note that the system equation only has two terms because we did not change  $E_1$ . The local equations are quite simple because  $\delta s = 0$  and as before we assume that  $\varepsilon_{e_i}^v = 1$ , therefore:

$$\frac{\delta v_2}{v_2} = \frac{\delta e_2}{e_2} \quad \text{and} \quad \frac{\delta v_3}{v_3} = \frac{\delta e_3}{e_3}$$

By substitution, the system equation can be written as:

$$C_{e_2}^{J_1} \frac{\delta v_2}{v_2} + C_{e_3}^{J_1} \frac{\delta v_3}{v_3} = 0$$

Since  $\delta J_1 = 0$ , it must be the case that the net change in flux downstream of  $S$  must also be zero. That is,  $\delta v_2 + \delta v_3 = 0$ , or  $\delta v_2 = -\delta v_3$ . We can therefore eliminate the  $\delta v_3$  term:

$$C_{e_2}^{J_1} \frac{\delta v_2}{v_2} - C_{e_3}^{J_1} \frac{\delta v_2}{v_2} \frac{v_2}{v_3} = 0$$

Cancelling terms we obtain:

$$C_{e_2}^{J_1} - C_{e_3}^{J_1} \frac{v_2}{v_3} = 1$$

We can substitute the absolute rates,  $v_2$  and  $v_3$  with the fractional fluxes,  $\alpha$  and  $1 - \alpha$  to give:

$$C_{e_2}^{J_1} - C_{e_3}^{J_1} \frac{\alpha}{1 - \alpha} = 0$$

and finally:

$$C_{e_2}^{J_1} (1 - \alpha) - C_{e_3}^{J_1} \alpha = 0$$

This result is called the **flux branch point theorem**. We can derive similar theorems with respect to  $J_2$  and  $J_3$ . In each case we carry out the same thought experiment such that the reference flux,  $J_2$  or  $J_3$ , is unchanged. The two additional theorems are given below with respect to  $J_2$  and  $J_3$ .

$$C_{e_1}^{J_2} (1 - \alpha) + C_{e_3}^{J_2} = 0$$

$$C_{e_1}^{J_3} \alpha + C_{e_2}^{J_3} = 0$$

We can also derive, using the same thought experiment, branch point theorems with respect to the species concentration,  $S$ . This time the system equation is:

$$C_{e_2}^s \frac{\delta e_2}{e_2} + C_{e_3}^s \frac{\delta e_3}{e_3} = \frac{\delta s}{s} = 0$$

Substituting in the same local equations as before and noting that  $\delta v_2 = -\delta v_3$ , we obtain after some rearrangement:

$$C_{e_2}^s(1-\alpha) + C_{e_3}^s\alpha = 0$$

This result is known as the **concentration branch point theorem**. It is very similar to the flux branch point theorem. There are also a set of variants that correspond to the flux branch theorems for  $J_2$  and  $J_3$ :

$$C_{e_1}^s(1-\alpha) + C_{e_3}^s = 0$$

$$C_{e_1}^s\alpha + C_{e_2}^s = 0$$

We can write out the theorems in matrix form (See equation (8.17)) using the theorems expressed in terms of  $J_2$ ; this includes one summation, one connectivity and one branch theorem:

$$\begin{bmatrix} C_1^{J_2} & C_2^{J_2} & C_3^{J_2} \\ C_1^s & C_2^s & C_3^s \end{bmatrix} \begin{bmatrix} 1 & -\varepsilon_1^1 & 0 \\ 1 & -\varepsilon_1^2 & 1-\alpha \\ 1 & -\varepsilon_1^3 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

Now solve for the control coefficient matrix by rearranging:

$$\begin{bmatrix} C_1^{J_2} & C_2^{J_2} & C_3^{J_2} \\ C_1^s & C_2^s & C_3^s \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & -\varepsilon_1^1 & 0 \\ 1 & -\varepsilon_1^2 & 1-\alpha \\ 1 & -\varepsilon_1^3 & 1 \end{bmatrix}^{-1}$$

Inverting the second matrix derives  $C_{e_2}^{J_2}$  and  $C_{e_3}^{J_2}$  [39]. In the following, we have simplified the notation by setting  $\varepsilon_1 = \varepsilon_s^1$ ,  $\varepsilon_2 = \varepsilon_s^2$ , and  $\varepsilon_3 = \varepsilon_s^3$ . The denominator,  $\varepsilon_2\alpha + \varepsilon_3(1-\alpha) - \varepsilon_1$  is **positive**, therefore the following equalities hold given that  $\varepsilon_1 < 0$ ,  $\varepsilon_2 > 0$  and  $\varepsilon_3 > 0$ :

$$C_{e_1}^{J_2} = \frac{\varepsilon_2}{\varepsilon_2\alpha + \varepsilon_3(1-\alpha) - \varepsilon_1} > 0$$

$$C_{e_2}^{J_2} = \frac{\varepsilon_3(1-\alpha) - \varepsilon_1}{\varepsilon_2\alpha + \varepsilon_3(1-\alpha) - \varepsilon_1} > 0$$

$$C_{e_3}^{J_2} = \frac{-\varepsilon_2(1-\alpha)}{\varepsilon_2\alpha + \varepsilon_3(1-\alpha) - \varepsilon_1} < 0$$

And for the concentration control coefficients:

$$C_{e_1}^s = \frac{1}{\varepsilon_2\alpha + \varepsilon_3(1-\alpha) - \varepsilon_1} > 0$$

$$C_{e_2}^s = \frac{-\alpha}{\varepsilon_2\alpha + \varepsilon_3(1-\alpha) - \varepsilon_1} < 0$$

$$C_{e_3}^s = \frac{-(1-\alpha)}{\varepsilon_2\alpha + \varepsilon_3(1-\alpha) - \varepsilon_1} < 0$$

With respect to the concentration control coefficient, note that  $C_1^s$  is positive while the two branch coefficients,  $C_2^s$  and  $C_3^s$  are negative. This is expected. The degree to which each of the output branches affects the concentration is in proportion to the amount of flux carried by the branch. This means that a branch that only carries a small amount of flux will have little effect on the branch species concentration.

Both flux control coefficients,  $C_1^{J_2}$  and  $C_2^{J_2}$ , are positive which we would expect. The flux control coefficient,  $C_3^{J_2}$  however is negative, indicating that changes in the activity of  $E_3$  decreases the flux in the other limb,  $J_2$ . This means there is **competition** in each output branch for flux. If one branch becomes more active, then it can ‘steal’ flux from the other branch. The amount stolen will depend on the various kinetic properties of the branch enzymes.

To answer what determines the competition between the output branches, we must look at the control equations in more detail. In particular, we must look at how the distribution of control is affected by different flux distributions, and the kinetics of the branch enzymes. In the following analysis,  $J_2$  will be the flux we observe as a result of perturbations to the enzymes in the branched pathway.

### Most Flux Through $J_3$

The first situation to consider is the case when most of the flux moves along  $J_3$  and only a small amount goes through the upper limb  $J_2$ , that is,  $\alpha \rightarrow 0$  and  $1 - \alpha \rightarrow 1$  – See Figure 11.2(b). Let us examine how the small amount of flux through  $J_2$  is influenced by the two branch limbs,  $E_2$  and  $E_3$ .

As  $\alpha \rightarrow 0$  and  $1 - \alpha \rightarrow 1$ , then:

$$C_{e_2}^{J_2} \rightarrow \frac{\varepsilon_1 - \varepsilon_3}{\varepsilon_1 - \varepsilon_3} = 1$$

$$C_{e_3}^{J_2} \rightarrow \frac{\varepsilon_2}{\varepsilon_1 - \varepsilon_3}$$

The first thing to note is that  $E_2$  tends to acquire proportional influence over its own flux,  $J_2$ . Since  $J_2$  only carries a very small amount of flux, any changes in  $E_2$  will have little effect on  $S$ , hence the flux through  $E_2$  is almost entirely governed by the activity of  $E_2$ . Because

Control Coefficient	Value
$C_{e_1}^{J_2}$	8.34
$C_{e_2}^{J_2}$	0.99
$C_{e_3}^{J_2}$	-8.51

**Table 11.2** Results showing high flux control coefficients in a simple branch model, see Tellurium script 11.1.

of the flux summation theorem and the fact that  $C_{e_2}^{J_2} = 1$ , it means that the remaining two coefficients must be equal and opposite in value. Since  $C_{e_3}^{J_2}$  is negative,  $C_{e_1}^{J_2}$  must be positive.

Unlike a linear pathway, the values for  $C_{e_2}^{J_2}$  and  $C_{e_1}^{J_2}$  are not bounded between zero and one. Depending on the values of the elasticities, it is possible for the control coefficients in a branched system to greatly exceed one [62, 75].

It is also possible to arrange the kinetic constants so that every step in the branch with respect to  $J_2$  has a control coefficient of unity (one of which must be -1 in order to satisfy the summation theorem). We could therefore claim that **every step** in the pathway is a rate limiting step with respect to  $J_2$ . This clearly shows again that rate limitation is not a simple concept as traditionally thought.

In a branched pathway it is possible to arrange the kinetic constants of the enzymes such that the feed branch has a flux control coefficient of +1, one of the output branch a coefficient of -1, and the other output branch a coefficient of +1. That is, **every step** in the pathway is equally rate limiting.

It is also possible to arrange the kinetic constants in the pathway such that the flux coefficients for  $E_1$  and  $E_3$  are much greater than one. This effect has been termed **ultrasensitivity** [75]. The Tellurium script 11.1 in the chapter Appendix illustrates a branched pathway with control coefficients over 8.0. Table 11.2 shows the results from the Tellurium script simulation.

The explanation for these high control coefficients is straight forward. Any changes in the two limbs that carry the high flux will have an adverse effect on the very small flux that is carried by  $J_2$ . Imagine a small stream coming off a large river. Any flooding in the large river is likely to have a huge impact on the small stream.

In a branched pathway it is possible to arrange the kinetic constants of the enzymes such the flux control coefficients in the feed and output branch can greatly **exceed one**.

Other than with an asymmetric distribution of flux, the ability to achieve high flux sensitivity

at a branch point depends on the relative values of the elasticities. For example, increasing the value  $\varepsilon_2$  relative to  $\varepsilon_3$  increases the sensitivity of the branch point. This could be achieved in a number of ways:

1.  $E_2$  can show positive cooperativity with respect to the branch species. That is, any changes in  $E_3$  become amplified through  $E_2$ .
2.  $v_3$  is operating in a more saturated regime compared to  $v_2$ . This will make  $\varepsilon_3$  smaller than  $\varepsilon_2$  and amounts to ensuring that the  $K_m$  for  $v_2$  is higher than the  $K_m$  of  $v_3$ .
3. Product inhibition on  $v_1$  is very small.

### Most Flux Through $J_2$

Let us now consider the other extreme, that is when most of the flux is through  $J_2$ . In other words,  $\alpha \rightarrow 1$  and  $1 - \alpha \rightarrow 0$  (See Figure 11.2(a)). Under these conditions the control coefficients yield:

$$\begin{aligned} C_{e_2}^{J_2} &\rightarrow \frac{\varepsilon_1}{\varepsilon_1 - \varepsilon_2} \\ C_{e_3}^{J_2} &\rightarrow 0 \end{aligned}$$

In this situation the pathway has effectively become a simple linear chain. The influence of  $E_3$  on  $J_2$  is negligible. By analogy, changing the flow of water in a small stream that comes off a large river will have a negligible effect on the rate of flow in the large river.

Figure 11.2 summarizes the changes in sensitivities at a branch point.

### Derivation by Implicit Differentiation

As was previously described in Chapter 7, we can also compute the control coefficients for a branched system by implicit differentiation. Start by writing out the rate of change of  $S$  at steady state for a simple branch as follows:

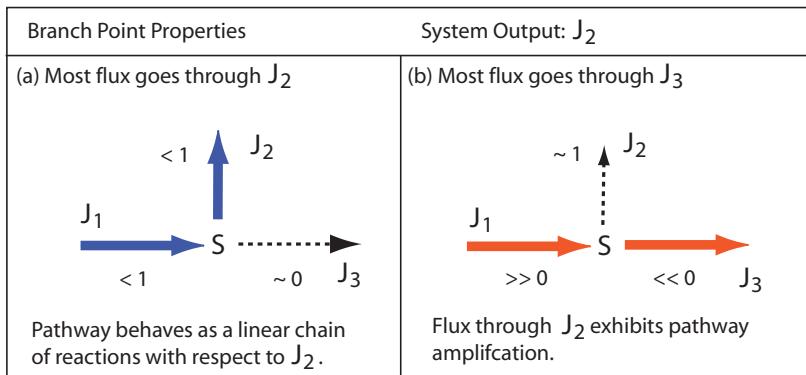
$$\frac{ds}{dt} = v_1 - v_2 - v_3 = 0$$

Assuming we wish to compute the coefficients with respect to  $E_1$ , we can write the equation as:

$$0 = v_1(s(e_1), e_1) - v_2(s(e_1)) - v_3 s(e_1))$$

Differentiating with respect to  $E_1$  gives:

$$0 = \frac{\partial v_1}{\partial s} \frac{ds}{de_1} + \frac{\partial v_1}{\partial e_1} - \frac{\partial v_2}{\partial S} \frac{ds}{de_1} - \frac{\partial v_3}{\partial s} \frac{ds}{de_1}$$



**Figure 11.2** The figure shows two flux extremes relative to the flux through branch  $J_2$ . In case (a) where most of the flux goes through  $J_2$ , the branch reverts functionally to a simple linear sequence of reactions comprised of  $J_1$  and  $J_2$ . In case (b), where most of the flux goes through  $J_3$ , the flux through  $J_2$  now becomes very sensitive to changes in activity at  $J_1$  and  $J_3$ . Given the right kinetic settings, the flux control coefficients can become ‘ultrasensitive’ with values greater than one (less than minus one for activity changes at  $J_3$ ). The values next to each reaction indicates the flux control coefficient for the flux through  $J_2$  with respect to activity at the reaction.

Scaling, setting  $\varepsilon_{e_1}^1 = 1$ , and solving for  $C_{e_1}^s$  yields:

$$C_{e_1}^s = \frac{1}{\varepsilon_2 \alpha + \varepsilon_3 (1 - \alpha) - \varepsilon_1}$$

where as before  $\alpha = J_2/J_1$ . The control coefficients for  $E_2$  and  $E_3$  can be derived in a similar manner.

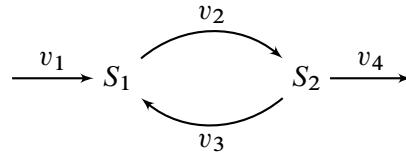
## 11.2 Futile or Substrate Cycles

Closely related to branched systems are cyclic pathways. A typical cyclic pathway is shown in Figure 11.3. For cycling to occur, both forward and back reactions must operate. It is typical to find that the forward and reverse reactions are chemically distinct. Often one reaction will be driven by ATP, while the other by hydrolysis of phosphate groups. Typical examples in metabolism include the cycle between glucose and glucose-6-phosphate, and the cycling between fructose-6-phosphate and fructose 1,6-bisphosphate. Such cycles have often been called futile cycles (or perhaps more accurately substrate cycles) because of the expenditure of free energy (as ATP) without any apparent benefit. A number of suggestions have been put forth to rationalize this apparent waste of energy. These include heat production, control of flux direction, metabolite buffering, and more sensitive control of the net

flux through the pathway. We will only consider the later here.

### Sensitivity Control

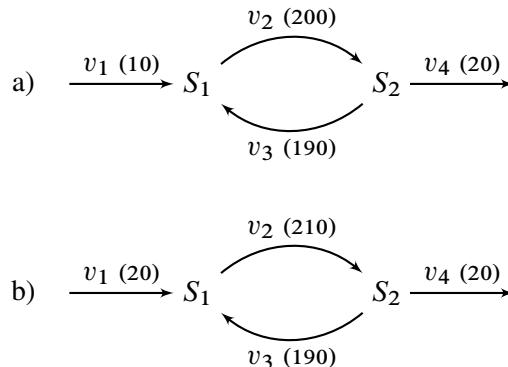
Figure 11.3 shows a typical cyclic pathway embedded in a linear chain. Of interest is the



**Figure 11.3** Cyclic Pathway.

sensitivity of the pathway flux,  $v_1$  or  $v_4$  to changes in  $v_2$ . The simplest assumption to make is that when we change  $v_2$ , there is no change in back flux,  $v_3$ . This could be for a number of reasons, for example  $v_3$  is saturated by its substrate  $S_2$ .

Figure 11.4 illustrates two situations – a reference state in (a), and a perturbation of 5% to  $v_2$  in (b). Assuming that all flux changes appear in output flux  $v_4$  and that  $v_3$  is not changed, the percentage change in  $v_4$  (or  $v_1$ ) is 100%, a twenty fold amplification.



**Figure 11.4** Amplification in a substrate cycle. (a) reference state, values refer to fluxes at various points, note that  $v_1 = v_2 - v_3$ . (b) activation of  $v_2$  by 5% leads to a 100% change in  $v_1$  and  $v_4$ . It assumes that  $v_3$  is not activated by any changes in  $S_2$ .

This effect can be easily quantified as follows. First, we note that the flux constraint is:

$$v_1 = v_2 - v_3$$

We then assume that a perturbation in  $v_2$  leads to the same change in  $v_1$ , that is:

$$\delta v_2 = \delta v_1$$

We can now compute the fractional changes in  $v_1$  and  $v_2$  as:

$$\frac{\delta v_1}{v_1} = \frac{\delta v_2}{v_2} \frac{v_2}{v_1}$$

The degree of amplification is then given by:

$$\frac{\delta v_1/v_1}{\delta v_2/v_2} = \frac{v_2}{v_1}$$

Since  $v_2 = v_1 + v_3$  then:

$$\frac{\delta v_1/v_1}{\delta v_2/v_2} = \frac{v_1 + v_3}{v_1} = 1 + \frac{v_3}{v_1} \quad (11.1)$$

This result shows that the higher the cycling rate ( $v_3$ ) compared to the through flux, the greater the amplification. This equation gives us the maximum degree of amplification possible. In practice,  $v_3$  will not remain unchanged because  $S_2$  rises. In addition,  $S_1$  will fall due to high consumption which will reduce  $v_2$  but increase  $v_1$  due to lower product inhibition. The resulting amplification is therefore a more complicated function than suggested by equation (11.1). However equation (11.1) gives the maximum possible amplification.

To carry out a more detailed analysis, we must turn to metabolic control analysis. We can examine the flux control coefficient for  $C_2^{J_1}$ :

$$C_2^{J_1} = \frac{\varepsilon_1^1 \varepsilon_2^4 (1 + v_3/v_1)}{D}$$

$$D = \varepsilon_1^1 \varepsilon_2^4 - \left(1 + \frac{v_3}{v_1}\right) (\varepsilon_1^1 \varepsilon_2^2 + \varepsilon_2^4 \varepsilon_1^2) + \frac{v_3}{v_1} (\varepsilon_1^1 \varepsilon_2^3 + \varepsilon_2^4 \varepsilon_1^3)$$

Let us simplify this equation by assuming that there is little or no product inhibition from  $S_2$  on to  $v_2$  and  $S_1$  on to  $v_3$ . This means that  $\varepsilon_1^3 = 0$  and  $\varepsilon_2^2 = 0$ . If we also multiply top and bottom by  $v_1$ , and using the relation  $v_1 + v_3 = v_2$ , we can simplify the control equation to:

$$C_2^{J_1} = \frac{\varepsilon_1^1 \varepsilon_2^4 v_2}{D}$$

$$D = \varepsilon_1^1 \varepsilon_2^4 v_1 - \varepsilon_2^4 \varepsilon_1^2 v_2 + \varepsilon_1^1 \varepsilon_2^3 v_3$$

Two things to note immediately from this equation. There must be product inhibition on the first step,  $\varepsilon_1^1$ , in order to get any sensitivity. If  $\varepsilon_1^1$  is zero then so is  $C_2^{J_1}$ . This is because all control is now on the first step. This highlights again the danger of using rate laws in models that are product insensitive because the use of such rate laws often give misleading or trivial results of no real interest. The second relatively simple statement to make from the above equation is the importance of  $\varepsilon_2^3$ . This elasticity is the activation of the reverse arm with respect to  $S_2$ . The larger this elasticity, the smaller the degree of amplification. This is

expected because any flux that flows back along the reverse cycle instead of into  $v_4$ , reduces the potential amplification factor. To analyze the equation further we can make additional simplifications.

We know that sensitivity increases when the cycling rate increases relative to the main flux,  $v_1$  and  $v_4$ . If  $v_2$  and  $v_3$  are much greater than  $v_1$ , we can simplify the equation further to:

$$C_2^{J_1} = \frac{v_2}{v_3\varepsilon_2^3/\varepsilon_2^4 - v_2\varepsilon_1^2/\varepsilon_1^1}$$

If the cycling rate is so high that  $v_2$  and  $v_3$  are almost indistinguishable, then we can see that maximal sensitivity is achieved when:

$$\frac{\varepsilon_2^3}{\varepsilon_2^4} + \frac{\varepsilon_1^2}{\varepsilon_1^1} \ll 1$$

This tells us that substrate activation of  $v_4$  by  $S_2$  should be stronger than substrate activation of  $S_2$  on  $v_3$ . Secondly, the product inhibition of  $S_1$  on  $v_1$  must be stronger than substrate activation of  $S_1$  on  $v_2$ . If we think about this in a thought experiment, these results are expected.

The requirements for amplification in substrate cycles is fairly complicated and questions remain whether real pathways use this mechanism *in vivo*. At this point we leave the topic of branches and cycles. In a subsequent chapter we will consider the dynamic properties of conserved cycles.

## Exercises

---

- Given the simple branch in Figure 11.1, prove the following theorems:

$$C_{e_1}^{J_2}(1 - \alpha) + C_{e_3}^{J_2} = 0$$

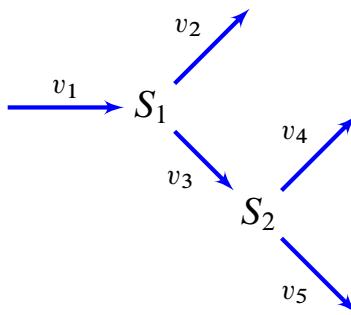
$$C_{e_1}^{J_3}\alpha + C_{e_2}^{J_3} = 0$$

- Prove that the following two theorems are true for the branch point in Figure 11.1:

$$C_{e_1}^s(1 - \alpha) + C_{e_3}^s = 0$$

$$C_{e_1}^s\alpha + C_{e_2}^s = 0$$

- Create a simulation of a simple branched system and arrange the rate constants so that one of the branches is hyper sensitive to changes in the other branch.
- Derive the flux branch theorems for the following multibranched system:



**Figure 11.5** Multi-Branched Pathway.

## Appendix

See Appendix D for more details of the Python tool Tellurium.

```

import tellurium as te
import numpy as np
r = te.loada (''''
    var S;
    ext Xo, w;
    J1: $Xo -> S; Vm1/Km1*(Xo-S/Keq) / (1+Xo/Km1+S/Km2);
    J2: S -> $w; Vm2*S^4/ (Km3+S^4);
    J3: S -> $w; Vm3*S/ (Km4+S);

    Xo = 9;
    S = 0.2;
    Vm1 = 1.4;
    Km1 = 0.4;
    Keq = 4.5;

    Km2 = 0.6;
    Vm2 = 0.05;
    Km3 = 0.8;
    Vm3 = 2.3;
    Km4 = 0.3;
'''')

r.steadyState()
print "Flux Control Coefficients:";
print r.getCC ('J2', 'Vm1');
print r.getCC ('J2', 'Vm2');
print r.getCC ('J2', 'Vm3');
print "Elasticities:";
```

```
e1 = r.getEE ('J1', 'S');
e2 = r.getEE ('J2', 'S');
e3 = r.getEE ('J3', 'S');
print e1, e2, e3;
print "Fluxes: ", r.J1, r.J2, r.J3;
```

**Listing 11.1** Simple Branched Pathway showing Flux Amplification.



# 12

## ***Negative Feedback***

### **12.1 Historical Background**

---

Feedback is widespread in biochemical networks and physiological systems. Some form of feedback permeates almost every known biological process. On the face of it, feedback is a simple process that involves sending a portion of the output back to the input. If the portion sent back reduces the input then the feedback is called negative feedback. Otherwise, it is called positive feedback.

Negative feedback is where part of the output of a system is used to reduce the magnitude of the system input.

#### **Water Clocks**

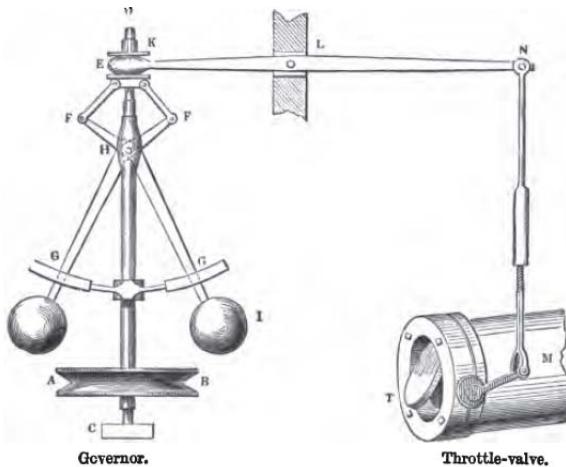
The concept of feedback control goes back at least as far as the Ancient Greeks [83]. Of some concern to the ancient Greeks was the need for accurate time keeping. In about 270 BC the Greek Ktesibios invented a float regulator for a water clock. The role of the regulator was to keep the water level in a tank at a constant depth. This constant depth yielded a constant flow of water through a tube at the bottom of the tank which filled a second tank at a constant rate. The level of water in the second tank thus depended on time elapsed.

Philon of Byzantium in 250 BC [28] is known to have kept a constant level of oil in a lamp using a float regulator and in the first century AD Heron of Alexandria experimented with float regulators for water clocks. Philon and particularly Heron (13 AD) [95] have left us

with an extensive book (*Pneumatica*) detailing many amusing water devices that employed negative feedback.

## Governors

It wasn't until the industrial revolution that feedback control, or devices for automatic control, became economically important. Probably the most famous modern device that employed negative feedback was the governor. Thomas Mead in 1787 took out a patent on a device that could regulate the speed of windmill sails. His idea was to measure the speed of the mill by the centrifugal motion of a revolving pendulum and use this to regulate the position of the sail. Very shortly afterwards in early 1788, James Watt is told of this device in a letter from his partner, Matthew Boulton. Watt recognizes the utility of the governor as a device to regulate the new steam engines that were rapidly becoming an important source of new power for the industrial revolution.



**Figure 12.1** A typical governor from J. Farley, *A Treatise on the Steam Engine: Historical, Practical, and Descriptive* (London: Longman, Rees, Orme, Brown, and Green, 1827, p436)

The novel device (Figure 12.1) employed two pivoted rotating fly-balls which were flung outward by centrifugal force. As the speed of rotation increased, the flyweights swung further out and up, operating a steam flow throttling valve which slowed the engine down. Thus, a constant engine speed was achieved automatically in the face of varying loads or steam pressure. So popular was this innovation that by 1868 it is estimated that 75,000 governors<sup>1</sup> were in operation in England. Many similar devices were subsequently invented to control a wide range of processes, including water wheels, telescope drives, and temperature and pressure controls.

<sup>1</sup> A History of Control Engineering, 1800-1930 By Stuart Bennett, 1979

The description of the governor illustrates the operational characteristics of **negative feedback**. The output of the device, in this case the steam engine speed, is “fed back” to control the rate of steam entering the steam engine and thus influence the engine speed.

In the early days devices for automatic control were designed through trial and error and little theory existed to understand the limits and behavior of feedback control systems. One of the difficulties with feedback control is the potential for instability. As the governor became more widespread, improvements were made in manufacturing mechanical devices which reduced friction. As a result engineers began to notice a phenomena they termed hunting. This was where after a change in engine load, the governor would begin to ‘hunt’ in an oscillatory fashion for the new stream rate that would satisfy the load. This effect caused considerable problems with maintaining a stable engine speed and resulted in James Maxwell and independently Vyshnegradskii, undertaking the first theoretical analysis of a negative feedback system.

Until the 20th century, feedback control was generally used as a means to achieve automatic control, that is to ensure that a variable, such as a temperature or a pressure was maintained at some set value. However, entirely new applications for feedback control emerged early in the 20th century, these included artillery tracking on ships and communication. In naval warfare a major issue is being able to accurately fire a weapon from a moving ship to a moving target. In the early part of the 20th century mechanical analog computers called rangekeepers (<https://en.wikipedia.org/wiki/Rangekeeper>) were developed that could continuously compute a target bearing, predict the future target position and make adjustments to the weapon to account for other factors such as wind. Tracking uses negative feedback and when employed in this mode it is called a **servomechanism** or servo for short.<sup>2</sup> Possibly the most important application of tracking came with the development of the feedback amplifier in the 1920s and 30s.

## Feedback Amplifiers

Amplification is one of the most fundamental tasks one can demand of an electrical circuit. One of the challenges facing engineers in the 1920’s was how to design amplifiers whose performance was robust with respect to the internal parameters of the system and which could overcome inherent nonlinearities in the implementation. This problem was especially critical to the effort to implement long distance telephone lines across the USA.

These difficulties were overcome by the introduction of the feedback amplifier, designed in 1927 by Harold S. Black [84] who was an engineer for Western Electric (the forerunner of Bell Labs). The basic idea was to introduce a negative feedback loop from the output of the amplifier to its input. At first sight, the addition of negative feedback to an amplifier might seem counterproductive. Indeed Black had to contend with just such opinions when introducing the concept. His director at Western Electric dissuaded him from following up on the idea and his patent applications were at first dismissed. In his own words, “our patent

<sup>2</sup>from the Latin servus ‘slave’

application was treated in the same manner as one for a perpetual motion machine” [7].

While Black’s detractors were correct in insisting that the negative feedback would reduce the gain of the amplifier, they failed to appreciate his key insight that the reduction in gain is accompanied by increased robustness of the amplifier and improved fidelity of signal transfer.

Unlike the steam engine governor which is used to stabilize some system variable, negative feedback in amplifiers is used to accurately track and amplify an external signal. These two applications highlight the two main ways in which negative feedback can be used, namely as a **regulator** or as a **servomechanism**.

Two modes of negative feedback:

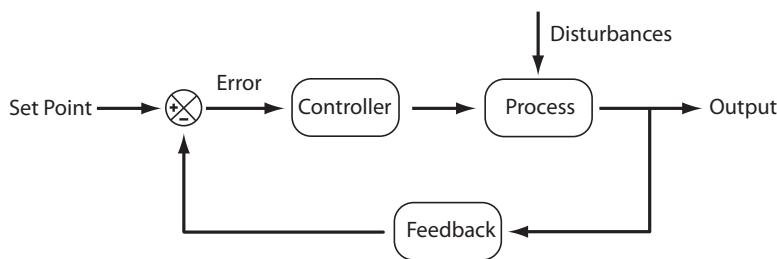
**Regulator:** Maintain a given variable at a constant level, e.g. Thermostat

**Servo:** Track a reference input, e.g. Op amp (Operational Amplifier) acting as a voltage follower.

As a regulator, negative feedback is used to maintain a controlled output at some constant desired level, whereas a servomechanism will slavishly track a reference input. We can see both applications at work in the eye. On the one hand there is the need to control the level of light entering the pupil. The diameter of the pupil is controlled by two antagonistic muscles. If the external light intensity increases, the muscles respond by reducing the pupil diameter, whereas the muscles increase the pupil diameter if the light intensity falls. The pupil reflex serves as an example of negative feedback using in a regulator mode. In contrast tracking an object involves maintaining the eyeball fixed on the object. In this mode the eye functions as a servomechanism.

Both regulator and servomechanism are implemented using the same operational mechanism. Figure 12.2 shows a generic negative feedback circuit. On the left of the figure can be found the input, sometimes called the desired value or more often the **set point**. If the circuit is used as a servomechanism then the output tracks the set point. As the set point changes the output follows. If the circuit is used as a regulator or homeostatic device then the set point is held constant and the output is maintained at or near the set point even in the face of disturbances elsewhere in the system.

The central mechanism in the feedback circuit is the generation of the error signal, that is the difference between the desired output (set point) and the actual output. The error is fed into a controller (often something that amplifies the error) which is used to increase or decrease the process. For example, if a disturbance on the process block reduces the output, then the feedback operates by generating a positive error, this in turn increases the process and restores the original drop in the output.

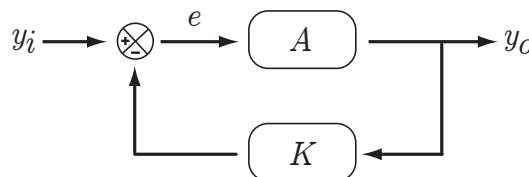


**Figure 12.2** Generic structure of a negative feedback system.

## 12.2 Simple Quantitative Analysis

The figure of the generic negative feedback circuit (Figure 12.2) is highly stylized which makes it difficult to identify the various component in a real biological system. In addition, biological system are invariably more complex with multiply nested feedback loops and multiple inputs and outputs. It is remarkable that even after 50 or 60 years of research, the role of many of the feedback systems in biochemical networks take is still speculative.

In the remainder of this section we will consider some basic properties of negative feedback systems. The simplest way to think about feedback quantitatively is by reference to Figure 12.3.



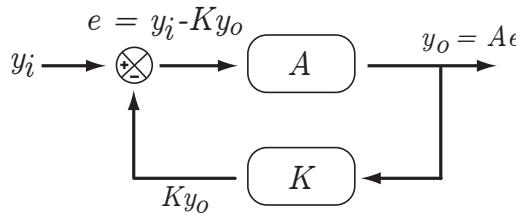
**Figure 12.3** Generic structure of a negative feedback system.

We will assume some very simple rules that govern the flow of information in this feedback system. For example, the output signal,  $y_o$  will be given by the process  $A$  multiplied by the error,  $e$ . The feedback signal will be assumed to be proportional to  $y_o$ , that is  $Ky_o$ . Finally, the error signal,  $e$  will be given by the difference between the set point,  $y_i$  and the feedback signal,  $Ky_o$  (Figure 12.4).

$$y_o = Ae, \quad e = y_i - Ky_o$$

From these simple relations we can eliminate  $e$  to obtain:

$$y_o = \frac{Ay_i}{1 + AK} \quad \text{or more simply} \quad y_o = G y_i \quad (12.1)$$



**Figure 12.4** Generic structure of a negative feedback system.

where  $G = A/(A + AK)$  is called the gain of the feedback loop, often called the **closed loop gain**. **Gain** is a term that is commonly used in control theory and refers to the scalar change between an input and output. Thus a gain of two means that a unit change in the input will result in twice the change in the output.

The **gain** is a measure of the change that occurs between a signal output and its input. A gain of two means that the output will change two times in magnitude compared to a change in the input.

In addition to the closed loop gain, engineers also define two other gain factors, the **open loop gain** and the **loop gain**. The open loop gain is the gain from process,  $A$ , alone. It is the gain one would achieve if the feedback loop were absent.

The loop gain is the gain from the feedback and process  $A$  combined,  $AK$ . The loop gain is an important quantity when discussing the stability and performance of feedback circuits. Figure 12.4 illustrates the different types of gain in a feedback circuit.

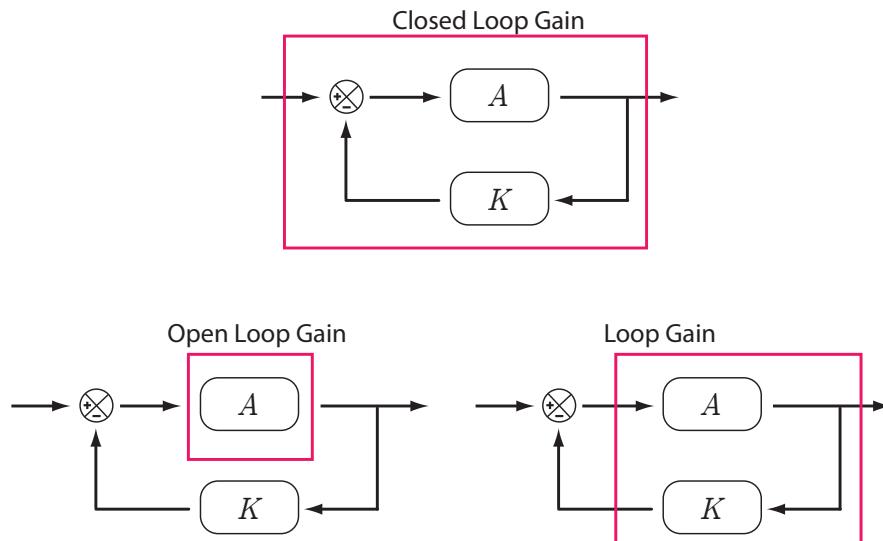
We can use equation (12.1) to discover some of the basic properties of a negative feedback circuit. The first thing to note is that as the loop gain,  $AK$ , increases, the system behavior becomes more dependent on the feedback loop and less dependent on the rest of the system:

$$\text{when } AK \gg 1 \text{ then } G \simeq \frac{A}{AK} = \frac{1}{K}$$

This apparently innocent effect has significant repercussions on other aspects of the circuit. To begin with, as the system becomes less dependent on  $A$ , so does variation in the properties of  $A$ . Feedback makes the performance of the system independent of variation in  $A$ . Such variation might include noise or variation as a result of the manufacturing process or in the case of biological systems, genetic variation. To be more precise we can compute the sensitivity of the gain  $G$  with respect to variation in  $A$ .

$$\frac{\partial G}{\partial A} = \frac{\partial}{\partial A} \frac{A}{1 + AK} = \frac{1}{(1 + AK)^2}.$$

That is the sensitivity drops as the loop gain,  $AK$  increases. If we consider the relative



**Figure 12.5** Generic structure of a negative feedback system.

sensitivity we find:

$$\frac{\partial G}{\partial A} \frac{A}{G} = \frac{1}{1 + AK}$$

In addition to resistance to parameter variation, feedback also confers resistance to disturbances in the output. Suppose that a nonzero disturbance  $d$  changes the output. The system behavior is then described by

$$y = Ae - d \quad e = u - Ky.$$

Eliminating  $e$ , we find

$$y = \frac{Au - d}{1 + AK}.$$

The sensitivity of the output to the disturbance is then

$$\frac{\partial y}{\partial d} = -\frac{1}{1 + AK}.$$

The sensitivity decreases as the loop gain  $AK$  is increased. In practical terms, this means that the imposition of a load on the output, for example a current drain in an electronic circuit, protein sequestration on a signaling network or increased demand for an amino acid will have less of an effect on the circuit as the feedback strength increases. In electronics, this property allows engineers to divide a circuit into functional modules.

Last but not least, feedback also improves the fidelity of the response. That is, for a given change in the input, a system with feedback is more likely to faithfully reproduce the input at the output than a circuit without feedback. An ability to faithfully reproduce signals is

critical in electronics communications and in fact it was this need that was the inspiration for the development of negative feedback in the early electronics industry.

Consider the case where the amplifier  $A$  is nonlinear. For example a cascade pathway exhibiting a sigmoid response. Then the behavior of the system  $G$  (now also nonlinear) is described by

$$G(y_i) = y_o = A(e) \quad e = y_i - Ky_o = y_i - KG(y_i).$$

Differentiating we find

$$G'(y_i) = A'(y_i) \frac{de}{dy_i} \quad \frac{de}{dy_i} = 1 - KG'(y_i).$$

Eliminating  $\frac{de}{dy_i}$ , we find

$$G'(y_i) = \frac{A'(y_i)}{1 + A'(y_i)K}.$$

We find then, that if  $A'(y_i)K$  is large ( $A'(y_i)K \gg 1$ ), then

$$G'(y_i) \approx \frac{1}{K},$$

so, in particular,  $G$  is approximately linear. In this case, the feedback compensates for the nonlinearities  $A(\cdot)$  and the system response is not distorted. Another feature of this analysis is that the slope of  $G(\cdot)$  is less than that of  $A(\cdot)$ , i.e. the response is “stretched out”. For instance, if  $A(\cdot)$  is saturated by inputs above and below a certain “active range”, then  $G(\cdot)$  will exhibit the same saturation, but with a broader active range.

A natural objection to the implementation of feedback as described above is that the system sensitivity is not actually reduced, but rather is shifted so that the response is more sensitive to the feedback  $K$  and less sensitive to the amplifier  $A$ . However, in each of the cases described above, we see that it is the nature of the loop gain  $AK$  (and not just the feedback  $K$ ) which determines the extent to which the feedback affects the nature of the system. This suggests an obvious strategy. By designing a system which has a stable feedback gain but a large “sloppy” amplifier, one ensures that the loop gain is large and the behavior of the system is satisfactory. Engineers employ precisely this strategy in the design of electrical feedback amplifiers, regularly making use of amplifiers with gains several orders of magnitude larger than the feedback gain (and the gain of the resulting system).<sup>3</sup> A typical operational amplifier such as the common 741 will have an open loop gain of approximately 200,000. Given this level of gain, a 741 without any feedback is almost useless because with only the smallest change in voltage at the input the output voltage will experience a huge voltage swing.<sup>4</sup>

---

<sup>3</sup>I wish to acknowledge Brian Ingalls for useful discussions on this topic.

<sup>4</sup>Even a 10  $\mu$ V change at the input will in theory result in change of 200 volts at the output. Note that op amps are generally powered by  $\pm 15$  volts, meaning that a small voltage change will result in a full swing from -15 to +15 volts resulting in the op amp saturating. Only with some form of feedback do op amps become useful.

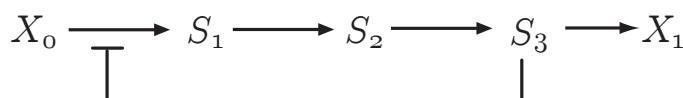
Summary of useful properties from negative feedback:

1. Amplification of signal.
2. Robustness to internal component variation.
3. High fidelity of signal transfer.
4. Low output impedance meaning the load does not affect the performance of the circuit.

## 12.3 Negative Feedback in Biochemical Systems

It was Umbarger (Umbarger, 1956) and Yates and Pardee (Yates & Pardee, 1956) who discovered feedback inhibition in the isoleucine biosynthesis pathway and the inhibition of aspartate transcarbamylase in *E. coli*. It wasn't long afterwards that some researchers began to investigate such feedback systems mathematically. Probably the most extensive mathematical analysis of biochemical feedback was conducted by Savageau (Savageau, 1972; Savageau, 1974; Savageau, 1976) and Burns and Kacser (Burns, 1971; Kacser & Burns, 1973) and Othmer and Tyson (Othmer, 1976; Tyson & Othmer, 1978) in the 1970s and Dibrov et. al. in the early 1980s (Dibrov et al., 1982). More recently, Cinquin and Demongeot have published an interesting review on the roles of feedback in biological systems (Cinquin & Demongeot, 2002).

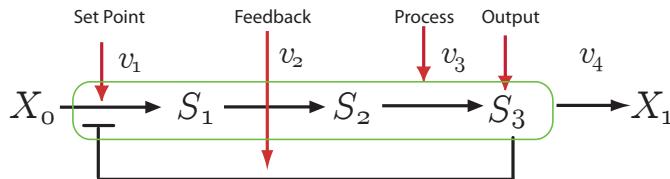
In the last section we considered a simple analysis of negative feedback and its behavioral effects. The treatment was however very generic and the question we wish to address here is how can we apply the same kind of analysis to biochemical feedback systems? This question is harder to answer than it seems. To begin with, biochemical systems are governed by nonlinear rate laws not the simple linear rules we used in the previous analysis. Secondly how do we map the generic diagram (Figure 12.2) onto a biochemical feedback circuit (Figure 12.6), on the surface they look similar but the initial impression is misleading.



**Figure 12.6** Simple four step pathway with negative feedback.

In order to be clear we need to identify the input (set point), output, the feedback loop and the process block in the biochemical network Figure 12.6.

In naturally evolved systems it is sometimes difficult to identify the various parts in a negative feedback circuit. Control is often achieved using allosteric enzymes which have



**Figure 12.7** Simple four step pathway with negative feedback.

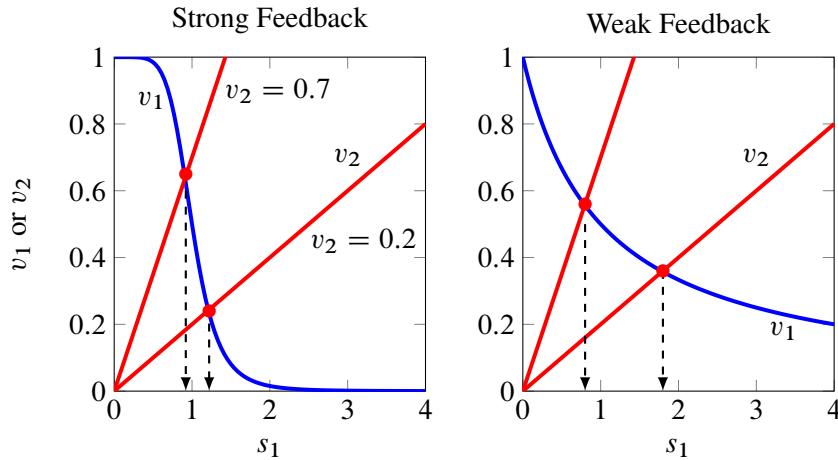
distinct binding sites for the controlling species that are separate from the main active site. The most difficult element to identify is the set point. In biochemical networks the set point is embedded in the regulated enzyme, often in the form of the half saturation constant of the allosteric regulator.

When comparing the block diagram (Figure 12.2) to the biological pathway in Figure 12.7 it may not be apparent how the two representations can be matched. It is however possible to pair each component in the block diagram to an equivalent component in the biological pathway. Thus the output,  $y_o$  in the block diagram corresponds to the concentration of  $S_3$ . The negative feedback component  $k$  corresponds to the interaction of  $S_3$  with the allosteric enzyme in the first step. The set point,  $y_i$  is more problematic but it is most likely embedded in the kinetic characteristics of the allosteric enzyme. Finally the controller  $A$  is represented by the steps  $v_1$ ,  $v_2$ , and  $v_3$ . The load on the system is represented by the last step,  $v_4$  and other disturbances can be assigned to  $v_1$ ,  $v_2$ ,  $v_3$ , and the input concentration,  $X_o$ .

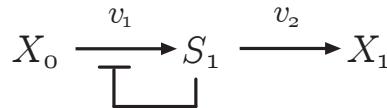
The flux through the pathway can also be considered an output because the reaction rate through the last step is a function  $S_3$ .

### Graphical understanding of feedback

If is possible to appreciate some of the effects of negative feedback using a graphical approach. Consider the much simplified negative feedback network shown in Figure 12.9. This network only has two steps,  $v_1$  and  $v_2$ . Figure 12.8 shows two plots, one with strong and the other with weak feedback. The strength of the feedback is modeled using a simple Hill like equation, for stronger feedback the Hill equation has a higher Hill coefficient. The plots show the reaction rates  $v_1$  and  $v_2$  as a function of the intermediate species,  $S_1$ , where  $S_1$  can negatively feedback on to the first step. We assume that the second step follows first-order kinetics so that the  $v_2$  is a straight line with rate law  $k_2 s_1$ . The feedback response curve shows a decline from high to low as  $S_1$  increases. For strong feedback the decline is much steeper (left plot). If the rate constant for the second step is changed, this changes the slope of  $v_2$ , this is equivalent to a perturbation in the system. In the case of weak feedback, changes in  $v_2$  result in significant changes to  $S_1$ , this is because the feedback response is shallow. In contrast, when we have strong feedback (left panel), where the slope is very steep, any changes in  $v_2$  results in only small changes in  $S_1$ . In this way we can see how strong negative feedback can buffer changes in  $v_2$ .



**Figure 12.8** Plot of  $v_1$  and  $v_2$  versus the concentration of  $S_1$  for a simple two step pathway with negative feedback.  $v_2$  is governed by the rate laws  $k_2 s_1$ . Two perturbations in  $k_2$  and its effect on  $v_2$  are shown. In the left panel where the feedback is strong, changes in  $k_2$  have hardly any effect on  $S_1$ . On the right panel, the same change in  $k_2$  results in a much larger change in  $S_1$ . This illustrates the homeostatic property of negative feedback. Left Panel:  $v_1 = 1/(1 + s_1^4)$ , Right Panel:  $v_1 = 1/(1 + s_1)$



**Figure 12.9** Simple two step pathway with negative feedback:  $v_1 = 1/(1 + s_1^n)$ ,  $v_2 = k_2 s_1$

### Control Analysis

Just as we did earlier, we can derive the flux and concentration control coefficients in terms of the elasticities. For convenience, we will write out the theorems in matrix form (See equation (8.17)), note the presence of the **feedback term**,  $\varepsilon_2^1$  in the matrix.

$$\begin{bmatrix} C_1^J & C_2^J & C_3^J \\ C_1^{s_1} & C_2^{s_1} & C_3^{s_1} \\ C_1^{s_2} & C_2^{s_2} & C_3^{s_2} \end{bmatrix} \begin{bmatrix} 1 & -\varepsilon_1^1 & -\varepsilon_2^1 \\ 1 & -\varepsilon_1^2 & -\varepsilon_2^2 \\ 1 & 0 & -\varepsilon_2^3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Rearranging the matrix equation yields:

$$\begin{bmatrix} C_1^J & C_2^J & C_3^J \\ C_1^{s_1} & C_2^{s_1} & C_3^{s_1} \\ C_1^{s_2} & C_2^{s_2} & C_3^{s_2} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -\varepsilon_1^1 & -\varepsilon_2^1 \\ 1 & -\varepsilon_1^2 & -\varepsilon_2^2 \\ 1 & 0 & -\varepsilon_2^3 \end{bmatrix}^{-1}$$

Inverting the elasticity matrix yields the following equations for the flux control coefficients with and without feedback to illustrate the difference in the results.

With Feedback	Without Feedback
$C_{e_1}^J = \frac{\varepsilon_1^2 \varepsilon_2^3}{\varepsilon_1^2 \varepsilon_2^3 - \varepsilon_1^1 \varepsilon_2^3 + \varepsilon_1^1 \varepsilon_2^2 - \varepsilon_1^2 \varepsilon_1^1}$	$C_{e_1}^J = \frac{\varepsilon_1^2 \varepsilon_2^3}{\varepsilon_1^2 \varepsilon_2^3 - \varepsilon_1^1 \varepsilon_2^3 + \varepsilon_1^1 \varepsilon_2^2}$
$C_{e_2}^J = \frac{-\varepsilon_1^1 \varepsilon_2^3}{\varepsilon_1^2 \varepsilon_2^3 - \varepsilon_1^1 \varepsilon_2^3 + \varepsilon_1^1 \varepsilon_2^2 - \varepsilon_1^2 \varepsilon_1^1}$	$C_{e_2}^J = \frac{-\varepsilon_1^1 \varepsilon_2^3}{\varepsilon_1^2 \varepsilon_2^3 - \varepsilon_1^1 \varepsilon_2^3 + \varepsilon_1^1 \varepsilon_2^2}$
$C_{e_3}^J = \frac{\varepsilon_1^1 \varepsilon_2^2 - \varepsilon_1^2 \varepsilon_1^1}{\varepsilon_1^2 \varepsilon_2^3 - \varepsilon_1^1 \varepsilon_2^3 + \varepsilon_1^1 \varepsilon_2^2 - \varepsilon_1^2 \varepsilon_1^1}$	$C_{e_3}^J = \frac{\varepsilon_1^1 \varepsilon_2^2}{\varepsilon_1^2 \varepsilon_2^3 - \varepsilon_1^1 \varepsilon_2^3 + \varepsilon_1^1 \varepsilon_2^2}$

Table 12.1 shows the corresponding concentration control coefficients in the presence of negative feedback. The first thing to note is that the addition of negative feedback adds a new term to the denominator,  $-\varepsilon_1^2 \varepsilon_2^1$ . This term is positive and therefore increases the value of the denominator.

Consider first the concentration control coefficients. To simplify matters assume that the product elasticities,  $\varepsilon_1^1$  and  $\varepsilon_2^2$  are zero. This greatly simplifies the equations.

Species or Flux			
Enzyme Step	$s_1$	$s_2$	$J$
$e_1$	$\frac{\varepsilon_2^3}{\varepsilon_1^2(\varepsilon_2^3 - \varepsilon_2^1)}$	$\frac{1}{\varepsilon_2^3 - \varepsilon_2^1}$	$\frac{\varepsilon_2^3}{\varepsilon_2^3 - \varepsilon_2^1}$
$e_2$	$\frac{-1}{\varepsilon_1^2}$	0	0
$e_3$	$-\frac{\varepsilon_2^1}{\varepsilon_1^2(\varepsilon_2^3 - \varepsilon_2^1)}$	$-\frac{1}{\varepsilon_2^3 - \varepsilon_2^1}$	$-\frac{\varepsilon_2^1}{\varepsilon_2^3 - \varepsilon_2^1}$

**Table 12.1** Control coefficients for simple negative feedback pathway. Assumes that production inhibition on  $v_2$  and  $v_3$  is absent.

The first observation to make is that  $C_{e_2}^{s_2}$  and  $C_{e_2}^J$  are zero. That is, disturbances inside the feedback have no effect on the signal species,  $S_2$  or the flux.

The second observation is that both  $C_{e_1}^{s_2}$  and  $C_{e_3}^{s_2}$  can be reduced to very small values if the feedback strength is significant, that is  $\varepsilon_2^1 \ll 0$ . This result illustrates the buffering capacity of the negative feedback loop, locking  $S_2$  into a very narrow range, see Figure (12.8).

The response of the pathway to  $X_o$  is given by the product  $C_{e_1}^{s_2} \varepsilon_{x_o}^1$ . If  $\varepsilon_{x_o}^1$  lies between 0.5 and 1.0 for a typical Michaelian enzyme, the source species,  $X_o$ , will have little influence

over the level of signal species,  $s_2$ . It is interesting to speculate that most enzymes that are regulated by feedback inhibition also show cooperativity with respect to the source species. This means that  $\varepsilon_{x_o}^1$  can be of the order of 4.0 or more suggesting that the cooperativity attempts to restore some control by the source species. In other words it suggests some degree of control by supply over the output  $s_2$ . This leads to complementary question of the how demand can affect the pathway.

If we make the feedback elasticity,  $\varepsilon_2^1 \ll 0$ , flux control coefficient for the final step,  $C_{e_3}^J$  tends to unity. That is, all control moves out of the feedback loop. In terms of a steam engine analogy, it is equivalent to being able to change the work load on the steam engine without loss of power. If  $C_{e_3}^J$  tends to one, then by the summation theorem, both  $C_{e_1}^J$  and  $C_{e_2}^J$  tend to zero. From this perspective the regulated steps therefore has no control over the flux.

Control over  $S_3$  and  $J$  is the opposite to the other, namely as the feedback strength increases:

$$C_{e_3}^J \rightarrow 1$$

$$C_{e_3}^{s_2} \rightarrow 0$$

Traditionally, allosteric enzymes have been considered flux controllers. The analysis here suggests the opposite. Allosteric enzymes, when part of a negative feedback loop are very poor controllers of flux. Instead the feedback allows distal steps to be good controllers of a pathway flux. For demand driven systems this is a logical arrangement.

The nagging suspicion remains however that intuitively the regulated step must have some kind of ability to enable the system to operate in the way it does. There are at least two answers to this. The first is that the feedback elasticity will be strongly negative. For example, if the regulated step were determined by a modified Hill like equation such as:

$$v = \frac{V_{\max} x_o}{s^n + x_o + K_m}$$

where  $S$  is the feedback signal, then the elasticity of the reaction rate with respect to the signal is given by:

$$\varepsilon_s^v = \frac{n x_o K_m}{K_m + (s/K_m)^n}$$

We see that at low signal,  $S$ , the elasticity is proportional to  $n$ , the Hill coefficient. The reaction is therefore very sensitive to changes in the signal molecule.

The second way to answer the question is to consider the pathway with and without the negative feedback loop. We can compare for example the flux control coefficient on the first step with and without negative feedback. When we remove the negative feedback the pathway will change to a new state where the concentrations of  $S_1$  and  $S_2$  are higher. To make the comparison fair we should adjust the level of enzyme in the first step to restore the levels of  $S_1$  and  $S_2$  to the values they had before the feedback was removed. When we do this we will also automatically restore the pathway flux to its original value. In practice

this means that all the elasticities except for the feedback elasticity are exactly in the same in both systems.

We can now ask the question what are the relative values for the flux control coefficients in step one? We can derive the ratio of the flux control of the regulated step of the open configuration,  $C_{e_1}^J$  to the closed configuration,  ${}^nC_{e_1}^J$ . This can be shown to be:

$$\frac{C_{e_1}^J}{{}^nC_{e_1}^J} = 1 - \frac{\varepsilon_1^2 \varepsilon_2^3 \varepsilon_{\text{fd}}}{\varepsilon_1^1 \varepsilon_2^2 \varepsilon_3^4 - \varepsilon_1^1 \varepsilon_2^2 \varepsilon_3^3 - \varepsilon_1^1 \varepsilon_2^3 \varepsilon_3^4 + \varepsilon_1^2 \varepsilon_2^3 \varepsilon_3^4}$$

The numerator term,  $\varepsilon_1^2 \varepsilon_2^3 \varepsilon_{\text{fd}}$  is negative and the denominator positive. The ratio is therefore negative. However we are subtracting this negative value from one. This means that overall the expression,  $C_{e_1}^J / {}^nC_{e_1}^J$  must be  $> 1$ . In other words, the flux control exerted by the open configuration will be **greater** than the control exerted by the closed configuration. This shows that the regulated step has the ability to throttle the flux. In the closed configuration, this property is hidden and operates silently by actively adjusting the flux so that the end product,  $X_3$  is stabilized.

To make the analysis simpler, us assume that the product inhibition elasticities,  $\varepsilon_1^1$  and  $\varepsilon_2^2$  are both zero. We now take the ratio of the flux control coefficient on the first step,  $C_1^J$  without feedback to the same coefficient with feedback, which will be denoted by  ${}^nC_1^J$ :

$$\frac{C_1^J}{{}^nC_1^J} = \frac{\varepsilon_1^2 \varepsilon_2^3 - \varepsilon_2^1 \varepsilon_1^2}{\varepsilon_1^2 \varepsilon_2^3} = 1 - \frac{\varepsilon_2^1 \varepsilon_1^2}{\varepsilon_1^2 \varepsilon_2^3} = 1 - \frac{\varepsilon_2^1}{\varepsilon_2^3}$$

Given that  $\varepsilon_2^1$  is negative, the term,  $-\varepsilon_2^1/\varepsilon_2^3$  is positive. That is:

$$C_1^J > {}^nC_1^J$$

This tells us that for a negative feedback loop to be effective, the flux control coefficient in the unregulated pathway must be higher than the flux control coefficient with feedback. There is however another way to measure the importance of a regulated step, and that is to look at the loop gain, something we'll examine in the next section.

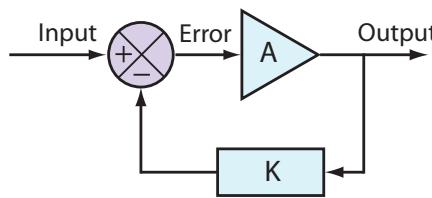
## 12.4 The PFK Paradox

---

In the last section a simple analysis was made of a feedback system using approaches developed from MCA and BST. How does this analysis and the broader literature on MCA and BST relate to the existing body of engineering control theory? This has been covered in detail by Ingalls [61] and Rao [100], here we will give a flavor of the connection particularly in relation to negative feedback. Not all effects of negative feedback will be discussed here and the reader is referred to some recent articles for additional details [29, 65].

As we saw earlier in the chapter, the classic diagram often used to depict negative feedback in control theory is that shown in Figure 12.10. In this diagram,  $K$  is the fraction of output

that is fed back to the summing junction. The summing junction computes the difference between the input (also called the set-point depending on the application), and the signal fed back from the output. The output is a function of the error, often a simple proportional relationship (hence called proportional control). To put it in more concrete terms, the  $A$  block might be a heater in a room and the output the room's temperature. The input is the desired temperature. If the room is hotter than the desired temperature, the error signal will be negative so that the heater is turned down. The opposite happens if the room is cooler than the desired temperature.



**Figure 12.10** Generic negative feedback system as depicted often by engineers.

As was done earlier we can derive the relationship between the input and output in Figure 12.10 by noting the following relationships. The error signal can be written as the difference between the input and the return loop:  $\text{error} = \text{input} - \text{output} \times K$ . The output can be written in terms of the gain  $A$  and the error:  $\text{output} = A \times \text{error}$ . We can combine both equations and solve for the output in terms of the input:

$$\text{output} = \frac{\text{input} \times A}{1 + A K} \quad (12.2)$$

In control theory this is called the closed-loop transfer function. In practice it is often expressed in the frequency domain but this need not concern us here. What is more interesting is that the control coefficient equations are directly related to the closed-loop transfer function [61]. The only difference is that whereas the full transfer function is defined over a range of input frequencies, the control coefficients are defined only at a single frequency of zero, the so-called DC response. For a two step pathway with feedback, the concentration control coefficient (8.16) is given by:  $1/(\varepsilon_s^2 - \varepsilon_s^1)$ . We therefore state the following equivalence [61, 60]:

$$C_{e_1}^s \frac{\delta e_1}{e_1} = \frac{\delta s}{s} = \frac{1}{\varepsilon_s^2 - \varepsilon_s^1} \frac{\delta e_1}{e_1} = \frac{\text{input} \times A}{1 + A K}$$

The input in the closed-loop transfer function matches the change in enzyme ( $\delta e_1/e_1$ ), but could be any input into  $v_1$ . In the case of  $e_1$  we assume that the elasticity of  $v_1$  with respect to  $e_1$ ,  $\varepsilon_{e_1}^1$ , is one. For other inputs this might not be the case and the corresponding elasticity would need to be explicitly given. For example if the input parameter is a signal  $P$ , then the term  $\delta e_1/e_1$  would be replaced with  $(\delta p/p) \varepsilon_p^1$ .

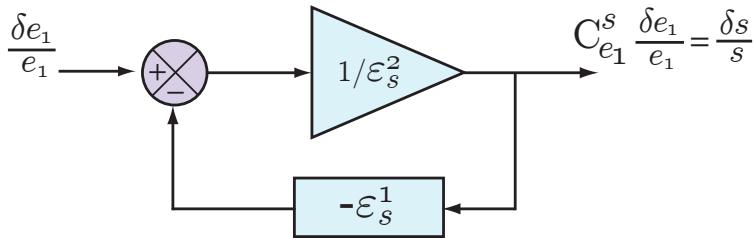
Rearranging terms by dividing top and bottom by  $\varepsilon_s^2$  we obtain:

$$C_{e_1}^s \frac{\delta e_1}{e_1} = \frac{\delta s}{s} = \frac{\delta e_1}{e_1} \frac{1/\varepsilon_s^2}{1 - \varepsilon_s^1/\varepsilon_s^2} = \text{input} \frac{A}{1 + A K}$$

From this we can determine that:

$$\text{input} = \frac{\delta e_1}{e_1}, \quad \text{output} = \frac{\delta s}{s} \quad A = \frac{1}{\varepsilon_s^2} \quad \text{and} \quad K = -\varepsilon_s^1$$

This generalizes to any length feedback loop. The feedback term,  $K$  will always be equal to the negative of the feedback elasticity. The expression for  $A$ ,  $1/\varepsilon_s^2$  is the control coefficient  $C_{e_1}^s$ , when there is **no feedback**.  $C_{e_1}^s$  is the control coefficient for the unregulated system (set  $\varepsilon_s^1$  to zero in  $C_{e_1}^s$ ) and is the gain of the system without feedback assuming everything else being equal. Figure 12.11 overlays the generalized feedback diagram with the equivalent elasticity terms. The error computed from the set point and feedback is the difference between the two terms  $\delta e_1/e_1$  and  $\varepsilon_s^1 C_{e_1}^s$  (recall that  $\varepsilon_s^1$  is negative).



**Figure 12.11** Mapping the standard engineering feedback control diagram to equivalent biochemical components.

As already introduced, the product  $AK$  is the loop gain which has a special status in control theory. It is one of the factors that determines the stability of a system with negative feedback and general performance of the system. In the reaction pathway,  $\varepsilon_s^1/\varepsilon_s^2$  is the loop gain. For the unregulated pathway,  $\varepsilon_s^1$  will be absent and under these conditions the control coefficient equals  $1/\varepsilon_s^2$ . We can therefore rewrite the loop gain as:

$$\text{Loop Gain} = AK = -\varepsilon_s^1 {}^u C_{e_1}^s$$

where  ${}^u C_{e_1}^s$  is computed for the *unregulated pathway*. The superscript  $u$  is used to indicate that this is the control coefficient for the unregulated pathway. The minus signs comes from the observation that  $K$  is equal to the negative of the feedback elasticity. For example for the four step pathway where  $S_3$  is the return signal (Figure 12.13), the loop gain will equal:

$$\text{Loop Gain} = -\varepsilon_3^1 {}^u C_{e_1}^{S_3} \tag{12.3}$$

where  ${}^u C_{e_1}^{S_3}$  is the control coefficient for the *unregulated* but equivalent system and  $\varepsilon_3^1$  is the elasticity of the feedback loop.

The analysis on the two step pathway can be extended to larger pathways with negative feedback and the conclusions from the two step study apply equally. For example, consider the four step pathway shown in Figure 12.13. In total there will be four flux control coefficients and twelve concentration control coefficients corresponding to the three species,  $S_1$ ,  $S_2$  and  $S_3$  and the four enzyme catalyzed reactions,  $E_1$  to  $E_4$ . The denominator for each control equation is the same and is given by equation (12.4).

$$D = \varepsilon_1^1 \varepsilon_2^2 \varepsilon_3^4 - \varepsilon_1^1 \varepsilon_2^2 \varepsilon_3^3 - \varepsilon_1^1 \varepsilon_2^3 \varepsilon_3^4 + \varepsilon_1^2 \varepsilon_2^3 \varepsilon_3^4 - \varepsilon_1^2 \varepsilon_2^3 \varepsilon_{\text{fd}} \quad (12.4)$$

where  $\varepsilon_{\text{fd}}$  is the feedback elasticity.  $\varepsilon_{\text{fd}}$  is equivalent to  $\varepsilon_3^1$ . The symbol  $\varepsilon_{\text{fd}}$  is used to help the reader identify the feedback elasticity more easily. Here we will focus on four of the control coefficients shown in Table 12.2.

Control Coefficient	Numerator
$C_{e_1}^{S_3}$	$\varepsilon_1^2 \varepsilon_2^3$
$C_{e_4}^{S_3}$	$\varepsilon_1^1 \varepsilon_2^3 - \varepsilon_1^1 \varepsilon_2^2 - \varepsilon_1^2 \varepsilon_2^3$
$C_{e_1}^J$	$\varepsilon_1^2 \varepsilon_2^3 \varepsilon_3^4$
$C_{e_4}^J$	$-\varepsilon_1^1 \varepsilon_2^2 \varepsilon_3^3 - \varepsilon_1^2 \varepsilon_2^3 \varepsilon_{\text{fd}}$

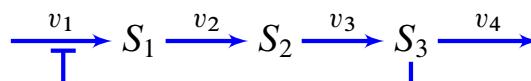
**Table 12.2** Control Coefficients and Corresponding Numerators of Control Equations. The feedback elasticity is highlighted as  $\varepsilon_{\text{fd}}$

$C_{e_1}^J$  is reduced by the presence of the feedback term  $-\varepsilon_1^2 \varepsilon_2^3 \varepsilon_{\text{fd}}$  in the denominator. As with the simple two step pathway, increasing the strength of the negative feedback reduces the flux control coefficient for regulated step  $e_1$ . At the same time the influence of the last step,  $C_{e_4}^J$ , tends to increase (to a maximum of one, see Figure 12.12).

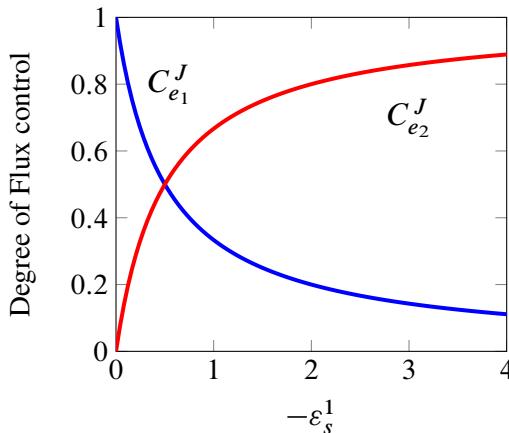
Likewise the concentration control coefficient,  $C_{e_4}^{S_3}$  tends to zero as the feedback strength increases. The negative feedback locks  $S_3$  into a narrow range. These results have been shown in previous sections but it is worth reiterating again. In general all steps within the signal loop will have little influence on either the flux or the level of  $S_3$  though the actual degree will be influenced by the elasticities inside the loop.

Regulated steps have low flux control coefficients and are therefore not rate-limiting.

What is troubling for many is that the regulated step,  $C_{e_1}^J$  has little in the way of flux control, that is, it is **not** rate-limiting. This would seem to raise a paradox. On the one hand, the regulated step must be important, Yet this importance is not reflected in the degree of influence the step has on the flux. How do we resolve this?



**Figure 12.13** Four Step unbranched Pathway with Negative Feedback.



**Figure 12.12** figure

Switch-over of flux control from the first step to the second step as a function of the feedback strength.

## Regulation and Negative Feedback

Most textbooks and online sites such as Wikipedia refer to phosphofructokinase as the rate-limiting or pacemaker step of glycolysis. There are many reasons why this is considered so. Phosphofructokinase is one of the earliest steps of glycolysis, *in vivo* it is a nonequilibrium reaction and most convincing of all, it is regulated by many effectors. What are the effectors there other than to control glycolytic flux? Many allosteric enzymes such as phosphofructokinase are considered flux controllers by the same reasoning.

From the previous sections we saw that regulated enzymes in unbranched pathways with end-product inhibition are in fact very poor flux controllers. When determined experimentally the flux control coefficient for phosphofructokinase is found to be invariable small [51, 116, 27, 17, 134, 106, 89, 140, 80, 88] and therefore phosphofructokinase is not rate-limiting in many situations. This matches the theoretical expectation even though intuitively it seems suspect.

We therefore have a paradox (sometimes called “The PFK Paradox” [130]). Intuition suggests that phosphofructokinase should be controlling glycolytic flux particularly given the multitude of effectors that regulate it. On the other hand, experimental evidence and theory suggests the opposite. The question is how to reconcile these two opposite views?

## The Phosphofructokinase Paradox

We first restate that phosphofructokinase is not rate-limiting when operating *in situ*. This has been shown experimentally many times [51, 116, 27, 17, 134, 106, 89, 140, 80, 88] as well as being consistent with theory. This cannot be easily disputed. And yet the literature,

textbooks and online resources still claim that phosphofructokinase is rate-limiting [86]. To reconcile this difference we must introduce a different measure that describes the strength of the regulated step and its ability to change flux.

One clear possibility is to look at the loop gain [4]. As described before, the loop gain is the overall gain around the feedback loop. It was shown previously (12.2) that the loop gain is the product of the forward gain,  $A$ , and the feedback gain,  $K$ , that is  $A \cdot K$ . In terms of elasticities the loop gain is the product of the feedback elasticity,  $\varepsilon_{\text{fd}}$ , and the concentration control coefficient,  $C_{e_1}^{x_3}$ , for the *unregulated pathway*. In this case this equals the product of the forward elasticities,  $\varepsilon_1^2 \varepsilon_2^3$  divided by the denominator,  $D$ , of the *unregulated pathway*, equation (12.3):

$$\text{Loop Gain} = -\varepsilon_1^2 \varepsilon_2^3 \varepsilon_{\text{fd}} / D \quad (12.5)$$

We can illustrate the use of this measure with an example. Using the four step pathway with negative feedback, the elasticities are set to the following values. All substrate elasticities ( $\varepsilon_1^2, \varepsilon_2^3, \varepsilon_3^4$ ) are set to 0.5. This corresponds to substrate levels set to the  $K_m$  of each enzyme. The product inhibition elasticities ( $\varepsilon_1^1, \varepsilon_2^2, \varepsilon_3^3$ ) are assumed to be small but not negligible and are set to -0.1. Lastly, the feedback inhibition elasticity,  $\varepsilon^{\text{fd}}$  is set to -4.0. Given these values we can compute the flux control coefficient for the regulated step using the expressions in Table 12.2:

$$C_{e_1}^J = 0.11$$

This tells us that from the perspective of the flux control coefficient, the reaction is not rate-limiting. However we can compute the loop gain using equation (12.5), this yields:

$$\text{Loop Gain} = 6.4$$

Note that the loop gain is significantly higher than the flux control coefficient. It means that changes to the regulator,  $S_3$ , will have a significant effect on throttling the pathway. We can compare these calculations to the same pathway but where the negative feedback strength is weak. If we set  $\varepsilon_{\text{fd}}$  to a small value such as -0.1 we obtain the following values for the flux control and loop gain:

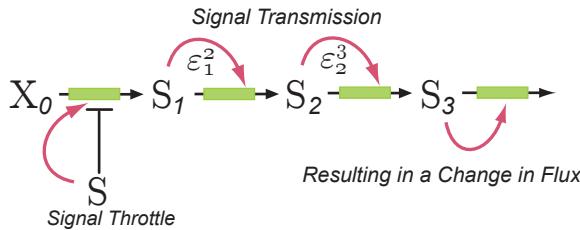
$$C_{e_1}^J = 0.8$$

$$\text{Loop Gain} = 0.16$$

The flux control has increased about eight fold and is now rate-limiting with respect to an external perturbation. What is more interesting is that the loop gain has been reduced forty fold, indicating little or no regulation from the feedback loop. We can access the degree to which a regulated enzyme is actually regulating by looking at the loop gain. The higher the loop gain the stronger the degree of regulation.

It is worth noting that there are two contributions to the loop gain (12.5), the action of the signal on the regulated step,  $\varepsilon_{\text{fd}}$ , and the transmission of that signal to cause a change (Figure 12.14),  $-\varepsilon_1^2 \varepsilon_2^3 / D$ . The effectiveness of the overall regulation is therefore not just a

function of the regulated step but of the entire loop. If the transmission elasticities, in this case  $\varepsilon_1^2$  and  $\varepsilon_2^3$ , are small then the loop gain could be significantly reduced. An examination of the elasticity of the regulated step is therefore insufficient to ascertain the effectiveness of the regulation and it is quite possible that with weak signal transmission, even in the presence of a strongly regulated step, effective regulation could be minimal.



**Figure 12.14** Transmission from Signal to Flux Change.

## 12.5 Robustness and Supply/Demand

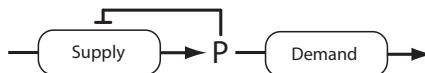
One way to look at metabolic systems is to divide them into two separate but connected blocks, the supply block and the demand block. Negative feedback makes this division very straight forward. Consider a factory that make cars. Should the rate of car production be controlled by the demand or the supply of cars? Economically the most efficient strategy is to let demand decide how many cars to make, this ensure that excess cars do not build up and thereby waste resources. In certain metabolic situations the same reasoning can be used. For example the production of amino acids would best be determined by the demand from protein synthesis. This means that if protein synthesis slows, amino acid production should also slow. A metabolic system that supplies amino acids should be able to supply amino acids effectively at both high and low demand. One way to do this is to maintain the amino acid level at a relatively constant level independent of the demand block. Let us assume for the moment that there is no feedback regulation in the supply/demand pathway (Figure 12.15). If demand rises, this will result in the intermediate metabolite,  $P$ , falling. As  $P$  falls the ability to supply the increased demand becomes more and more difficult. If demand falls, the flux through the pathway will fall. This will cause the intermediate metabolite,  $P$  to rise. Since the equilibrium constant across the supply block is likely to be large, the concentration of  $P$  could raise to toxic high levels as the supply block approach equilibrium at low fluxes.



**Figure 12.15** A system divided into supply and demand blocks.

The solution to avoid both problems at high and low demand is to use negative feedback (Figure 12.16). With negative feedback, high demand will result in a decrease in the inter-

mediate metabolite,  $P$ , which in turn will release repression in the supply block to restore some the loss in  $P$ . Alternatively, at low demand, the increase in  $P$  will suppress its own production preventing excessive production of  $P$ .



**Figure 12.16** A system divided into supply and demand blocks with negative feedback.

## 12.6 Instability

Although negative feedback offers considerable advantages to a system, too much feedback and delays can result instability in the form of sustained oscillations. We will discuss stability of negative feedback systems in more later in the book in Chapter 13.

## Further Reading

1. Hofmeyr, JH and Rohwer, JM (2011) Supply-demand analysis a framework for exploring the regulatory design of metabolism. *Methods Enzymology*, Vol 500, 533-554
2. Sauro HM (2017) Control and regulation of pathways via negative feedback. *J R Soc Interface*. 2017 Feb; 14 (127)

## Exercises

1. A regulated step via a negative feedback look has a flux control coefficient of 0.9. Would you consider this system to be a well regulated pathway? Explain your answer.
2. A coal mine is driven by the supply of coal it finds in the mine. What could go wrong with such a system when compared to one that is driven by demand?



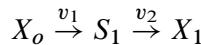
# 13

## Stability

Section 4.7 briefly mentions the concept of stability in a biochemical system. In this chapter we will look at stability in much more detail and how MCA can tell us something about whether a given steady state is stable or not.

In the simplest terms, a stable steady state is one where if a *small* perturbation is made to a species levels, then over time, that perturbation relaxes back to the original steady state. Figure 13.1 show the effect of two perturbations on a system. In both cases the system relaxes back to the original steady state.

To understand the concept of stability more fully, consider the following simple system:



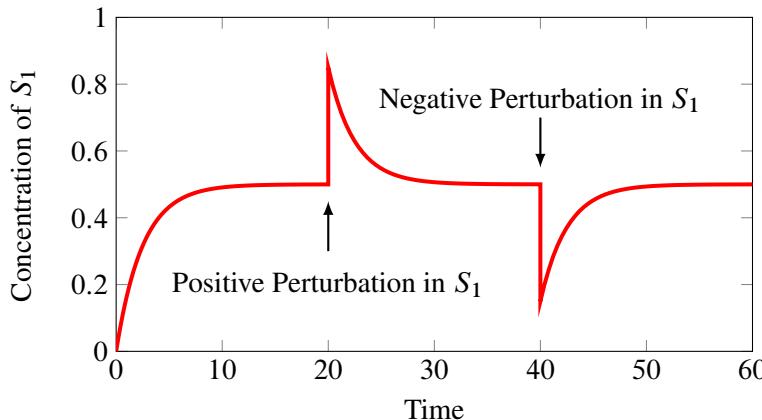
where  $X_o$  and  $X_1$  are fixed species,  $v_1$  is given by the rate law  $k_1 x_o$  and  $v_2$  by  $k_2 s_1$ . The differential equation for the single floating species,  $S_1$ , was given by:

$$\frac{ds_1}{dt} = k_1 x_o - k_2 s_1 \quad (13.1)$$

By setting the differential equation to zero, the steady state solution can be determined to be:

$$s_1 = k_1 x_o / k_2 \quad (13.2)$$

An important question to ask is whether the steady state is stable or not, that is, whether a small perturbation will decay and return to the steady state. To test this we can do the following: When the system is at steady state, make a small perturbation to the steady state



**Figure 13.1** Multiple Perturbations. The steady state concentration of the species  $S_1$  is 0.5 and a perturbation is made to  $S_1$  by adding an additional 0.35 units of  $S_1$  at time = 20 and removing 0.35 units at time = 40. In both cases the system relaxes back.

concentration of  $S_1$ , by an amount  $\delta s_1$  and ask how  $\delta s_1$  subsequently changes as a result of this perturbation. That is, what is  $d(\delta s_1)/dt$ ? To determine this, the new rate of change equation is rewritten as follows:

$$\frac{d(s_1 + \delta s_1)}{dt} = k_1 x_o - k_2(s_1 + \delta s_1)$$

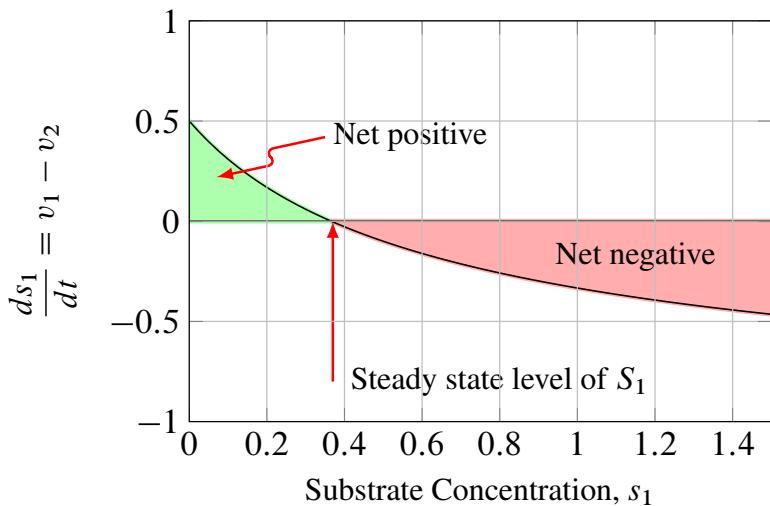
If we insert the steady state solution for  $S_1$ , equation (13.2), into the above equation we are left with:

$$\frac{d\delta s_1}{dt} = -k_2 \delta s_1 \quad (13.3)$$

Note that the rate of change of the *disturbance itself*,  $\delta s_1$ , is negative. The system reduces the disturbance so that the system returns back to the original steady state. Systems with this kind of behavior are called **stable**. If the rate of change in  $S_1$  had been positive instead of negative, the perturbation would have continued to diverge away from the original steady state and the system would then be considered **unstable**.

Let's look at this graphically by plotting the rate of change,  $ds_1/dt$ , as a function of  $s_1$ , shown in Figure 13.2. The steady state occurs when the net rate of change is zero, marked by the straight line arrow. If the substrate level falls below this value, the net rate goes positive, thereby increasing the level of  $S_1$ . If the substrate rises above the steady state level, the graph shows the net rate of change going negative, so that  $S_1$  decreases. The system is therefore stable.

This kind of stability is also called the **internal stability** because it describes the system's stability to perturbations in the internal state. We can informally define a stable system as:



**Figure 13.2** Rate of change as a function of  $s_1$ . The arrow indicates the steady state for  $S_1$ . When  $S_1$  is below the steady state value, the net change is positive meaning that  $S_1$  will increase. When  $S_1$  is above the steady state value, the net change is negative meaning that  $S_1$  will decrease. The system is therefore stable.

**Internal Stability:** A biochemical pathway is internally stable if at steady state, small perturbations to the floating species relax back to the steady state.

**Caveat:** If the perturbed species is part of a conserved cycle (see section 3.8), then the total mass in the cycle must remain constant during the perturbation. This may require perturbing one species in a positive direction and another in a negative direction.

Divide both sides of equation (13.3) by  $\delta s_1$  and taking the limit,  $\delta s_1 \rightarrow 0$ , we find that:

$$\frac{\partial(ds_1/dt)}{\partial s_1} = -k_2$$

The left-hand side is the normalized rate of change of the perturbation, in this case normalized with respect to the steady state concentration of  $S_1$ . The sign of the left-hand side will still tell us whether the perturbation is expanding or contracting. The stability of this simple system can therefore be determined by inspecting the sign of  $\partial(ds_1/dt)/\partial s_1$ . In this case  $\partial(ds_1/dt)/\partial s_1 = -k_2$  is negative, meaning the system is *stable*. It is worth noting that the larger the rate constant,  $k_2$ , the quicker the system relaxes back to steady state.

For systems with more than one species, a system's stability can be determined by looking at all the terms  $\partial(ds_i/dt)/\partial s_j$  which are given collectively by the matrix expression:

$$\frac{d(d\mathbf{s}/dt)}{d\mathbf{s}} = \mathbf{J} \quad (13.4)$$

where  $\mathbf{J}$  is called the **Jacobian matrix** containing elements of the form  $\partial(ds_i/dt)/\partial s_j$ . Equation (13.3) can also be generalized to:

$$\frac{d(\delta \mathbf{s})}{dt} = \mathbf{J} \delta \mathbf{s} \quad (13.5)$$

where  $\mathbf{J}$  is given by:

$$\begin{bmatrix} \frac{\partial(ds_1/dt)}{\partial s_1} & \dots & \frac{\partial(ds_1/dt)}{\partial s_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial(ds_m/dt)}{\partial s_1} & \dots & \frac{\partial(ds_m/dt)}{\partial s_m} \end{bmatrix}$$

Equation (13.5) is an example of an *unforced* linear differential equation and has the general form:

$$\frac{dx}{dt} = Ax$$

Solutions to unforced linear differential equations are well known and take the form:

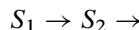
$$x_j(t) = c_1 \mathbf{K}_1 e^{\lambda_1 t} + c_2 \mathbf{K}_2 e^{\lambda_2 t} + \dots + c_n \mathbf{K}_n e^{\lambda_n t}$$

The solution involves a sum of exponentials,  $e^{\lambda_i t}$ , constants  $c_i$ , and vectors,  $\mathbf{K}_i$ . The exponents of the exponentials are given by the eigenvalues (see Appendix B.9) of the matrix,  $A$ , and  $\mathbf{K}_i$ , the corresponding eigenvectors. The  $c_i$  terms are related to the initial conditions assigned to the problem. It is possible for the eigenvalues to be complex, but in general if the real parts of the eigenvalues are negative, the exponents will decay. If they are positive, the exponents will grow. We can therefore determine the stability properties of a given model by computing the eigenvalues of the Jacobian matrix and looking for any positive eigenvalues. Note that the elements of the Jacobian matrix will often be a function of the species levels; it is therefore important that the Jacobian be evaluated at the steady state of interest.

### Example 13.1

---

Assume that following system:



is governed by the set of differential equations:

$$\frac{ds_1}{dt} = -2s_1$$

$$\frac{ds_2}{dt} = 2s_1 - 4s_2$$

If we assume that at time zero, the concentrations of  $s_1$  and  $s_2$  equal 1 and 2 respectively, then the solution to this system can be derived using Mathematica or by using standard algebraic methods for solving linear homogeneous systems. The solution can be found to be:

$$\begin{pmatrix} s_1 \\ s_2 \end{pmatrix} = c_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{-2t} + c_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} e^{-4t}$$

$$s_1 = c_1 e^{-2t}$$

$$s_2 = c_1 e^{-2t} + c_2 e^{-4t}$$

Since the exponents are all negative (-2, -2 and -4), the system is stable to perturbations in  $S_1$  and  $S_2$ .

We can formally define the internal stability of a biochemical system as follows:

The steady state for the biochemical system:

$$\frac{d\mathbf{s}}{dt} = \mathbf{Nv(s, p)} \quad (13.6)$$

is stable if all the eigenvalues of the system's Jacobian matrix have negative real parts. The system is unstable if at least one of the eigenvalues has a positive real part.

There are many software packages that compute the eigenvalues of a matrix, and there are a small number of packages that can compute the Jacobian directly from a biochemical model. For example, the script below is taken from Tellurium. It defines a simple model, initializes the model values, computes the steady state, and then prints out the eigenvalues of the Jacobian matrix (Listing 13.1). For a simple one variable model, the Jacobian matrix only has a single entry and the eigenvalue corresponds to that entry. The output from running the script is given below, showing that the eigenvalue is  $-0.3$ . Since we have a negative eigenvalue, the pathway must be stable to perturbations in  $S_1$ .

```
import tellurium as te

rr = te.loada '''
$Xo -> S1; k1*Xo;
S1 -> $X1; k2*S1;

// Set up the model initial conditions
Xo = 1; X1 = 0;
k1 = 0.2; k2 = 0.3;
''')

# Evaluation of the steady state
```

```

rr.getSteadyStateValues()

# print the eigenvalues of the Jacobian matrix
print rr.getEigenvalues()

# Output follows:
[[-0.3  0. ]]

```

**Listing 13.1** Computing eigenvalues to determine stability.

---

**Example 13.2**

The following system:

$$\rightarrow S_1 \rightarrow S_2 \rightarrow$$

is governed by the set of differential equations:

$$\frac{ds_1}{dt} = 3 - 2s_1$$

$$\frac{ds_2}{dt} = 2s_1 - 4s_2$$

The Jacobian matrix is computed by differentiating the equations with respect to the steady state values of  $S_1$  and  $S_2$ :

$$\mathbf{J} = \begin{bmatrix} -2 & 0 \\ 2 & -4 \end{bmatrix}$$

The eigenvalues for this matrix are:  $-2$  and  $-4$ , respectively. Since both eigenvalues are negative, the system is stable to small perturbations in  $S_1$  and  $S_2$ .

---

**Example 13.3**

Consider the system:

$$\begin{aligned}
X_o &\rightarrow S1 \quad v_o \\
S1 &\rightarrow X1 \quad s_1 \\
S1 &\rightarrow S2 \quad s_1(1 + s_2^3) \\
S2 &\rightarrow X2 \quad 5s_2
\end{aligned}$$

where  $X_o$ ,  $X_1$  and  $X_2$  are fixed species. At steady state  $s_1 = 2.295$  and  $s_2 = 1.14$  with parameter values  $v_o = 8$ . Determine whether this steady state is stable or not. The differential equations for the system are given by:

$$\frac{ds_1}{dt} = v_o - s_1 - s_1(1 + s_2^3)$$

$$\frac{ds_2}{dt} = s_1(1 + s_2^q) - 5s_2$$

The Jacobian matrix is computed by differentiating the equations with respect to the steady state values of  $S_1$  and  $S_2$ :

$$\mathbf{J} = \begin{bmatrix} -2 - s_2^3 & -3s_1s_2^2 \\ 1 + s^3 & -5 + 3s_1s_2^2 \end{bmatrix} = \begin{bmatrix} -3.4815 & -8.948 \\ 2.482 & 3.948 \end{bmatrix}$$

The eigenvalues for this matrix are:  $0.2333 + 2.9i$  and  $0.2332 - 2.9i$ , respectively. Since the real parts of the eigenvalues are positive, the system is unstable to small perturbations in  $S_1$  and  $S_2$ . The eigenvalues are also complex, meaning they have imaginary terms ( $2.9i$ ). This indicates that the system displays some kind of periodic behavior.

---

The pattern of eigenvalues tells us a lot about stability, but also about the kind of transients that occur after a perturbation. The following sections will investigate this subject further.

## 13.1 Jacobian for Biochemical Systems

---

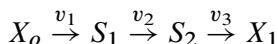
For a given set of differential equations, we can compute the Jacobian by differentiating the equations with respect to the model variables. However for biochemical networks, the Jacobian can be written in a special way that highlights the importance of the network structure and kinetics of the biochemical reaction steps. To do this, recall that the unscaled elasticity (see (6.3)) is given by:

$$\mathcal{E}_s^v = \frac{\partial v}{\partial s}$$

where  $v$  is a reaction rate and  $s$  an effector of the reaction. For example, if  $v = k_1 s$ , the unscaled elasticity,  $\mathcal{E}_s^v = k_1$ . The matrix of unscaled elasticities can be defined as:

$$\frac{\partial \mathbf{v}}{\partial \mathbf{s}} = \begin{bmatrix} \mathcal{E}_{s_1}^{v_1} & \mathcal{E}_{s_2}^{v_1} & \dots & \mathcal{E}_{s_m}^{v_1} \\ \mathcal{E}_{s_1}^{v_2} & \mathcal{E}_{s_2}^{v_2} & \dots & \mathcal{E}_{s_m}^{v_2} \\ \vdots & \vdots & & \vdots \\ \mathcal{E}_{s_1}^{v_n} & \mathcal{E}_{s_2}^{v_3} & \dots & \mathcal{E}_{s_m}^{v_n} \end{bmatrix}$$

where  $n$  is the number of reactions and  $m$  the number of species. An elasticity matrix has  $n$  rows representing  $n$  reactions, and  $m$  columns represent  $m$  species, representing all possible combinations of reactions and species. Many entries in the elasticity matrix will often be zero because not all effectors influence every reaction. For example, consider the pathway:



where  $X_o$  and  $X_1$  are fixed species. The pathway has three reactions, which we will designate  $v_1$ ,  $v_2$ , and  $v_3$  and two floating species,  $S_1$  and  $S_2$ . The unscaled elasticity matrix will

therefore be a three by two matrix. If we assume reversibility or product inhibition in all three reactions, the entries in the matrix will be:

$$\frac{\partial \mathbf{v}}{\partial \mathbf{s}} = \begin{bmatrix} \mathcal{E}_{s_1}^{v_1} & \mathcal{E}_{s_2}^{v_1} \\ \mathcal{E}_{s_1}^{v_2} & \mathcal{E}_{s_2}^{v_2} \\ \mathcal{E}_{s_1}^{v_3} & \mathcal{E}_{s_2}^{v_3} \end{bmatrix} = \begin{bmatrix} \mathcal{E}_{s_1}^{v_1} & 0 \\ \mathcal{E}_{s_1}^{v_2} & \mathcal{E}_{s_2}^{v_2} \\ 0 & \mathcal{E}_{s_2}^{v_3} \end{bmatrix}$$

Note that the entries  $\mathcal{E}_{s_2}^{v_1}$  and  $\mathcal{E}_{s_1}^{v_3}$  are zero because  $S_2$  has no direct effect on  $v_1$ , and  $S_1$  has no direct effect on  $v_3$ . Some of the unscaled elasticities will also be negative. For example,  $\mathcal{E}_{s_2}^{v_2}$  will be negative because increases in  $S_2$  will slow down the  $v_2$  reaction rate due to product inhibition.

Recall that an element of the Jacobian is defined as:

$$\frac{\partial(ds_j/dt)}{\partial s_j}$$

that is the differential equation differentiated with respect to a species. However, we also know that the vector of rates of change is given by the system equation:

$$\frac{ds}{dt} = \mathbf{N}\mathbf{v}(\mathbf{s}, \mathbf{p})$$

Differentiating this with respect to  $\mathbf{s}$  yields<sup>1</sup>:

$$\frac{\partial}{\partial \mathbf{s}} \left( \frac{ds}{dt} \right) = \mathbf{N} \frac{\partial \mathbf{v}}{\partial \mathbf{s}}$$

The left-hand side is the Jacobian, hence the Jacobian is the product of the stoichiometry matrix and the unscaled elasticity matrix:

$$\mathbf{J} = \mathbf{N} \frac{\partial \mathbf{v}}{\partial \mathbf{s}} \quad (13.7)$$

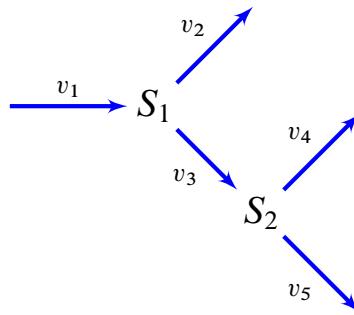
This is a very important result. Given that stability is determined from the Jacobian, this result indicates that stability is a function of network topology **and** the kinetics of the individual reactions. The result indicates that it is not always possible to discern the functional dynamics of a motif just from the topological pattern. The dynamics also depends on the kinetics of the constituent parts.

The dynamics of a network is a function of the network topology **and** the kinetics of its constituent parts.

<sup>1</sup>This is modified if there are conserved cycled in the pathway.

**Example 13.4**

Consider following branched system:



The stoichiometry matrix  $\mathbf{N}$ , and the matrix of unscaled elasticities,  $\partial \mathbf{v} / \partial \mathbf{s}$ , are given by:

$$\mathbf{N} = \begin{bmatrix} v_1 & v_2 & v_3 & v_4 & v_5 \\ 1 & -1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 & -1 \end{bmatrix} \begin{matrix} S_1 \\ S_2 \end{matrix} \quad \frac{\partial \mathbf{v}}{\partial \mathbf{s}} = \begin{bmatrix} \mathcal{E}_{s_1}^{v_1} & 0 \\ \mathcal{E}_{s_1}^{v_2} & 0 \\ \mathcal{E}_{s_1}^{v_3} & \mathcal{E}_{s_2}^{v_3} \\ 0 & \mathcal{E}_{s_2}^{v_4} \\ 0 & \mathcal{E}_{s_2}^{v_5} \end{bmatrix}$$

The Jacobian matrix can be determined from the product,  $\mathbf{N} \partial \mathbf{v} / \partial \mathbf{s}$  which yields:

$$\mathbf{N} \frac{\partial \mathbf{v}}{\partial \mathbf{s}} = \begin{bmatrix} v_1 & v_2 & v_3 & v_4 & v_5 \\ 1 & -1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} \mathcal{E}_{s_1}^{v_1} & 0 \\ \mathcal{E}_{s_1}^{v_2} & 0 \\ \mathcal{E}_{s_1}^{v_3} & \mathcal{E}_{s_2}^{v_3} \\ 0 & \mathcal{E}_{s_2}^{v_4} \\ 0 & \mathcal{E}_{s_2}^{v_5} \end{bmatrix} = \begin{bmatrix} \mathcal{E}_{s_1}^{v_1} - \mathcal{E}_{s_1}^{v_2} - \mathcal{E}_{s_1}^{v_3} & -\mathcal{E}_{s_2}^{v_3} \\ \mathcal{E}_{s_1}^{v_3} & \mathcal{E}_{s_2}^{v_3} - \mathcal{E}_{s_2}^{v_4} - \mathcal{E}_{s_2}^{v_5} \end{bmatrix}$$

## 13.2 External Stability

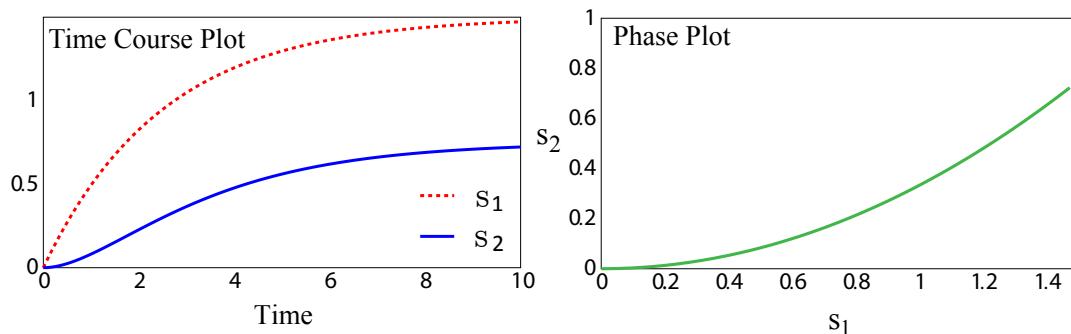
There is one other type of stability that is useful with respect to biochemical systems, namely **external stability**. This refers to the idea that if a system is externally stable, then a finite change to an input of the system should elicit a finite change to the internal state of the system. In control theory this is called BIBO, or Bounded Input Bounded Output stability.

It is very important to bear in mind that the finite change in the internal state refers to a linearized system. When the system is nonlinear, the output may be bounded by physical constraints.

A system that is internally unstable will also be unstable to changes in the systems inputs. In biochemical systems such inputs could be the boundary species that feed a pathway, a drug intervention, or the total mass of a conserved cycle. External stability can be determined using the same criteria used for internal stability, that is the real parts of eigenvalues of the Jacobian matrix should all be negative.

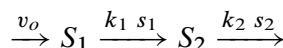
### 13.3 Phase Portraits

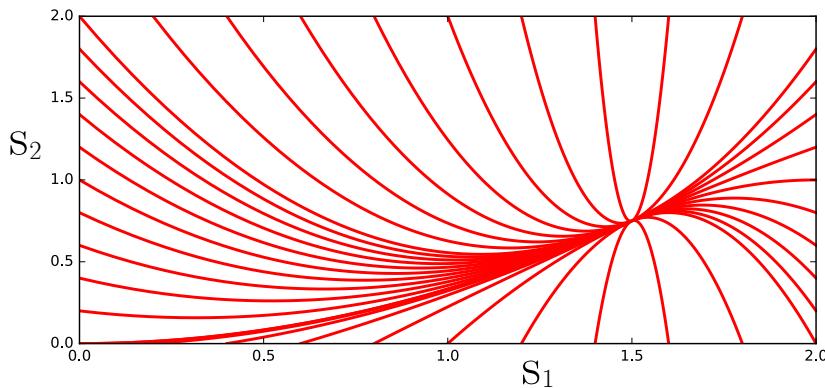
The word **phase space** refers to a space where all possible states are shown. For example, in a biochemical pathway with two species,  $S_1$  and  $S_2$ , the phase space consists of all possible trajectories of  $S_1$  and  $S_2$  in time. For a two dimensional system with species  $S_1$  and  $S_2$ , the phase space can be conveniently displayed with  $S_1$  on one axis and  $S_2$  on the other. A line on a two dimensional plane will represent how  $S_1$  and  $S_2$  move with respect to each other in time. Figure 13.3 shows a time course plot for a simple three-step pathway with two species and the corresponding trajectory in the phase plot. In a real phase plot we would have many trajectories shown rather than just one, each trajectory generated from a different set of initial conditions. Figure 13.4 shows the same phase plot but this time with forty-four trajectories. Note they all converge on a single point that represents the system's steady state.



**Figure 13.3** Time course simulation plot and corresponding phase plot.

A visual representation of the phase space is often called a **phase portrait** or **phase plane**. To illustrate a phase portrait consider the following simple reaction network:





**Figure 13.4** Multiple trajectories plotted on the phase plot. Each trajectory was started at a different initial condition for  $S_1$  and  $S_2$ . See Tellurium Listing: 13.5.

with two linear differential equations:

$$\frac{ds_1}{dt} = v_o - k_1 s_1$$

$$\frac{ds_2}{dt} = k_1 s_1 - k_2 s_2$$

We can assign particular values to the parameters, set up some initial conditions, and plot the evolution of  $S_1$  and  $S_2$  in phase space. If we replot the solution using many different initial conditions, we get something that looks like the plots shown in Figures 13.4 to 13.10.

The plots illustrate a variety of transient behaviors around the steady state. These particular transient behaviors apply specifically to linear differential equations. If we have a nonlinear system and we linearize the system around the steady state, the linearized system will also behave in a way shown by these plots.

Consider the general two dimensional linear set of differential equations:

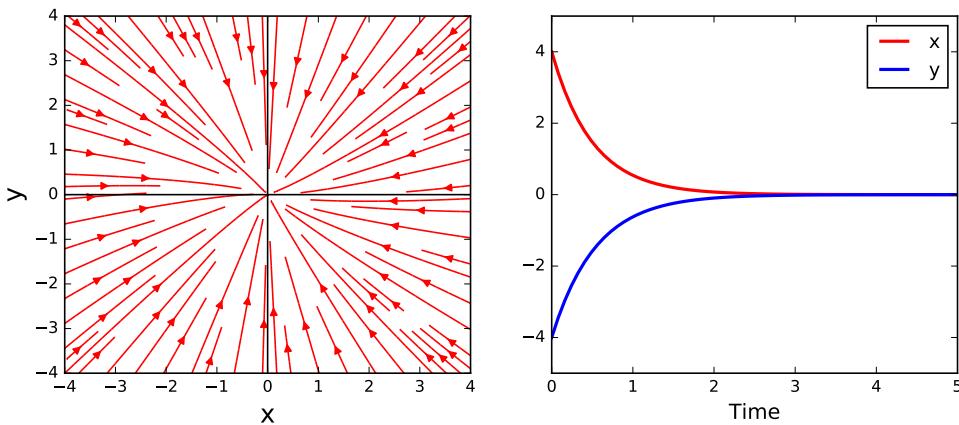
$$\frac{ds_1}{dt} = a_{11}s_1 + a_{12}s_2$$

$$\frac{ds_2}{dt} = a_{21}s_1 + a_{22}s_2$$

As we've already seen, a two dimensional linear system of differential equations has solutions of the form:

$$s_1 = c_1 k_1 e^{\lambda_1 t} + c_2 k_2 e^{\lambda_2 t}$$

$$s_2 = c_3 k_3 e^{\lambda_3 t} + c_4 k_4 e^{\lambda_4 t}$$



**Figure 13.5** Trajectories for a two species reaction network. On the left is the phase plot and on the right, a single transient as a function of time. This system illustrates a stable node corresponding to *Negative Eigenvalues* in the Jacobian. Jacobian:  $a_{11} = -2, a_{12} = 0, a_{21} = -0.15, a_{22} = -2$ . Corresponding eigenvalues:  $\lambda_1 = -2, \lambda_2 = -2$ . The symmetry in the trajectories is due to eigenvalues of the same magnitude. All phase portraits were made using script 13.6. This can be modified to accommodate the different plots by changing the coefficients in the differential equations.

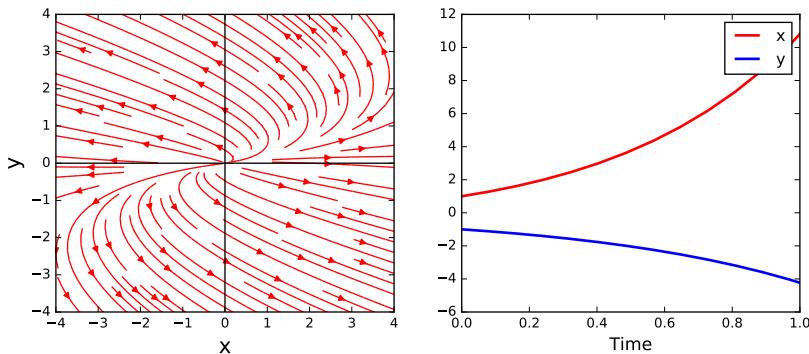
That is, a sum of exponential terms. The  $c_i$  and  $k_i$  terms are constants related to the initial conditions and eigenvectors, respectively, but the  $\lambda_i$  terms or eigenvalues determine the qualitative pattern that a given behavior might exhibit. Note that the eigenvalues can be complex or real numbers. In applied mathematics,  $e$  raised to a complex number immediately suggests some kind of periodic behavior. Let us now consider different possibilities for the eigenvalues<sup>2</sup>.

- **Both Eigenvalues have the same sign, different magnitude but are real.** If both eigenvalues are negative, the equations describe a system known as a **stable node**. All trajectories move towards the steady state point. If the eigenvalues have the same magnitude and the  $c_i$  terms have the same magnitude, the trajectories move to the steady state in a symmetric manner as shown in Figure 13.5.

If the two eigenvalues are both positive, the trajectories move out from the steady state reflecting the fact that the system is unstable. Such a point is called an **unstable node**. If the two eigenvalues have different magnitudes but are still positive, the trajectories twist as shown in Figure 13.6.

- **Real Eigenvalues but of opposite sign.** If the two eigenvalues are real but of opposite sign, we see behavior called a **saddle-node** shown in Figure 13.7. This is where the trajectories move towards the steady state in one direction, called the stable manifold, and form a stable

<sup>2</sup>See Appendix B.9 for a Math update on eigenvalues



**Figure 13.6** Trajectories for a two species reaction network. On the left is the phase plot and on the right a single transient as a function of time. This system illustrates an unstable node, also called an improper node corresponding to *Positive Eigenvalues*. Jacobian:  $a_{11} = 1.2, a_{12} = -2, a_{21} = -0.05, a_{22} = 1.35$ . Corresponding eigenvalues:  $\lambda_1 = 1.6, \lambda_2 = 0.95$ .

ridge. In all other directions trajectories move away, resulting in an unstable manifold. Since trajectories can only move towards the steady state if they are exactly on the stable ridge, saddle nodes are generally considered unstable.

Description	Eigenvalues	Behavior
Both Positive	$r_1 > r_2 > 0$	Unstable
Both Negative	$r_1 < r_2 < 0$	Stable
Positive and Negative	$r_1 < 0 < r_2$	Saddle point
Complex Conjugate	$r_1 > r_2 > 0$	Unstable spiral
Complex Conjugate	$r_1 < r_2 < 0$	Stable spiral
Pure Imaginary	$r_1 = r_2 = 0$	Center

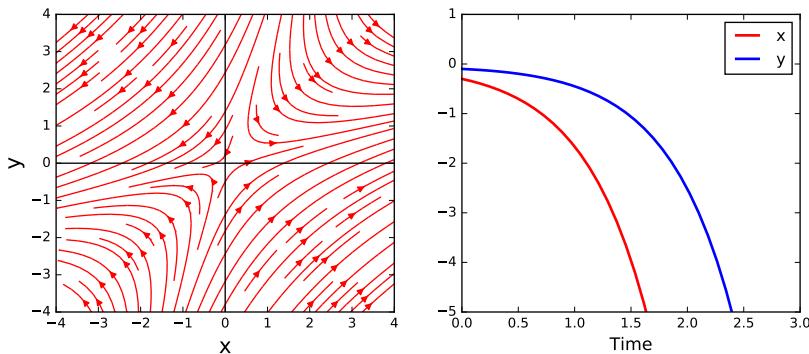
**Table 13.1** Summary of Node Behaviors.

- **Complex Eigenvalues.** Sometimes the eigenvalues can be complex, that is of the form  $a+ib$  where  $i$  is the imaginary number. It may seem strange that the solution to a differential equation that describes a physical system can admit complex eigenvalues. To understand what this means, we must recall Euler's formula:

$$e^{i\theta} = \cos(\theta) + i \sin(\theta)$$

which can be extended to:

$$e^{(a+bi)t} = e^{at} \cos(bt) + i e^{bt} \sin(bt) \quad (13.8)$$

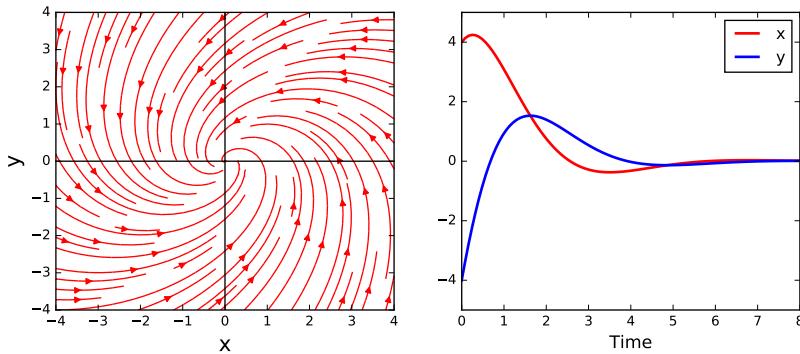


**Figure 13.7** Trajectories for a two species reaction network. On the left is the phase plot and on the right a single transient as a function of time. This system illustrates a saddle node corresponding to *One Positive and One Negative Eigenvalue*. Jacobian:  $a_{11} = 2, a_{12} = -1, a_{21} = 1, a_{22} = -2$ . Corresponding eigenvalues:  $\lambda_1 = -1.73, \lambda_2 = 1.73$ .

When the solutions are expressed in sums of sine and cosine terms, the imaginary parts cancel out, leaving just trigonometric terms with real parts (the proof is provided at the end of the chapter). This means that systems with complex eigenvalues show periodic behavior.

Figures 13.8, 13.9, and 13.10 show typical trajectories when the system admits complex eigenvalues. If the real parts are positive, the spiral trajectories move outwards away from the steady state. Such systems are unstable. In a pure linear system, the trajectories will expand out forever. They will only stop and converge to a stable oscillation if the system has nonlinear elements which limits the expansion. In these cases we observe limit cycle behavior.

If the real parts of the eigenvalues are negative, the spiral trajectory moves into the steady state and is therefore considered stable.



**Figure 13.8** Trajectories for a two species reaction network. On the left is the phase plot, and on the right, a single transient as a function of time. This system illustrates an stable spiral node corresponding to *Negative Complex Eigenvalues*. Jacobian:  $a_{11} = -0.5, a_{12} = -1, a_{21} = 1, a_{22} = -1$ . Corresponding eigenvalues:  $\lambda_1 = -0.75 + 0.97i, \lambda_2 = -0.75 - 0.97i$ .

### Conjugate Pair

A complex conjugate pair is a complex number of the form:  $a \pm bi$ . The eigenvalues for a two variable linear system with matrix  $\mathbf{A}$  can be computed directly using the relation:

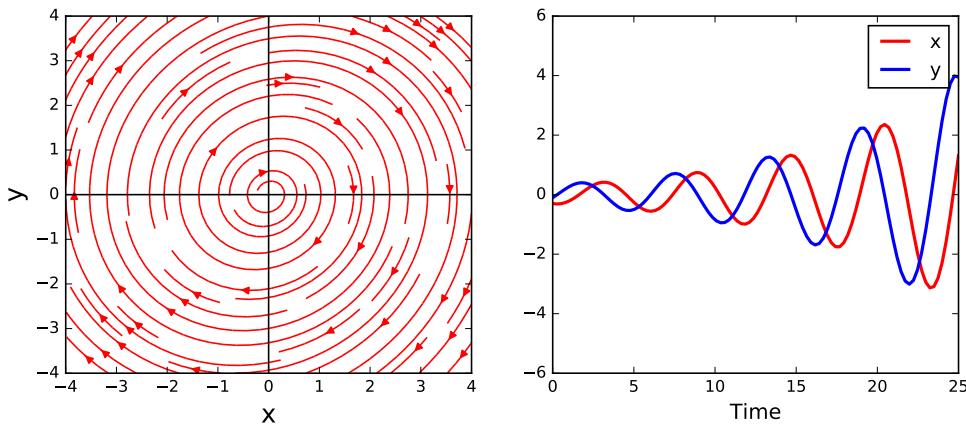
$$\lambda = \frac{\text{tr}(\mathbf{A}) \pm \sqrt{\text{tr}^2(\mathbf{A}) - 4 \det(\mathbf{A})}}{2}$$

where  $\text{tr}(\mathbf{A}) = a + d$ , and  $\det(\mathbf{A}) = ad - bc$ . If the term in the square root is negative, the eigenvalues will always come out as a conjugate pair owing to the  $\pm$  term. If  $\text{tr}^2(\mathbf{A}) - 4 \det(\mathbf{A}) < 0$ , then the solution will be the conjugate pair:

$$\lambda = \frac{\text{tr}(\mathbf{A})}{2} \pm \frac{\sqrt{\text{tr}^2(\mathbf{A}) - 4 \det(\mathbf{A})}}{2}$$

Therefore a complex eigenvalue will always be accompanied by its conjugate partner.

- **Imaginary Eigenvalues with Zero Real Parts.** It is possible for the pair of eigenvalues to have no real component but retain an imaginary part. In this situation the behavior is called a center. This is where the trajectory orbits the steady state. The oscillation is an unusual one in the sense that it implies zero dampening in the system, in other words, zero energy loss. Such a situation is rare in biology and even in non-living systems, such behavior tends to be idealized. For example, a pendulum in a vacuum with no friction at the fulcrum. The other unusual aspect of a center is that the oscillation is depending on the starting condition. Again, we can relate this to a pendulum where the swing depends on how much force we initially apply to the pendulum. Biological oscillators are invariably energy dependent and



**Figure 13.9** Phase portrait for the two species reaction network. Unstable spiral node. **Positive Complex Eigenvalues.** Jacobian:  $a_{11} = 0, a_{12} = 1.0, a_{21} = -1.2, a_{22} = 0.2$ . Corresponding eigenvalues:  $\lambda_1 = 0.1 + 1.09i, \lambda_2 = 0.1 - 1.09i$ .

the frequency is independent of the initial conditions. Biological oscillators therefore tend not to be center types.

## 13.4 Bifurcation Plots

In its simplest form, a bifurcation plot is just a plot of the steady state value of a system variable, such as a concentration or flux versus a parameter of the system. For example, we know that the steady state solution for the simple system:

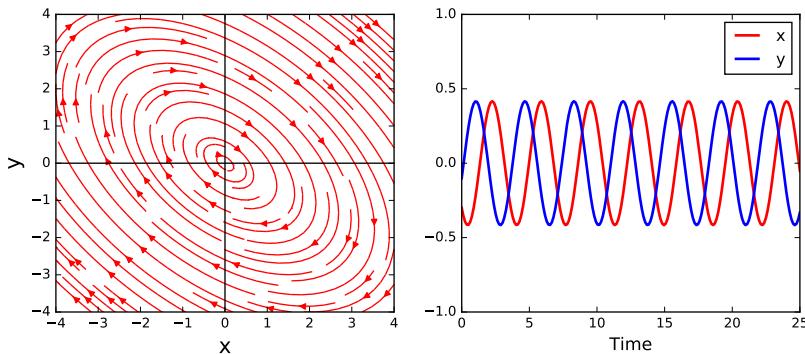
$$\frac{ds_1}{dt} = k_1 x_o - k_2 s_1 \quad (13.9)$$

was given by:

$$s_1 = k_1 x_o / k_2 \quad (13.10)$$

We can plot the steady state value of  $S_1$  as a function of  $k_2$  as shown in Figure 13.12. This isn't a particularly interesting bifurcation plot however, and misses one of the most important characteristics.

Equation (13.10) shows that the simple system (13.9) only has one steady state for a given set of parameters. That is, if we set values to  $X_o, k_1$ , and  $k_2$ , we find there is only *one value* of  $S_1$  that satisfies these parameter settings. This is what Figure 13.12 also demonstrates. What is more interesting is when a system admits multiple steady state values for a given set of parameter values. To illustrate this behavior, let us look at a common system that can admit three possible steady states. It is in these cases that bifurcation plots become particularly useful and more interesting.



**Figure 13.10** Phase portrait for the two species reaction network. Center node. **Complex Eigenvalues, Zero Real Part.** Jacobian:  $a_{11} = 1, a_{12} = 2.0, a_{21} = -2, a_{22} = -1$ . Corresponding eigenvalues:  $\lambda_1 = 0 + 1.76i, \lambda_2 = 0 - 1.76i$ . See Tellurium script:13.6

## Bistable Systems

Bifurcation plots can be useful for identifying changes in qualitative behavior, particularly for systems that have multiple steady states. Consider the system shown in Figure 13.13 where the species  $S$  stimulates its own production forming a positive feedback loop. As species  $S$  accumulates, the rate of formation of  $S$  increases.

At first glance this would seem to be a very unstable situation. One might imagine that the concentration of  $S$  would continue to increase without limit. However, physical constraints ultimately limits the upper value for the concentration of  $S$ . To investigate the properties of this networks, we will construct a simple model. This model uses the following kinetic laws for the synthesis and degradation steps:

$$v_1 = b + k_1 \frac{x^4}{k_2 + x^4}$$

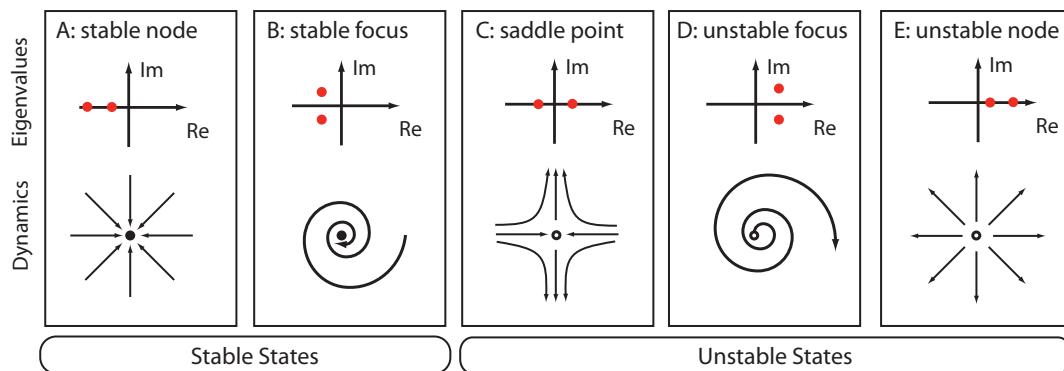
$$v_2 = k_3 x$$

$v_1$  is a Hill like equation with a Hill coefficient of four and a basal rate of  $b$ .  $v_2$  is a simple irreversible mass-action rate law. The differential equation for the model is:

$$\frac{dx}{dt} = v_1 - v_2$$

To find the steady state, set the differential equation to zero and attempt to analytically solve for  $x$ . If we try this we get a solution that is complex to interpret.

A better way to understand what is going on, is to plot both rate laws as a function of the species  $x$ . When we do this we obtain Figure 13.14. The intersection points marked with filled circles indicate the steady state solutions because at these points,  $v_1 = v_2$ . If we vary



**Figure 13.11** Summary of behaviors including dynamics and associated eigenvalues for a two dimensional linear system. Adapted from “Computational Models of Metabolism: Stability and Regulation in Metabolic Networks”, Adv in Chem Phys, Vol 142, Steuer and Junker.

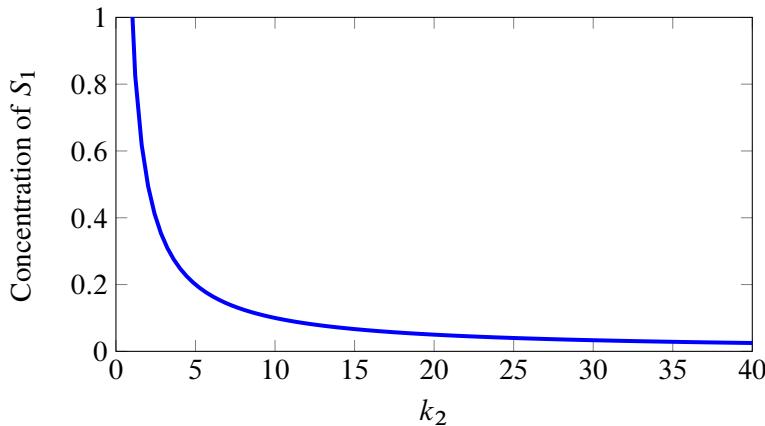
the slope of  $v_2$  by changing  $k_3$ , the intersection points will change (Figure 13.15). At a high  $k_3$  value, only one intersection point remains (Panel C), the low intersection point. If the value of  $k_3$  is low, only the high intersection point remains (Panel A). However, with the right set of parameter values, we can make a system with three steady state values (Panel B).

We can determine the three different steady state stabilities by doing a simple graphical analysis on Figure 13.14. Figure 13.16 shows the same plot but with perturbations.

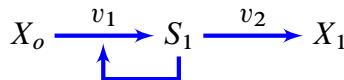
Starting with the first steady state in the low left corner of Figure 13.16, consider a perturbation made in  $x$ ,  $\delta s$ . This means that both  $v_1$  and  $v_2$  increase, however  $v_2 > v_1$  meaning that after the perturbation, the rate of change in  $s$  is *negative*. Since it is negative, this restores the perturbation back to the steady state. The same logic applies to the upper steady state. This tells us that the lower and upper steady states are both stable.

What about the middle steady state? Consider again a perturbation,  $\delta s$ . This time  $v_1 > v_2$  which means that the rate of change of  $s$  is *positive*. Since it is positive, the perturbation, instead of falling back, continues to grow until  $s$  reaches the upper steady state. We conclude that the middle steady state is therefore unstable. This system possess three steady states, one unstable and two stable. Such a system is known as a **bistable system** because it can rest in one of two stable states but not the third.

Another way to observe the different steady states is to run a time course simulation at many different starting points. Figure 13.17 shows the plots generated using the script in Listing 13.2. The plots show two steady states, a high state at around 40, and a low state at around 3. Notice that there is no third state observed. As we have already discussed, the middle steady state is unstable, and all trajectories diverge from this point. It is therefore not possible, when doing a time course simulation, to observe an unstable steady state since



**Figure 13.12** Steady state concentration of  $S_1$  as a function of  $k_2$  for the system,  $ds_1/dt = k_1 c_o - k_2 s_1$ .



**Figure 13.13** System with Positive Feedback.

there is no way to reach it.<sup>3</sup>

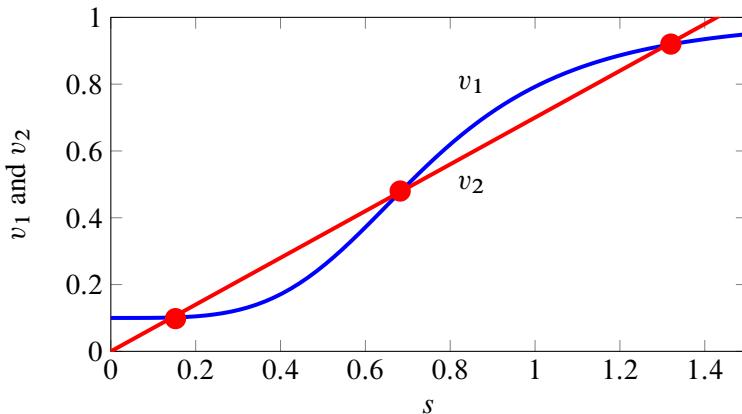
```
import tellurium as te
import numpy

rr = te.loada '''
J1: $Xo -> x; 0.1 + k1*x^4/(k2+x^4);
x -> $w; k3*x;

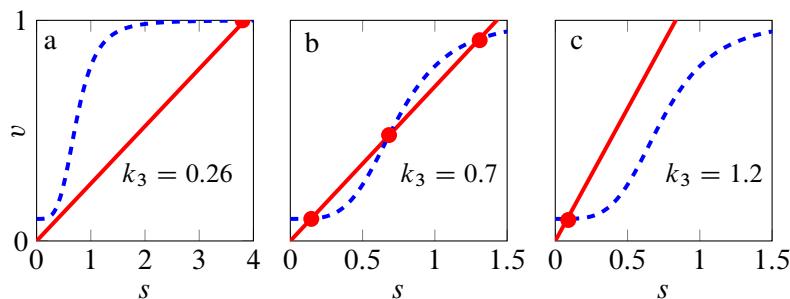
k1 = 0.9;
k2 = 0.3;
k3 = 0.7;
x = 0.05;
'''

m = rr.simulate(0, 15, 100)
for i in range(1, 10):
    rr.x = i*0.2
    mm = rr.simulate(0, 15, 100, ["x"])
    m = numpy.hstack((m, mm))
te.plotArray(m)
```

<sup>3</sup>The author has been reliably informed that running time backwards in a time course simulation will cause the simulation to converge on the unstable steady state. The author has not tried this himself, however.



**Figure 13.14** Reaction velocities,  $v_1$  and  $v_2$ , as a function of  $s$  for the system in Figure 13.12. The intersection points marked by full circles indicate possible steady states.  $k_1 = 0.9; k_2 = 0.3; k_3 = 0.7; b = 0.1$ .

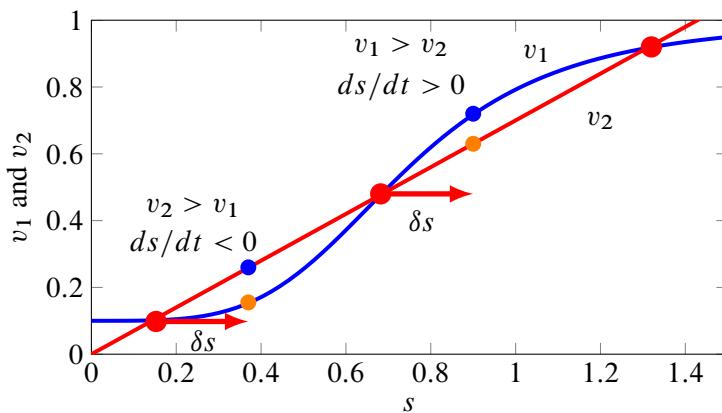


**Figure 13.15**  $v_1$  and  $v_2$  plotted against  $s$  concentration. Intersection points on the curves mark the steady state points. Panel a) One intersection point at a high steady state; b) Three steady states; c) One low steady state.

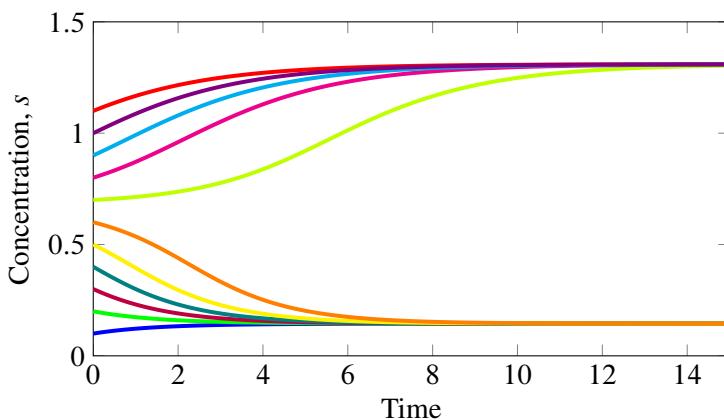
---

**Listing 13.2** Tellurium script used to generate Figure 13.17.

We can get an estimate for the values of all three steady states from Figure 13.14. Reading from the graph we find  $s$  values at 0.145, 0.683, and 1.309. It is also possible to use the steady state solver from Tellurium to locate the steady states. Listing 13.3 shows a simple script to compute them. By setting an appropriate initial condition, we can use Tellurium to pinpoint all three steady states. For example, if we use an initial value of  $s$  at 0.43, the steady state solver will locate the third steady state at 0.683. Steady state solvers such as the one included with Tellurium can be used to find unstable states, providing the initial starting point is close enough.



**Figure 13.16** A graphical understanding of the stability of steady states. See text for details. Computed using the SBW rate law plotter.  $k_1 = 0.9$ ;  $k_2 = 0.3$ ;  $k_3 = 0.7$ ;  $b = 0.1$ .



**Figure 13.17** Time course data generated from Tellurium model 13.2. Each line represents a different initial concentration for  $s$ . Some trajectories transition to the low state while others to the upper state.

```

import tellurium as te

rr = te.loada ('''
$X0 -> x; 0.1 + k1*x^4/(k2+x^4);
x -> $w; k3*x;

// Initialization here
k1 = 0.9; k2 = 0.3;
k3 = 0.7;
''')

# Compute steady state
print rr.getSteadyStateValues()

```

**Listing 13.3** Basic bistable model.

## Stability of Positive Feedback

What determines the stability of a positive feedback system? Let us consider the same network with positive feedback as before (Figure 13.13). The differential equation for this system is:

$$\frac{ds}{dt} = v_1(s) - v_2(s)$$

where we have explicitly shown that each reaction rate is a function of  $s$ . To determine whether the system is stable to small perturbations, we can differentiate the equation with respect to  $s$  to form the Jacobian. Notice there is only one element in the Jacobian because we only have one state variable:

$$\frac{ds/dt}{ds} = \frac{\partial v_1}{\partial s} - \frac{\partial v_2}{\partial s}$$

The terms on the right are unscaled elasticities (6.3). If the expression is positive, the system is unstable because it means that  $ds/dt$  is increasing if we increase  $s$ . We can scale both sides to yield:

$$J_s = \varepsilon_s^1 - \varepsilon_s^2$$

where the right-hand term now includes the scaled elasticities (6.2). The criteria for instability is that  $\varepsilon_s^1 - \varepsilon_s^2 < 0$ , or  $\varepsilon_s^1 > \varepsilon_s^2$ . Therefore, if the positive feedback is stronger than the effect of  $s$  on the degradation step  $v_2$ , the system will be unstable.

Recall that the elasticities are a measure of the kinetic order of the reaction. Thus an elasticity of one means the reaction is first-order. A saturable irreversible Michaelis-Menten reaction will have a variable kinetic order between one and zero (near saturation). A Hill equation can, depending on the Hill coefficient, have kinetic orders greater than one (Table 13.2). Knowing this information, there are at least two ways to make sure that the elasticity for the feedback elasticity,  $v_1$ , is greater than the elasticity for the degradation step,  $v_2$ :

Kinetic Order	Elasticity
First-Order	1.0
Zero-Order	0.0
Sigmoidal	> 1.0

**Table 13.2**

1.  $v_1$  is modeled using a Hill equation with a Hill coefficient  $> 1$  and  $v_2$  is first-order or less.
2. A Hill coefficient = 1 on  $v_1$ , with Michaelis-Menten saturable kinetics on  $v_2$  to ensure less than first-order kinetics on  $v_2$ .

By substituting the three possible steady state values for  $s$  into the equation for  $ds/dt$ , we can compute the value for the Jacobian element in each case (Table 13.3).

Steady State $s$	Jacobian: $(ds/dt)/ds$	Elasticity, $\varepsilon_s^1$
0.145	-0.664	0.052
0.683	0.585	1.835
1.309	-0.47	0.33

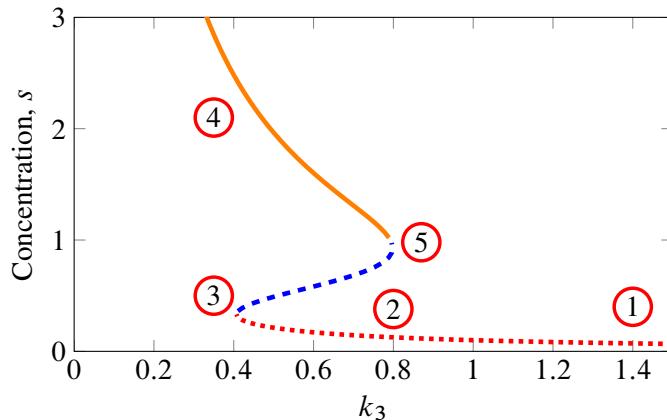
**Table 13.3** Table of steady state values of  $s$  and corresponding values for the Jacobian element. Negative Jacobian values indicate a stable steady state, positive elements indicate an unstable steady state. The table shows one stable and two unstable steady states.

The unstable steady state at  $x = 0.683$  has an elasticity for  $v_1$  of 1.835. Note this value is greater than the elasticity of the first-order degradation reaction,  $v_2$ , which equals one. Therefore this state is unstable.

## Bifurcation Plot

Let's now return to the question of plotting a bifurcation graph for the bistable system in Figure 13.13. Figure 13.15 shows both reaction rates,  $v_1$  and  $v_2$  plotted as a function of the intermediate species,  $S$ . In this figure we see three intersection points, marking the three possible steady states. By varying the degradation constant  $k_3$ , we can change the behavior of the system so that it exhibits a single high steady state, three separate steady states, or a single low steady state (See Figure 13.14).

If we track the intersection points as we vary the value of the rate constant  $k_3$ , we obtain the bifurcation plot shown in Figure 13.18.



**Figure 13.18** Plotting intersection points from Figure 13.14 as a function of  $k_3$ . Dotted line marks the lower intersection point, dashed line the middle intersection points, and solid line the upper intersection point. Computed using the SBW AUTO C# Tool.

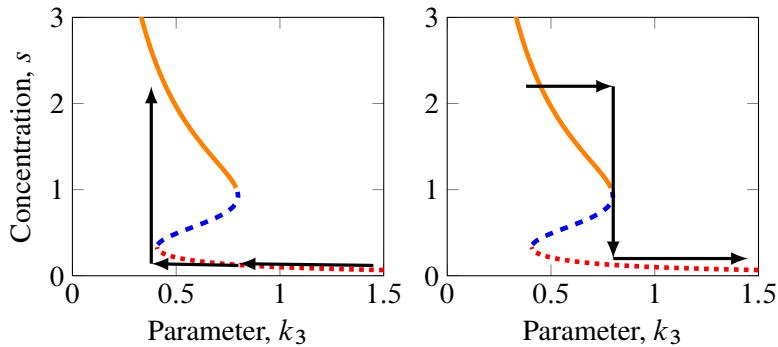
Figure 13.18 shows that at some value of the parameter  $k_3$ , the system has three possible steady states, outside this range only a single steady state persists. Bifurcation diagrams are extremely useful for uncovering and displaying such information. Drawing bifurcation diagrams is not easy, however, and there are some software tools that can help. Figure 13.18 for example was generated using the SBW Auto C# tool<sup>4</sup>. Another useful tool for drawing bifurcation diagrams is Oscill8<sup>5</sup>. Both tools can read SBML. Figure 13.18 was generated first by entering the model into Tellurium (Shown in Listing 13.3) to generate the SBML. The model was then passed to Auto C# to produce the bifurcation diagram.

The bifurcation plot shows how the steady state changes as a function of a parameter, in this case  $k_3$ . Of interest is the following observation. If we start  $k_3$  at a high value of 1.4 (Marker 1), we see that there is only one low steady state. As  $k_3$  is lowered, we pass the point at approximately  $k_3 = 0.8$  (Marker 2) where three steady states emerge. We continue lowering  $k_3$ , and see that the concentration of  $s$  rises very slowly until about 0.4 (marker 3). At this point the system jumps to a single steady state, but now at a high level (Marker 4). The interesting observation is that if we now increase the value of  $k_3$ , we do not traverse the same path. As we increase  $k_3$  beyond 0.4, we do not drop back to the low state, but continue along the high state until we reach  $k_3 = 0.8$  (Marker 5), at which point we jump down to the low state (Marker 2). The direction in which we traverse the parameter  $k_3$  affects the type of behavior we observe. This special phenomena is called **hysteresis**, Figure 13.19.

Hysteresis is where the behavior of a system depends on its past history.

<sup>4</sup> [http://jdesigner.sourceforge.net/Site/Auto\\_C.html](http://jdesigner.sourceforge.net/Site/Auto_C.html)

<sup>5</sup> <http://oscill8.sourceforge.net/>



**Figure 13.19** Depending on whether we increase or decrease  $k_3$ , the steady state path we traverse will be different. This is a characteristic of hysteresis.

### Irreversible Bistability

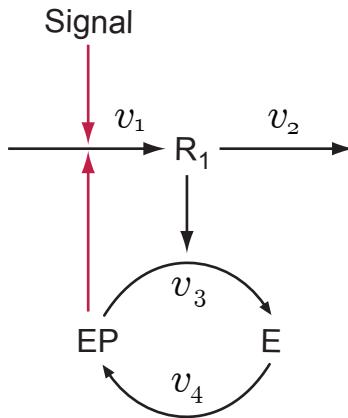
It is possible to design an irreversible bistable switch. Figure 13.21 shows the bifurcation plot for such a system. This is modified from the ‘Mutual activation’ model in the review by Tyson [139], Figure 1e. In this example increasing the signal results in the system switching to the high state at around 2.0. If we reduce the signal from a high level, we traverse the top arc. If we assume the signal can never be negative, we will remain at the high steady state even if the signal is reduced to zero. The bifurcation plot in the negative quadrant of the graph is physically inaccessible. This means it is not possible to transition to the low steady state by decreasing signal. As a result, the bistable system is **irreversible**, that is, once it is switched on, it will always remain on.

```
import tellurium as te

r = te.loada ('''
$X -> R1;    k1*EP + k2*Signal;
R1 -> $w;    k3*R1;
EP -> E;      Vm1*EP / (Km + EP);
E -> EP;     ((Vm2+R1)*E) / (Km + E);

Vm1 = 12; Vm2 = 6;
Km = 0.6;
k1 = 1.6; k2 = 4;
E = 5; EP = 15;
k3 = 3; Signal = 0.1;
''')

result = r.simulate(0, 20, 500)
r.plot()
```

**Listing 13.4** Script for Figure 13.21.**Figure 13.20** System with Positive Feedback using a covalent modification cycle,  $E$ ,  $EP$ .

The Tellurium script for the model is shown in Listing 13.4.

## Further Reading

---

1. Edelstein-Keshet L (2005) Mathematical Model sin Biology. SIAM Classical In Applied Mathematics. ISBN-10: 0-89871-554-7.
2. Fall CP, Marland ES, Wagner JM, Tyson JJ (2000) Computational Cell Biology. Springer: Interdisciplinary Applied Mathematics. ISBN 0-387-95369-8.
3. Steuer R and Junker BH (2009). Computational models of metabolism: stability and regulation in metabolic networks. Advances in chemical physics, 142, 105.

## Exercises

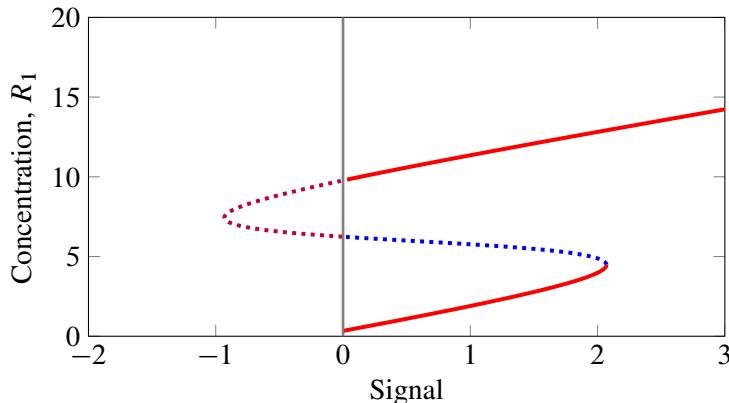
---

1. Determine the Jacobian matrix for the following systems:

$$\text{a) } \frac{dx}{dt} = x^2 - y^2 \quad \frac{dy}{dt} = x(1-y)$$

$$\text{b) } \frac{dx}{dt} = y - xy \quad \frac{dy}{dt} = xy$$

2. Compute the steady state solutions to the two systems in the previous question.

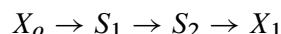


**Figure 13.21** Bifurcation diagram for species  $R_1$  with respect to the signal. Signal from the model shown in Tellurium script 13.4. The continuous line represents stable steady state points, the dotted line the unstable steady states. Plotted using Oscil8 <http://oscil8.sourceforge.net/>.

3. Determine the stability of the solutions from the previous question.
4. Determine the Jacobian in terms of the elasticities and stoichiometry matrix for the following systems:
  - a)  $X_o \rightarrow S_1; S_1 \rightarrow S_2; S_2 \rightarrow S_3; S_3 \rightarrow X_1$
  - b)  $X_o \rightarrow S_1; S_1 \rightarrow S_2; S_2 \rightarrow S_1; S_2 \rightarrow X_1$
  - c)  $S_1 \rightarrow S_2; S_2 \rightarrow S_3$

Assume all reactions are product insensitive,  $X_i$  species are fixed, and in c)  $S_3$  regulates the first step.

5. Show that the following system is stable to perturbations in  $S_1$  and  $S_2$  by computing the eigenvalues at steady state (See Listing 13.1):



The three rate laws are given by:

$$v_1 = \frac{V_{m1} x_o}{Km_1 + x_o + s_1/K_i}$$

$$v_2 = \frac{V_{m2} s_2}{Km_2 + s_1 + s_2/K_j}$$

$$v_2 = \frac{V_{m3} s_3}{Km_3 + s_2}$$

Assign the following values to the parameters:  $x_o = 1; x_1 = 0; V_{m1} = 1.5; V_{m2} = 2.3; V_{m3} = 1.9; K_{m1} = 0.5; K_{m2} = 0.6; K_{m3} = 0.45; K_i = 0.1; K_j = 0.2$ .

6. Show that the following system is unstable or stable depending on the value of  $v_2$ .

- a)  $\rightarrow S_1; v_o$
- b)  $S_1 \rightarrow; v_2 = k_2 s_1$
- c)  $S_1 \rightarrow; v_3 = k_3 \frac{s_1}{1 + (s_1/K_I)^n}$

7. Show that the following system is unstable. What kind of unstable dynamics does it have?

```
import tellurium as te

rr = te.loada """
J0: $X0 -> S1; VM1*(X0-S1/Keq1)/(1+X0+S1+pow(S4,h));
J1: S1 -> S2; (10*S1-2*S2)/(1+S1+S2);
J2: S2 -> S3; (10*S2-2*S3)/(1+S2+S3);
J3: S3 -> S4; (10*S3-2*S4)/(1+S3+S4);
J4: S4 -> $X1; Vm4*S4/(KS4+S4);

X0 = 10;           X1 = 0;
S1 = 0.973182;   S2 = 1.15274;
S3 = 1.22721;    S4 = 1.5635;
VM1 = 10;         Keq1 = 10;
h = 10;           Vm4 = 2.5;
KS4 = 0.5;
"""

```

8. Show that the following system is unstable. What kind of unstable dynamics does it have?

```

import tellurium as te

r = te.loada ('''
// var R, X, E, EP;
// ext src, S, Waste;

J0: $src -> X;      k1*S;
J1: X -> R;          (kop + ko*EP)*X;
J2: R -> $waste;    k2*R;
J3: E -> EP;         Vmax_1*R*E/ (Km_1 + E);
J4: EP -> E;         Vmax_2*EP/ (Km_2 + EP);

src = 0;             kop = 0.01;
ko = 0.4;            k1 = 1;
k2 = 1;              R = 1;
EP = 1;              S = 0.2;
Km_1 = 0.05;         Km_2 = 0.05;
Vmax_2 = 0.3;        Vmax_1 = 1;
KS4 = 0.5;
''')

result = r.simulate(0, 300, 1000)
r.plot()

```

## 13.5 Appendix

---

Prove that the presence of imaginary numbers in the solution to a set of differential equations means that the solution is periodic (Equation (13.8)). Consider the system:

$$x(t) = c_1 z_1 e^{(\lambda+i\mu)t} + c_2 z_2 e^{(\lambda-i\mu)t}$$

where  $z_1$  and  $z_2$  are corresponding conjugate eigenvectors. Using Euler's formula,  $e^{i\mu} = \cos(\mu) + i \sin(\mu)$  and that  $e^{(\lambda+i\mu)t} = e^{\lambda t} e^{i\mu t}$  we obtain:

$$\begin{aligned} x(t) &= c_1 z_1 e^{\lambda t} (\cos(\mu t) + i \sin(\mu t)) \\ &\quad + c_2 z_2 e^{\lambda t} (\cos(\mu t) - i \sin(\mu t)) \end{aligned}$$

Writing the conjugate eigenvectors as  $z_1 = a + bi$  and  $z_2 = a - bi$ , we get:

$$\begin{aligned} x(t) &= c_1(a + bi)e^{\lambda t}(\cos(\mu t) + i \sin(\mu t)) \\ &\quad + (a - bi)e^{\lambda t}(\cos(\mu t) - i \sin(\mu t)) \end{aligned}$$

Multiply out and separate the real and imaginary parts:

$$\begin{aligned} x(t) &= e^{\lambda t} [c_1(a \cos(\mu t) - b \sin(\mu t) + i(a \sin(\mu t) + b \cos(\mu t))) \\ &\quad + c_2(a \cos(\mu t) - b \sin(\mu t) - i(a \sin(\mu t) + b \cos(\mu t)))] \end{aligned}$$

The complex terms cancel leaving only the real parts. If we set  $c_1 + c_2 = k_1$  and  $(c_1 - c_2)i = k_2$  then:

$$x(t) = e^{\lambda t} [k_1(a \cos(\mu t) - b \sin(\mu t)) \\ k_2(a \sin(\mu t) + b \cos(\mu t))]$$

The solution is real when the constants  $c_1$  and  $c_2$  are real. This will only be the case when the eigenvalues are a conjugate pair,  $(a \pm ib)$ , which is the case we are considering. Therefore, systems that admit a complex pair of conjugate eigenvalues result in periodic real solutions.

```
# Plot a phase portrait for a simple two pecies pathway
import tellurium as te
import matplotlib.pyplot as plt

rr = te.loada '''
$Xo -> S1; k1*Xo;
S1 -> S2; k2*S1;
S2 -> $X1; k3*S2;
k1 = 0.6; Xo = 1;
k2 = 0.4; k3 = 0.8;
'''

plt.figure(figsize=(9, 4))
S1Start = 0
S2Start = 0
for i in range(1, 11):
    rr.S1 = S1Start
    rr.S2 = S2Start
    m = rr.simulate(0, 10, 120, ["S1", "S2"])
    p = te.plotArray(m, show=False)
    plt.setp (p, color='r')
    S1Start = S1Start + 0.2
S1Start = 2
S2Start = 0
for i in range(1, 11):
    rr.S1 = S1Start
    rr.S2 = S2Start
    m = rr.simulate(0, 10, 120, ["S1", "S2"])
    p = te.plotArray(m, show=False)
    plt.setp (p, color='r')
    S2Start = S2Start + 0.2
S2Start = 0
S1Start = 0
for i in range(1, 11):
    rr.S1 = S1Start
    rr.S2 = S2Start
    m = rr.simulate(0, 10, 120, ["S1", "S2"])
```

```
p = te.plotArray(m, show=False)
plt.setp (p, color='r')
S2Start = S2Start + 0.2
S1Start = 0
S2Start = 2
for i in range(1, 11):
    rr.S1 = S1Start
    rr.S2 = S2Start
    m = rr.simulate(0, 10, 120, ["S1", "S2"])
    p = te.plotArray(m, show=False)
    plt.setp (p, color='r')
    S1Start = S1Start + 0.2
plt.xlabel ('S1', fontsize=16)
plt.ylabel ('S2', fontsize=16)
plt.savefig ("plot.pdf")
plt.show()
```

**Listing 13.5** Script for Figure 13.4.

```
import numpy as np
import matplotlib.pyplot as plt
import tellurium as te

Y, X = np.mgrid[-4:4:100j, -4:4:100j]

r = te.loada '''
    x' = 1*x + 2*y;
    y' = -2*x - 1*y;

    x = -0.3; y = -0.1;
'''

m = r.simulate (0, 25, 200)

U = 1*X + 2*Y
V = -2*X - 1*Y

plt.subplots(1,2, figsize=(10,4))
plt.subplot(121)
plt.xlabel('x', fontsize='16')
plt.ylabel('y', fontsize='16')
plt.streamplot(X, Y, U, V, density=[1, 1])

plt.ylim((-4,4))
plt.xlim((-4,4))

plt.axhline(0, color='black')
```

```
plt.axvline(0, color='black')

plt.subplot(122)

plt.ylim((-1,1))
plt.xlim((0,25))
plt.xlabel('Time', fontsize='13')

plt.plot (m[:,0], m[:,1], color='r', linewidth=2, label='x')
plt.plot (m[:,0], m[:,2], color='b', linewidth=2, label='y')
plt.legend()

plt.savefig ('c:\\tmp\\phase.pdf')

plt.show()
```

**Listing 13.6** Script for Figure 13.10.

# 14

## ***Stability of Negative Feedback Systems***

### **14.1 Introduction**

---

A key property of negative feedback systems that was touched upon in Chapter 12 is their propensity to become unstable. In a negative feedback system most disturbances are damped due to the action of the feedback. However, what if the feedback mechanism takes too long to respond, so that by the time the negative feedback acts, the disturbance has already abated? In this situation the feedback would attempt to restore a disturbance that is no longer present, resulting in an incorrect action. Now imagine that the feedback system acts in the opposite direction to the disturbance because of the delay. This would cause the disturbance to grow rather than fall. If there is sufficient gain in the feedback loop to amplify or at least maintain the original disturbance, the result would be an unstable situation. If the loop gain amplifies, then the disturbance will grow until it reaches the physical limits of the system, at which point the loop gain is likely to fall and the disturbance will fall. The net result is a continuous growth and decline in the original disturbance, that is, a sustained oscillation.

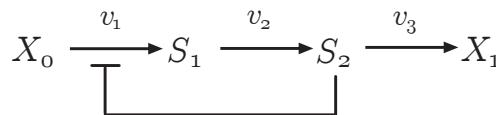
The key elements for sustaining an oscillation is a sufficient loop gain (at least 1.0) and a delay in the system response of exactly  $-360^\circ$ . Delays are measured by inputting a periodic signal such as a sine wave, and measuring the shift in phase that results. A  $-360^\circ$  phase shift means that the sine wave has been delayed by a full cycle. In a system with negative feedback, the delay is caused by two components. The first shift is  $-180^\circ$ , which comes from the inverting negative feedback. That is, if a rising disturbance reaches the feedback signal, the reaction rate at the signal target must fall and vice versa. The second component comes from the delay as the disturbance propagates through the reaction steps of the system. Each reaction step can contribute up to  $-90^\circ$  of delay; two reaction steps can contribute up

to  $-180^\circ$ , and three reaction steps up to  $-270^\circ$  and so on. However, the degree of delay caused by the reaction steps is a function of the frequency components in the disturbance. Random disturbances in the system will occur at all frequencies, therefore disturbances in a system that has sufficient loop gain will quickly locate the point where the delay caused by the reaction steps is exactly  $-180^\circ$ . Adding this to the negative feedback delay, we arrive at a total of  $-360^\circ$ . This results in a rapid destabilization of the system as disturbances are reinforced instead of damped.

Another way to look at this is that if a  $-360^\circ$  delay occurs, it means that the negative feedback is effectively behaving as a positive feedback, and positive feedbacks are normally destabilizing.

## 14.2 Stability using MCA

In this section we will look at the stability of negative feedback systems using metabolic control analysis. Consider first a simple three step pathway with negative feedback. This system (Figure 14.1) has just two steps within the feedback loop.  $s_2$  is the feedback signal that inhibits  $v_1$ .



**Figure 14.1** Negative Feedback with Three Reaction Steps.

We can compute the Jacobian using the expression introduced in the last chapter (13.7):

$$\mathbf{N} \frac{\partial \mathbf{v}}{\partial \mathbf{s}}$$

where  $\mathbf{N}$  is the stoichiometry matrix, and  $\partial \mathbf{v} / \partial \mathbf{s}$  is the matrix of unscaled elasticities. The stoichiometry matrix  $\mathbf{N}$  for the model in Figure 14.1 is:

$$\mathbf{N} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix}$$

As before, the symbol  $\mathcal{E}_j^i$  is used to denote the unscaled elasticity for reaction  $i$  with respect to species  $j$ , so that the elasticity matrix for the feedback model is given by:

$$\frac{\partial \mathbf{v}}{\partial \mathbf{s}} = \begin{bmatrix} 0 & \mathcal{E}_{s2}^{v1} \\ \mathcal{E}_{s1}^{v2} & 0 \\ 0 & \mathcal{E}_{s2}^{v3} \end{bmatrix}$$

where  $\mathcal{E}_{s_2}^{v_1}$  is the feedback elasticity. Replace each unscaled elasticity,  $\mathcal{E}$ , with the equivalent term involving the scaled elasticity,  $\varepsilon$ . This can be done by substituting each unscaled elasticity with the term:

$$\mathcal{E}_j^i = \varepsilon_j^i \frac{v_i}{s_j}$$

If the reactions  $v_2$  and  $v_3$  are assumed to be first-order and irreversible, then the corresponding elasticities,  $\varepsilon_1^2$ , and  $\varepsilon_2^3$ , can be set to one. This means that the  $\partial\mathbf{v}/\partial\mathbf{s}$  matrix can be written as:

$$\frac{\partial\mathbf{v}}{\partial\mathbf{s}} = \begin{bmatrix} 0 & \varepsilon_{s_2}^{v_1} F_1 \\ F_2 & 0 \\ 0 & F_3 \end{bmatrix}$$

where  $F_i$  are the scaling factors,  $v_i/s_j$ . Given the stoichiometry and unscaled elasticity matrix, the Jacobian matrix can be written as:

$$\mathbf{N} \frac{\partial\mathbf{v}}{\partial\mathbf{s}} = \begin{bmatrix} -F_2 & \varepsilon_{s_2}^{v_1} F_1 \\ F_2 & -F_3 \end{bmatrix}$$

To determine the stability of this system, the eigenvalues of the Jacobian must be evaluated. The standard approach (see Appendix B.9) to evaluating the eigenvalues is to compute the determinant of the expression:

$$\lambda \mathbf{I} - \mathbf{A} \quad (14.1)$$

where  $\mathbf{A}$  is the Jacobian and  $\lambda$  the eigenvalue. Using (14.1) it follows that:

$$\begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} - \begin{bmatrix} -F_2 & \varepsilon_{s_2}^{v_1} F_1 \\ F_2 & -F_3 \end{bmatrix} = \begin{bmatrix} \lambda + F_2 & -\varepsilon_{s_2}^{v_1} F_1 \\ -F_2 & \lambda + F_3 \end{bmatrix} \quad (14.2)$$

The determinant for a 2 by 2 matrix is given by:  $a_{11}a_{22} - a_{21}a_{12}$ , so that from (14.2):

$$\lambda^2 + \lambda(F_2 + F_3) + F_2 F_3 - F_1 F_2 \varepsilon_{s_2}^{v_1} = 0$$

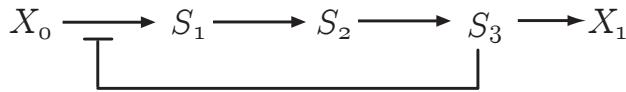
This is a quadratic equation so there will be at most two eigenvalues. The criterion for negative roots (and hence stability) in a quadratic equation is that all the coefficients have the same sign (see Appendix B.6). We should note that all  $F_i$  factors are positive, therefore it should be clear that the first two coefficients (attached to  $\lambda^2$  and  $\lambda$ ) are positive. The right most coefficient is  $-F_1 F_2 \varepsilon_{s_2}^{v_1}$  where the elasticity  $\varepsilon_{s_2}^{v_1}$  is negative because it represents the negative feedback. The expression,  $-F_1 F_2 \varepsilon_{s_2}^{v_1}$ , is therefore positive overall. Since all the coefficients are positive the solutions,  $\lambda$  must be both negative therefore system must be stable. There is no possibility of this system **ever** showing instability. Although it was assumed that the reaction  $v_2$  and  $v_3$  were first-order, relaxing this condition makes no

difference to the signs of the terms in the quadratic, therefore stability is also ensured under these less stringent conditions.

A linear pathway of three steps with a negative feedback loop from the second species to the first reaction will always be stable.

### Four Step Pathway

Let us look at a slightly modified system shown in Figure 14.2 where we have added an additional step in the feedback loop so that the pathway is four steps long.



**Figure 14.2** Negative Feedback in a Four Species Pathway.

The stoichiometry matrix and elasticity matrix are given by:

$$\mathbf{N} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}$$

$$\frac{\partial \mathbf{v}}{\partial \mathbf{x}} = \begin{bmatrix} 0 & 0 & \mathcal{E}_{s_3}^{v_1} \\ \mathcal{E}_{s_1}^{v_2} & 0 & 0 \\ 0 & \mathcal{E}_{s_2}^{v_3} & 0 \\ 0 & 0 & \mathcal{E}_{s_3}^{v_4} \end{bmatrix}$$

As before replace the unscaled elasticities,  $\mathcal{E}$ , using the expression:

$$\mathcal{E}_j^i = \varepsilon_j^i \frac{v_i}{s_j}$$

where the ratio  $v_i/s_j$  is in turn replaced with the factor  $F_i$ . As a point of reference the replacements will be:

$$\mathcal{E}_3^1 = \varepsilon_3^1 \frac{v_1}{s_3} = \varepsilon_3^1 F_1 \quad \mathcal{E}_1^2 = \varepsilon_1^2 \frac{v_2}{s_1} = \varepsilon_1^2 F_2$$

$$\mathcal{E}_2^3 = \varepsilon_2^3 \frac{v_3}{s_2} = \varepsilon_2^3 F_3 \quad \mathcal{E}_3^4 = \varepsilon_3^4 \frac{v_4}{s_3} = \varepsilon_3^4 F_4$$

To make matters simpler, assume first-order kinetics for  $\varepsilon_1^2$  and  $\varepsilon_3^4$ , so that these elasticities equal one. The product of the stoichiometry and elasticity matrix with therefore yield:

$$\mathbf{N} \frac{\partial \mathbf{v}}{\partial s} = \begin{bmatrix} -F_2 & 0 & \varepsilon_3^1 F_1 \\ F_2 & -F_3 & 0 \\ 0 & F_3 & -F_4 \end{bmatrix}$$

To compute the eigenvalues we evaluate the determinant of  $\lambda \mathbf{I} - \mathbf{A}$ :

$$\lambda \mathbf{I} - \mathbf{A} = \begin{bmatrix} \lambda + F_2 & 0 & \varepsilon_3^1 F_1 \\ F_2 & \lambda + F_3 & 0 \\ 0 & F_3 & \lambda + F_4 \end{bmatrix}$$

The determinant of the above matrix yields the following cubic equation:

$$\begin{aligned} \lambda^3 + v^2(F_2 + F_3 + F_4) + s(F_2 F_4 + F_2 F_3 + F_3 F_4) \\ + F_1 F_2 F_3 - F_1 F_2 F_3 \varepsilon_3^1 = 0 \end{aligned}$$

Since the pathway is linear, at steady state we can assert that the reaction rates,  $v_1, v_2, v_3$ , and  $v_4$  are equal to each other. This allows us to make a useful simplification. Recall that  $F_1 = v_1/s_3$  and  $F_4 = v_4/s_3$ , meaning  $F_1 = F_4$ . This changes the characteristic equation to:

$$\begin{aligned} \lambda^3 + \lambda^2(F_2 + F_3 + F_4) + s(F_2 F_4 + F_2 F_3 + F_3 F_4) \\ + F_2 F_3 F_4(1 - \varepsilon_3^1) = 0 \end{aligned}$$

where  $F_1$  has been replaced with  $F_4$ . For a general cubic equation such as (see Appendix B.6):

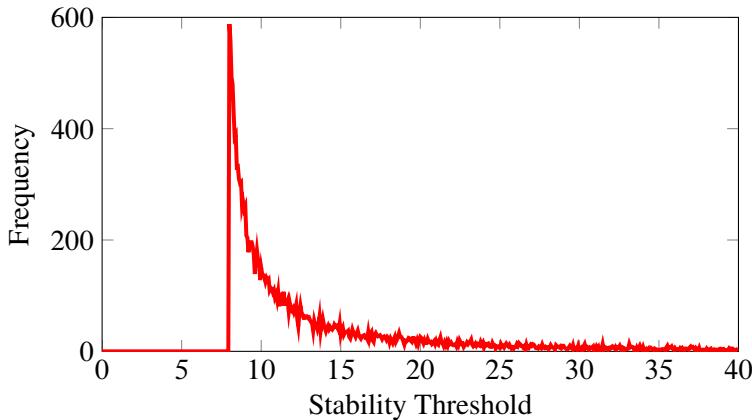
$$a_0 x^3 + a_1 x^2 + a_2 x + a_3 = 0$$

the condition for all negative roots is:

$$a_0 > 0, a_1 > 0, a_3 > 0 \quad \text{and} \quad a_1 a_2 > a_0 a_3$$

Negative roots of course indicate stability. We note that the first three conditions are satisfied because the first two coefficients are positive (all  $F_i$  terms are positive), and the expression  $(1 - \varepsilon_{s_3}^{v_1})$  is net positive (note  $\varepsilon_{s_3}^{v_1} < 0$ ). With these conditions satisfied, the requirement for stability ( $a_1 a_2 > a_0 a_3$ ) becomes:

$$(F_2 + F_3 + F_4)(F_2 F_4 + F_2 F_3 + F_3 F_4) > F_2 F_3 F_4(1 - \varepsilon_3^1)$$



**Figure 14.3** Stability Threshold for equation (14.3). Only feedback elasticities greater than eight result in instability.

Dividing both sides by  $F_2 F_3 F_4$  (this is positive, we therefore do not flip the inequality) yields a new inequality:

$$\frac{F_2}{F_3} + \frac{F_2}{F_4} + \frac{F_3}{F_4} + \frac{F_3}{F_2} + \frac{F_4}{F_3} + \frac{F_4}{F_2} + 3 > 1 - \varepsilon_3^1$$

Subtracting one from both sides yields:

$$\frac{F_2}{F_3} + \frac{F_2}{F_4} + \frac{F_3}{F_4} + \frac{F_3}{F_2} + \frac{F_4}{F_3} + \frac{F_4}{F_2} + 2 > -\varepsilon_3^1 \quad (14.3)$$

This is the condition for stability. So long as the left-hand side is greater than the negative of the feedback elasticity,  $\varepsilon_3^1$ , the system will be stable. The question now is what values can the left-hand side have? One way to pin down the inequality further is to try all combinations of  $F_i$  to discover what range of values the left-hand side can yield. Figure 14.3 shows combinations of  $F_i$  terms. The graph clearly shows that no combination falls below 8. Eight is therefore a lower bound for the left-hand term. This means that the feedback elasticity must be greater than -8 in order to achieve stability. Values less than -8 will cause instability.

Another way to look at equation (14.3) is to recall the relation:

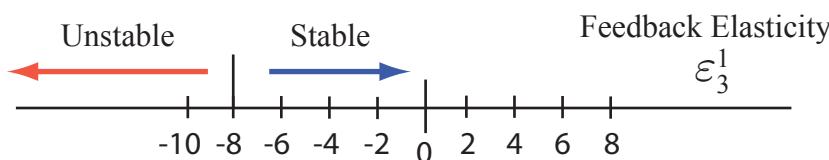
$$\frac{1}{a} + a \geq 2$$

Looking more closely at (14.3), we see that each pair of terms is in the form  $(1/a) + a$ . Therefore each pair of terms must be  $\geq 2$ . The lowest value each term can have is 2, therefore the minimum value for the sum of three of the paired terms must be 6. With the addition of the remaining 2 in the equation, it means that the minimum value for the left

hand side of equation ((14.3)) is 8. This matches the result from the numerical study in Figure 14.3. Swapping the negative signs in the inequality to make interpretation easier, we can summarize that if

$$\varepsilon_3^1 > -8 \quad (14.4)$$

then the system is stable.



**Figure 14.4** Threshold for Stability.

Figure 14.4 illustrates the threshold in diagrammatic form. Recall that the more negative  $\varepsilon_3^1$  is, the stronger the feedback. For example if  $\varepsilon_3^1 = -2.0$ , then the system will be stable since  $-2$  is greater than  $-8$ . If the strength of the negative feedback is less than  $-8$ , the system will be unstable. For example if  $\varepsilon_3^1 = -10$ , then the system will be unstable because  $-10 < -8$ . This means that instability in this system requires a fairly strong degree of feedback. Note that many allosteric enzymes tend to have elasticities for effectors in the range of  $-1$  to  $-4$ . A similar analysis can be done on pathways of other lengths, which will be considered in the next section.

From this analysis we can also make a few more observations. We have assumed that the elasticities for the effect of  $s_1$  on  $v_2$  and  $s_2$  on  $v_3$  was first-order. What if these values are closer to 0.5 where the substrate levels are near the  $K_m$  of the enzymes? The elasticities can be reintroduced so that the threshold for stability becomes:

$$\varepsilon_3^1 \varepsilon_1^2 \varepsilon_2^3 > -8$$

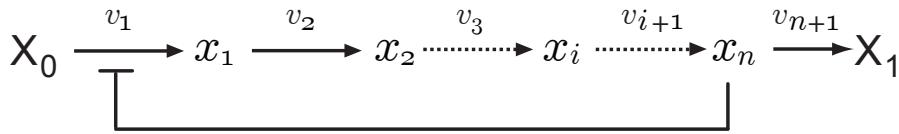
If the two elasticities have values around 0.5, then the threshold for instability will be greater and it becomes more difficult to reach an unstable situation. To illustrate this, if  $\varepsilon_1^2$  and  $\varepsilon_2^3$  have values of 0.5 each, the threshold equation becomes:

$$\frac{1}{4} \varepsilon_3^1 > -8$$

That is for stability:

$$\varepsilon_{s_3}^1 > -32$$

The threshold for stability is now much higher. It requires significant feedback strength to make such a system unstable.



**Figure 14.5** Negative feedback system with  $n$  species.

### 14.3 Effect of Pathway Length

Consider the negative pathway shown in Figure 14.5, which has  $n$  species and  $n + 1$  steps: The stoichiometry and elasticity matrices have the general form:

$$\mathbf{N} = \begin{bmatrix} 1 & -1 & 0 & 0 & \dots \\ 0 & 1 & -1 & 0 & \dots \\ 0 & 0 & 1 & -1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \quad \frac{\partial \mathbf{v}}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial v_1}{\partial s_1} & 0 & 0 & 0 & \dots & \frac{\partial v_1}{\partial s_n} \\ \frac{\partial v_2}{\partial s_1} & \frac{\partial v_2}{\partial s_2} & 0 & 0 & 0 & \dots \\ 0 & \frac{\partial v_3}{\partial s_2} & \frac{\partial v_3}{\partial s_3} & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \\ 0 & 0 & 0 & 0 & 0 & \frac{\partial v_{n+1}}{\partial s_n} \end{bmatrix}$$

Replacing  $\partial v / \partial s$  with  $\mathcal{E}$  and assuming all reactions are product insensitive, we obtain:

$$\frac{\partial \mathbf{v}}{\partial \mathbf{x}} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & \mathcal{E}_n^1 \\ \mathcal{E}_1^2 & 0 & 0 & 0 & 0 & \dots \\ 0 & \mathcal{E}_2^3 & 0 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \\ 0 & 0 & 0 & 0 & 0 & \mathcal{E}_n^{n+1} \end{bmatrix}$$

The Jacobian,  $\mathbf{N} \frac{\partial \mathbf{v}}{\partial \mathbf{x}}$  is given by:

$$\mathbf{N} \frac{\partial \mathbf{v}}{\partial \mathbf{x}} = \begin{bmatrix} -\mathcal{E}_1^2 & 0 & 0 & 0 & \dots & \mathcal{E}_n^1 \\ \mathcal{E}_1^2 & -\mathcal{E}_2^3 & 0 & 0 & \dots & 0 \\ 0 & \mathcal{E}_2^3 & -\mathcal{E}_3^4 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \\ 0 & 0 & 0 & 0 & \dots & -\mathcal{E}_n^{n+1} \end{bmatrix}$$

We now replace terms  $\mathcal{E}_j^i$  with  $\varepsilon_j^i F_j$  and as before compute  $\lambda \mathbf{I} - \mathbf{N} \frac{\partial \mathbf{v}}{\partial \mathbf{x}}$  from which we can derive the characteristic equation. Note that in the dynamical systems community the matrix is computed using  $(\mathbf{A} - \lambda \mathbf{I})$  so that the signs will be reversed, however the result is identical:

$$\lambda \mathbf{I} - \mathbf{N} \frac{\partial \mathbf{v}}{\partial \mathbf{x}} = \begin{bmatrix} \lambda + \varepsilon_1^2 F_2 & 0 & 0 & 0 & \dots & -\varepsilon_n^1 F_1 \\ -\varepsilon_1^2 F_2 & \lambda + \varepsilon_2^3 F_3 & 0 & 0 & \dots & 0 \\ 0 & -\varepsilon_2^3 F_3 & \lambda + \varepsilon_3^4 F_4 & 0 & \dots & 0 \\ 0 & 0 & -\varepsilon_3^4 F_4 & \lambda + \varepsilon_4^5 F_5 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \\ 0 & 0 & 0 & 0 & 0 & \lambda + \varepsilon_n^{n+1} F_{n+1} \end{bmatrix}$$

To make matters simpler, let us assume that all reactions other than the feedback step are first-order with *equal* rate constants. In such a situation the substrate elasticities,  $\varepsilon_j^i = 1$  and all species concentrations,  $s_j$ , are the same. This means that  $\varepsilon_1^2 = \varepsilon_2^3 = \varepsilon_3^4 = \varepsilon_n^i = 1$  and  $F_1 = F_2 = F_3 = F_n = F$ . The matrix can therefore be rewritten as:

$$\begin{bmatrix} \lambda + F & 0 & 0 & 0 & \dots & 0 & -\varepsilon_n^1 F \\ -F & \lambda + F & 0 & 0 & \dots & 0 & 0 \\ 0 & -F & \lambda + F & 0 & \dots & 0 & 0 \\ 0 & 0 & -F & \lambda + F & \dots & 0 & \lambda + F \end{bmatrix}$$

As before, the characteristic equation is obtained from the determinant which in this case has a simple form:

$$(\lambda + F)^n - \varepsilon_n^1 F^n = 0$$

Following Savageau [115], we now solve for  $\lambda$  using De Moivre's theorem (See Appendix B.9).

We first express the above equation by rearranging and taking the  $n^{th}$  root on both sides.

$$\begin{aligned} (\lambda + F)^n &= \varepsilon_n^1 F^n \\ \sqrt[n]{(\lambda + F)^n} &= \sqrt[n]{\varepsilon_n^1 F^n} \\ \lambda + F &= (\varepsilon_n^1)^{1/n} F \\ \lambda &= (\varepsilon_n^1)^{1/n} F - F \\ \lambda &= ((\varepsilon_n^1)^{1/n} - 1) F \end{aligned}$$

Because  $\varepsilon_n^1$  is negative, when taking the  $n^{th}$  root, we must take into account that there are  $n$  possible roots. To include this fact in the expression, we add a negative sign to  $\varepsilon_n^1$  and insert

$(-1)^{1/n}$  to compensate. The multiple roots are now found in the term  $(-1)^{1/n}$ :

$$\lambda = [(-1)^{1/n} (-\varepsilon_n^1)^{1/n} - 1]F$$

The  $n$  roots of  $(-1)^{1/n}$  can be obtained using de Moivre's theorem (See Appendix B.9):

$$\lambda_m = \left[ (-\varepsilon_n^1)^{1/n} \left( \cos \left( \frac{2m+1}{n}\pi \right) + i \sin \left( \frac{2m+1}{n}\pi \right) \right) - 1 \right] F \quad \text{for } m = 0, 1, 2, \dots, n-1$$

For the system to be stable, the *real part* of the  $\lambda_m$  must be negative. Since  $F$  is positive, for stability it must be true the real part of the expression must be negative:

$$(-\varepsilon_n^1)^{1/n} \cos \left( \frac{2m+1}{n}\pi \right) - 1 < 0$$

The transition to *instability* will occur when the largest  $\lambda_m$  becomes positive. This will happen at  $m = 0$  since the cosine term is maximum at this point. Setting  $m = 0$ , stability is assured when:

$$(-\varepsilon_n^1)^{1/n} \cos \left( \frac{\pi}{n} \right) - 1 < 0$$

For stability, the first term, not including the -1 on the left-hand side, must be *less* than +1 to make the overall expression on the left be less than 0. That is, for the system to be stable it must be true that:

$$(-\varepsilon_n^1)^{1/n} \cos \left( \frac{\pi}{n} \right) < 1 \tag{14.5}$$

Recall that  $n + 1$  is the number of steps in the pathway. Let us consider the cases where  $n = 1$  (two steps, one species), and  $n = 2$  (three steps, two species) which we know to be stable from the previous sections. When  $n = 1$ , the cosine term on the left-hand side of equation ((14.5)) is -1. Since the elasticity term will always be positive, the left-hand side will always be negative and thus never exceed 1. That is, the system will *always* be stable.

When  $n = 2$  (three steps), the left-hand side is 0 and again the left-hand side will *always* be less than one. That is, the system will *always* be stable.

When  $n = 3$  (four steps), the cosine term is 0.5. Therefore for the expression to exceed one, the elasticity term must equal at least 2. This happens when the elasticity has a value of -8 because  $8^{1/3} = 2$ . That is, the system becomes unstable when the elasticity is less than -8.

Another way to look at expression (14.5), and one commonly found in the literature, is to rearrange the inequality, noting the sign change rules, to obtain a slightly different expression:

$$-\varepsilon_n^1 < \frac{1}{\cos^n(\pi/n)} = \sec^n(\pi/n) \tag{14.6}$$

We therefore arrive at the central result [47, 114, 138, 135]:

**Condition for stability where  $n$  is the number of species:**

$$-\varepsilon_n^1 < \sec^n \left( \frac{\pi}{n} \right)$$

Number of Steps in Pathway	Instability Threshold $-\varepsilon_{\text{feedback}}$
2	stable
3	stable
4	8
5	4.0
6	2.9
7	2.4
8	2.1
9	1.9
:	:
$\infty$	1.0

**Table 14.1** Relationship between pathway length and the degree of feedback inhibition on the threshold for stability (See Figure 14.5).  $-\varepsilon_{\text{feedback}}$  is the elasticity of the feedback inhibition. The analysis assumes first-order kinetics for all reactions other than the feedback reaction. Additionally, it also assumes that the rate constants for the first-order reactions all have the same value. This ensures that the species levels are the same which simplifies the algebra.

When  $n = 3$  (four steps),  $\sec^3(\pi/3)$  equals 8, that is:

$$\varepsilon_3^1 < -8$$

for stability. This confirms the results we obtained previously in equation ((14.4)). That is, as long as the elasticity is larger than -8, the network is stable.

To give another example, if the number of steps in the pathway is eight ( $n = 7$ ), then  $\sec^7(\pi/7) = 2.08$ . This the feedback elasticity must be less than -2.08 for the system to be stable. Note this threshold is much less stringent. As the length of the pathway increases, it becomes easier for the system to be unstable.

Sometimes this condition is expressed as:

$$n_H > \frac{A}{\cos^n\left(\frac{\pi}{n}\right)}$$

When expressed this way,  $n_H$  is the term in a specific inhibition rate law,  $v_o/(1 + kS^{n_H})$ , and it therefore is always positive (this explains the change in the inequality sign). Equation (14.6) is however more general and does not depend on a specific mechanism for the negative feedback.

Table 14.1 summarizes some of the threshold points for a negative feedback system with varying pathway lengths.

## Further Reading

---

1. Edelstein-Keshet L (2005) Mathematical Models in Biology. SIAM Classical In Applied Mathematics. ISBN-10: 0-89871-554-7
2. Fall CP, Marland ES, Wagner JM, Tyson JJ (2000) Computational Cell Biology. Springer: Interdisciplinary Applied Mathematics. ISBN 0-387-95369-8.
3. Steuer R and Junker BH (2009). Computational models of metabolism: stability and regulation in metabolic networks. *Advances in chemical physics*, 142, 105.
4. M Savageau (1976) Biochemical systems analysis: a study of function and design in molecular biology, Addison-Wesley. Note: The text was republished in 2010 and is available on Amazon.

## Exercises

---

1. Run simulations of a four step pathway to show that the threshold for stability is when  $\varepsilon_3^1 < -8$ . Assume first-order irreversible kinetics for each reaction except for the regulated step,  $v_1$ . The rate law for the first step can be set to  $v_o/(1 + ks_3^h)$  where  $h$  is the Hill coefficient. Set all rate constants for the reactions steps to the same value. Investigate the stability of the systems as a function of changes to the Hill coefficient and hence  $\varepsilon_3^1$ .
2. Investigate the stability threshold for longer pathways using simulation. For example investigate a pathway with six steps and confirm the theoretical prediction made in Table 14.1 are correct.

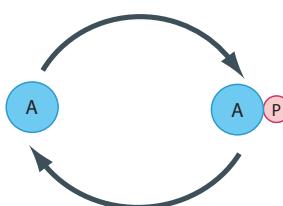
# 15

## *Moiety Conservation Laws*

### 15.1 Moiety Constraints

---

Many cell processes operate on different time scales. For example, metabolic processes tend to operate on a faster time scale than protein synthesis and degradation. Such time scale differences have a number of implications to model builders, software designers and model behavior. In this chapter we will examine these aspects in relation to species conservation laws. To introduce this topic, consider a simple protein phosphorylation cycle such as the one shown in Figure 15.1. This shows a protein undergoing phosphorylation (upper limb) and dephosphorylation (lower limb) via a kinase and phosphatase, respectively.

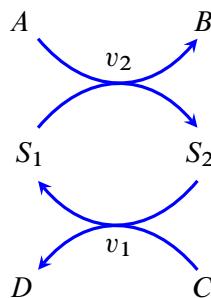


**Figure 15.1** Phosphorylation and dephosphorylation cycle forming a moiety conservation cycle between unphosphorylated (left species) and phosphorylated protein (right species).

The depiction in Figure 15.1 is a simplification. The ATP used during phosphorylation is not shown, as well as the release of free phosphate during the dephosphorylation event. Furthermore, synthesis and degradation of protein is also absent. In many cases we can leave these aspects out of the picture. ATP for instance is held at a relatively constant level by

strong homeostatic forces from metabolism so that within the context of the cycle, changes in ATP isn't something we need worry about. More interestingly is that within the time scale of phosphorylation and dephosphorylation, we can assume that the rate of protein synthesis and degradation is negligible. This assumption leads to the emergence of a new property of the cycle called **moiety conservation** [103].

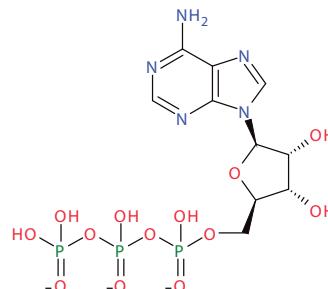
In chemistry a moiety is described as a subgroup of a larger molecule. In Figure 15.1 the moiety is a protein. During the interconversion between the phosphorylated and unphosphorylated protein, the amount of moiety (protein) remains constant. More abstractly, we can draw a cycle in the following way (Figure 16.2), where  $S_1$  and  $S_2$  are the cycle species:



**Figure 15.2** Simple conserved cycle where  $s_1 + s_2 = \text{constant}$ . Note: This assumes unit volume for both  $S_1$  and  $S_2$  so that concentrations can be summed.

The two species,  $S_1$  and  $S_2$ , are conserved because the total  $s_1 + s_2$  remains constant over time (at least over a time scale shorter than protein synthesis and degradation). Such cycles are collectively called **conserved cycles**.

Protein signalling pathways abound with conserved cycles such as these although many are more complex and may involve multiple phosphorylation reactions. In addition to protein networks, other pathways also possess conservation cycles. One of the earliest conservation cycles to be recognized was the adenosine triphosphate (ATP) cycle. ATP is a chain of three phosphate residues linked to a nucleoside adenosine group, Figure 15.3.

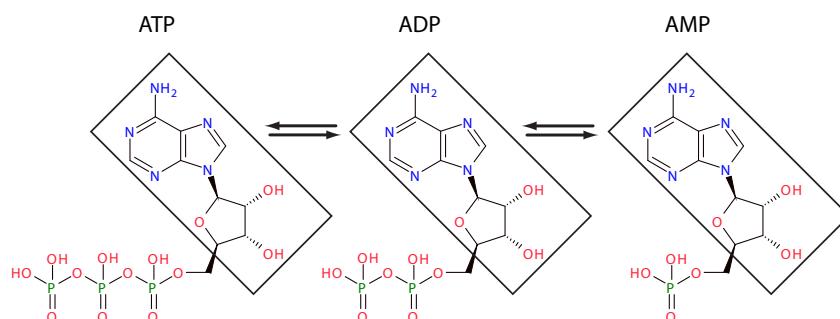


**Figure 15.3** Adenosine Triphosphate: Three phosphate groups plus an adenosine subgroup.

The linkage between the phosphate groups involves phosphoric acid anhydride bonds that can be cleaved by hydrolysis one at a time leading in turn to the formation of adenosine diphosphate (ADP) and adenosine monophosphate (AMP), respectively. The hydrolysis provides much of the free energy to drive endergonic processes in the cell. Given the insatiable need for energy, there is a continual and rapid interconversion between ATP, ADP and AMP as energy is released or captured. One thing that is constant during these interconversions is the amount of adenosine group (Figure 15.4). That is, adenosine is a conserved moiety. Over longer time scales there is also the slower process of AMP degradation and biosynthesis via the purine nucleotide pathway, but for many models, we assume that this process is negligible compared to the ATP turn over by energy metabolism.



There are many other examples of conserved moieties such enzyme/enzyme-substrate complexes, NAD/NADH, phosphate and coenzyme A. In all these cases the basic assumption is that the interconversions of the subgroups is rapid compared to their net synthesis and degradation. We should emphasize that in reality conserved moieties do not exist since all molecular subgroups will at some point be subject to synthesis and degradation. However, over sufficiently short time scales, the sum total of these groups can be considered constant. In this chapter we will consider conserved moieties in detail. In particular, we will look at how to detect them in our models and how they influence the design of simulation software. We will wait until Chapter 16 to discuss their effect on pathway dynamics and their relationship to MCA.

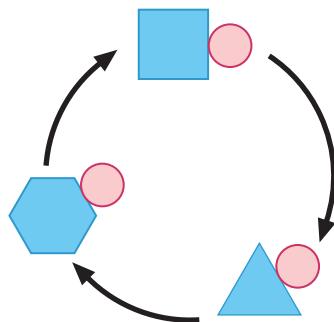


**Figure 15.4** The adenosine moiety, indicated by the boxed molecular group, is conserved during the interconversion of ATP, ADP and AMP.

**Moiety:** A subgroup of a larger molecule.  
**Conserved Moiety:** A subgroup whose interconversion through a sequence of reactions leaves it unchanged.

## 15.2 Moiety Conserved Cycles

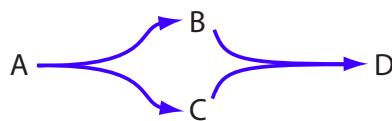
Any chemical group that is preserved during a cyclic series of interconversions is called a **conserved moiety**. Examples of conserved moiety subgroups include species such as phosphate, acyl, nucleoside groups or covalently modifiable proteins. As a moiety gets redistributed through a network, the **total amount** of the moiety is constant and does not change during the time evolution of the system. For any particular subgroup, the total amount is determined solely by the initial conditions imposed on the model.



**Figure 15.5** Conserved Moiety in a Cyclic Network. The darker species are modified as they traverse the reaction cycle, but the lighter subgroup (smaller circle), remains unchanged. This creates a conserved cycle, where the total number of moles of moiety (smaller circle) stays constant.

There are rare cases when a ‘conservation’ relationship arises out of a non-moiety cycle. This does not affect the mathematic analysis, only the physical interpretation of the relationship. For example, in Figure 15.6 the constraint  $b - c = T$  applies even though there is no moiety involved.

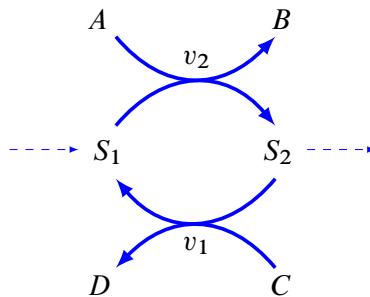
The presence of conserved moieties is an approximation introduced into a model, however, over the time scale in which the conservation holds, their existence can have a profound effect on the dynamic behavior of the model. For example, the hyperbolic response of a simple enzyme (in the form of enzyme conservation between  $E$  and  $ES$ ), or the sigmoid behavior observed in protein signalling networks, is due in significant part to moiety conservation laws (see Chapter 16).



**Figure 15.6** Conservation due to stoichiometric matching. In this system,  $b - c = \text{constant}$ .

In the subsequent discussion it will be assumed that all species reside in a unit volume. This makes it possible to express the conservation laws using concentrations. If the volumes are non-unity and of varying magnitude, the conservation laws would have to be adjusted with respect to the volume because it is amounts that are strictly conserved, not concentration. For an introductory discussion non-unity volumes are an unnecessary distraction to the focus of the chapter and we will therefore assume unit volume in all cases.

Figure 15.7 illustrates the simplest possible network which displays a conserved moiety, the total mass,  $s_1 + s_2$ , is constant during the evolution of the network.



**Figure 15.7** Simple conserved cycle. The dotted lines signify negligible levels of synthesis and degradation, therefore over short time scales,  $s_1 + s_2 = \text{constant}$ .

The system equations for the simple conserved cycle in Figure 15.7 are written as:

$$\frac{ds_1}{dt} = v_1 - v_2$$

$$\frac{ds_2}{dt} = v_2 - v_1$$

From these equations it should be evident that the rate of appearance of  $S_1$  must equal the rate of disappearance of  $S_2$ , that is  $ds_1/dt = -ds_2/dt$ . This means that whenever  $S_1$  changes,  $S_2$  must change in the opposite direction by **exactly** the same amount. During this time, the sum of  $s_1$  and  $s_2$  will therefore remain unchanged.

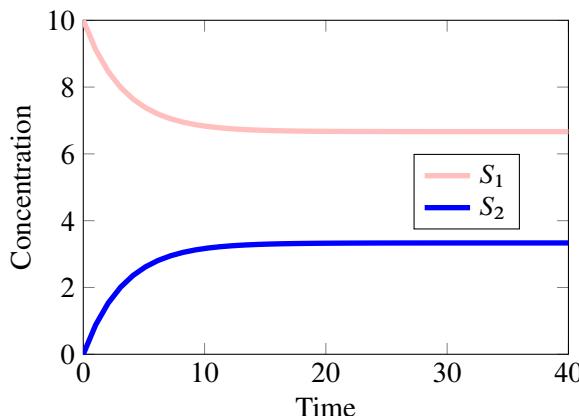
When solving the system by computer we only need to numerically integrate one of the species because the other species can be computed from the conservation relationship.

Whichever differential equation is chosen however, the species left out must be computed algebraically using the conservation law. Therefore, the system can be reduced (15.1) to one differential and one linear algebraic equation (15.1) compared to the two differential equations in the original formulation.

$$s_2 = T - s_1$$

$$\frac{ds_1}{dt} = v_1 - v_2 \quad (15.1)$$

The term  $T$  in the algebraic equation (15.1) refers to the total amount of  $S_1$  and  $S_2$ . This value is computed from the initial amounts given to  $S_1$  and  $S_2$  at the start of a simulation.



**Figure 15.8** Simulation of the simple cycle shown in Figure 15.7. The total moiety remains constant at 10 concentration units. Model:  $S_1 \rightarrow S_2; k_1 * S_1$ ;  $S_2 \rightarrow S_1; k_2 * S_2$ ;  $S_1=10$ ;  $k_1=0.1$ ;  $k_2=0.2$ .

### 15.3 Basic Theory

The question we want to address here is how to determine whether a given network contains conserved cycles and if so what are they? The key to this question is the stoichiometry matrix,  $\mathbf{N}$ . In the example shown in Figure 15.7 the stoichiometry matrix is given by:

$$\mathbf{N} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

Since either row can be derived from the other by multiplication by  $-1$ , the rows are called **linearly dependent rows** (See Box 15.0). The linear dependence among the rows reflects the earlier relationship  $ds_1/dt = -ds_2/dt$ . Any linear dependence among the rows of the stoichiometry matrix means that there are linear relationships among the rates of change,

### Box 15.0 Linear Dependence and Independent - Recap

One of the most important ideas in linear algebra is the concept of linear dependence and independence. Take three vectors, say  $[1, -1, 2]$ ,  $[3, 0, -1]$  and  $[9, -3, 4]$ . If we look at these vectors carefully it should be apparent that the third vector can be generated from a combination of the first two, that is  $[9, -3, 4] = 3[1, -1, 2] + 2[3, 0, -1]$ . Mathematically we say that these vectors are *linearly dependent*.

In contrast, the following vectors,  $[1, -1, 0]$ ,  $[0, 1, -1]$  and  $[0, 0, 1]$ , are independent because there is no combination of these vectors that can generate even one of them. Mathematically we say that these vectors are *linearly independent*.

$ds/dt$ . In turn, relationships among the rates of change means there are mass constraints on how the individual amounts can change. In the previous example it was shown that for every mole of  $S_1$  of that was consumed, one mole of  $S_2$  was produced and for every mole of  $S_2$  consumed one mole of  $S_1$  was produced. Therefore since  $ds_1/dt + ds_2/dt$  must be zero, this implies that  $s_1 + s_2$  remains constant.

In conclusion, whenever a network exhibits conserved moieties, there will be dependencies among the rows of  $\mathbf{N}$ . A measure of such dependencies is the rank, denoted  $rank(\mathbf{N})$  (See Box 15.1). Any dependencies among the rows will be reflected in a rank that is less than  $m$ , the number of rows of  $\mathbf{N}$ .

It is possible to arrange the rows of  $\mathbf{N}$  so that the first  $rank(\mathbf{N})$  rows are linearly independent. The metabolites which correspond to these rows are called the **independent species** ( $s_i$ ). The remaining  $m - rank(\mathbf{N})$  rows correspond to the **dependent species** ( $s_d$ ).

In the simple conserved cycle, Figure 15.7, there is one independent species,  $S_1$  and one dependent species,  $S_2$ .

#### Example 15.1

Figure 15.5 illustrates a three species cycle. What is the conservation law for this pathway? The stoichiometry matrix for this system is given by:

$$\mathbf{N} = \begin{bmatrix} v_1 & v_2 & v_3 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix} \begin{array}{l} S_1 \\ S_2 \\ S_3 \end{array} \quad (15.2)$$

Inspection reveals that the sum of the three rows is zero meaning that:

$$\frac{ds_1}{dt} + \frac{ds_2}{dt} + \frac{ds_3}{dt} = 0$$

or that the total  $s_1 + s_2 + s_3$  is constant. There are no other relationships between the rows other than this one.

**Example 15.2**

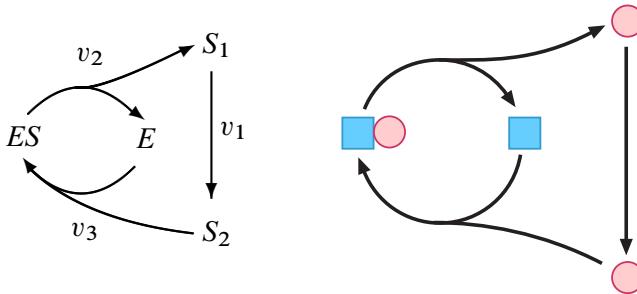
A linear pathway has the following stoichiometry matrix:

$$\mathbf{N} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix}$$

Does the pathway contain any conserved cycles? No, because neither row in the matrix can be derived from the other by a linear combination. The rows are linearly independent, therefore the pathway has no conserved cycles.

---

To illustrate this idea on a more complicated example, consider the pathway shown in Figure 15.9. This pathway includes four species,  $S_1$ ,  $S_2$ ,  $E$  and  $ES$ .



**Figure 15.9** Linked Conserved Cycles. The network rendered on the right shows the moiety composition of the participating species.

The mass-balance equations of this model can be written down as:

$$\begin{aligned}\frac{de}{dt} &= v_2 - v_3 & \frac{des}{dt} &= v_3 - v_2 \\ \frac{ds_1}{dt} &= v_2 - v_1 & \frac{ds_2}{dt} &= v_1 - v_3\end{aligned}$$

A visual inspection of the mass-balance equations reveals the following two relationships:

$$\begin{aligned}\frac{de}{dt} + \frac{des}{dt} &= 0 \\ \frac{des}{dt} + \frac{ds_1}{dt} + \frac{ds_2}{dt} &= 0\end{aligned}\tag{15.3}$$

### Box 15.1 The Rank of a Matrix - Recap

Closely related to linear independence (Box 15.0) is the concept of Rank. Consider the three vectors described in Box 15.0,  $[1, -1, 2]$ ,  $[3, 0, 1]$  and  $[9, -3, 4]$  and stack them one atop each other to form a matrix:

$$\begin{bmatrix} 1 & -1 & 2 \\ 3 & 0 & 1 \\ 9 & -3 & 4 \end{bmatrix}$$

then the Rank is the number of linear independent vectors that make up the matrix. In this case the Rank is 2, because there are only two linear independent row vectors in the matrix.

These relationships tell us that there are two conservation laws,  $e + es$  and  $es + s_1 + s_2$ . This means that given the amount of  $ES$ , the amount of  $E$  can be computed. In addition, given the amount of  $ES$  and  $S_1$ , the amount of  $S_2$  can be computed. Therefore  $ES$  and  $S_1$  can be designated the independent species and  $E$  and  $S_2$  the dependent species. What this means in practical terms is that in a modeling program only two differential equations need be solved instead of four. The reduced model equations will look like:

$$e = T_1 - es$$

$$s_2 = T_2 - s_1 - es$$

$$\frac{des}{dt} = v_3 - v_2$$

$$\frac{ds_1}{dt} = v_2 - v_1$$

where  $T_1$  is the total amount of  $E$  type moiety, and  $T_2$  is the total amount of  $S$  type moiety.

The stoichiometry matrix for the model in Figure 15.9 is given by:

$$\mathbf{N} = \begin{bmatrix} v_1 & v_2 & v_3 \\ 1 & 0 & -1 \\ 0 & -1 & 1 \\ -1 & 1 & 0 \\ 0 & 1 & -1 \end{bmatrix} \begin{array}{l} S_2 \\ ES \\ S_1 \\ E \end{array} \quad (15.4)$$

Examining the stoichiometry matrix reveals conservation laws as relationships among the matrix rows. The 4th row ( $E$ ) can be formed by multiplying the 2nd row ( $ES$ ) by -1, and the 3rd row ( $S_1$ ) can be formed by multiplying the first row by -1 and adding it to the 4th row ( $ES$ ).

These simple examples show that it is possible to derive conservation laws by looking for dependencies among the rows of the stoichiometry matrix. For simple cases this can be done by inspection but for large pathways this approach is not practical. Instead a more systematic theory for deriving the conservation laws must be developed.

## 15.4 Computational Approaches

---

There are a number of related methods for computing the conservation laws of a given pathway, some are simple such as the one shortly to be described, while others are more sophisticated and are used to determine the conservation laws in very large stoichiometry matrices.

The easiest method to derive conservation laws is to use row reduction [96, 24, 25]. This is based on forward elimination which is the first part of Gaussian Elimination. Gaussian Elimination is a traditional way to solve simultaneous linear equations by eliminating one unknown at a time, and is a technique often taught in high school. Elimination is carried out by applying a series of simple manipulations called **elementary operations**. These operations include interchanging two equations (exchange), multiplying an equation through by a nonzero number (scaling), and adding an equation one or more times to another equation (replacement). In practice the equations are cast into a matrix form so that the elementary operations are applied to the values in the matrix where each row of the matrix represents an equation. Thus, interchanging two equations is equivalent to swapping two rows in the matrix. The elementary operations are carried out on the matrix until a particular arrangement, called the echelon form, is established (See Box 15.3).

Elementary operations are often represented in matrix form and are then called elementary matrices (See Box 15.1). Applying a particular elementary operation then becomes equivalent to multiplying by an elementary matrix.

The technique for finding conservation laws works as follows. Consider the network in Figure 15.9. The system equation for this network is:

$$\begin{matrix} S_2 \\ ES \\ S_1 \\ E \end{matrix} \left[ \begin{array}{ccc} 1 & 0 & -1 \\ 0 & -1 & 1 \\ -1 & 1 & 0 \\ 0 & 1 & -1 \end{array} \right] \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} ds_2/dt \\ des/dt \\ ds_1/dt \\ de/dt \end{bmatrix}$$

We will cast the equation in the following form where an identity matrix has been added to the right-hand side.

$$\mathbf{N}v = I \frac{ds}{dt}$$

Written out fully the system equation will look like:

### Box 15.1 Elementary Matrices - Recap

Elementary matrix operations such as row exchange, row scaling or row replacement can be represented by simple matrices called elementary matrices, called Type I, II and III, respectively. Elementary matrices can be constructed from the identity matrix. For example, a scaling operation can be represented by replacing one of the elements of the main diagonal of an identity matrix by the scaling factor. The following matrix represents a type II matrix which will scale the second row of a given matrix by the factor  $k$ :

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Type I elementary matrices will exchange two given rows in a given matrix and are constructed from an identity matrix where rows in an identity matrix are exchanged that correspond to the rows exchanged in the target matrix. The following type I matrix will exchange rows 2 and 3 in a target matrix:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Type III elementary matrices will add/subtract a given row in a target matrix to another row in the same matrix. Type III matrices are constructed from an identity matrix where a single off diagonal element is set to the multiplication factor, and the specific location represents the two rows to combine. If an elementary matrix adds a row  $i$  to a row  $j$  multiplied by a factor  $\alpha$ , then the identity matrix with entry  $i, j$  is set to  $\alpha$ . In the following example, the type III elementary matrix will subtract five times the 2nd row from the 3rd row.

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -5 & 0 \end{bmatrix}$$

A particularly important property of elementary matrices is that they can all be inverted. In addition, pre-multiplying by an elementary matrix will modify the rows of a target matrix while post-multiplying will operate on the columns.

$$\begin{matrix} S_2 \\ ES \\ S_1 \\ E \end{matrix} \left[ \begin{array}{ccc} 1 & 0 & -1 \\ 0 & -1 & 1 \\ -1 & 1 & 0 \\ 0 & 1 & -1 \end{array} \right] \left[ \begin{array}{c} v_1 \\ v_2 \\ v_3 \end{array} \right] = \left[ \begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array} \right] \left[ \begin{array}{c} ds_2/dt \\ des/dt \\ ds_1/dt \\ de/dt \end{array} \right]$$

Forward elimination is first applied to the stoichiometry matrix. To do this a series of elementary operations to the left-hand side is applied such that the stoichiometry matrix is reduced to echelon form. For consistency we apply the same set of elementary operations to the right-hand side so that the identity matrix records whatever operations we carry out. This amounts to multiplying both sides by a set of elementary matrices. We only need to reduce the matrix to its row echelon form, not to its reduced echelon form (See Box 15.3).

Reducing a matrix to echelon form raises the possibility of generating zero rows in the matrix if there are dependencies in the rows (See Box 15.3).

With this being the case, the system equation after forward elimination can be expressed in the following way:

$$\left[ \begin{array}{c} M \\ \mathbf{0} \end{array} \right] v = E \frac{ds}{dt} \quad (15.5)$$

where the identity matrix has been shown transformed into the matrix  $E$  which represents the product of all elementary operations that were applied to the left-hand side. The left-hand side has itself been transformed into an echelon form which is represented as a partitioned matrix of  $M$  and  $\mathbf{0}$ . The  $E$  matrix can also be partitioned row-wise to match the partitioning in the echelon matrix, that is:

$$\left[ \begin{array}{c} M \\ \mathbf{0} \end{array} \right] v = \left[ \begin{array}{c} X \\ Y \end{array} \right] \frac{ds}{dt} \quad (15.6)$$

Multiplying out the lower partition, we obtain:

$$Y \frac{ds}{dt} = \mathbf{0} \quad (15.7)$$

This general result is equivalent to the equations shown in (15.3), that is (15.7) represents the set of conservation laws. Determining the conservation laws therefore involves reducing the stoichiometry matrix and extracting the lower portion of the modified identity matrix.

Let us now proceed with an example to illustrate this method. We will use the stoichiometry matrix from equation (15.4). For convenience, the stoichiometry and identity matrix are placed next to each other in the following sequence of elementary operations. An elementary operation carried out on the stoichiometry matrix is simultaneously applied to the identity matrix.

### Box 15.3 Echelon Forms - Recap

There are two kinds of matrices that one frequently encounters in the study of linear equations. These are the **row echelon** and **reduced echelon forms**. Both matrices are generated when solving sets of linear equations. The row echelon form is derived using forward elimination and the reduced echelon form by Gauss-Jordan Elimination.

A **row echelon matrix** is defined as follows:

1. All rows that consist entirely of zeros are at the bottom of the matrix.
2. In each non-zero row, the first non-zero entry is a 1, called the leading one.
3. The leading 1 in each row is to the right of all leading 1's above it. This means there will be zeros below each leading 1.

The following three matrices are examples of row echelon forms:

$$\begin{bmatrix} 1 & 4 & 3 & 0 \\ 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 5 & 3 & 0 \\ 0 & 1 & 7 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The **reduced echelon form** has one additional characteristic:

4. Each column that contains a leading one has zeros above and below it. The following three matrices are examples of reduced echelon forms:

$$\begin{bmatrix} 1 & 0 & 4 & 0 \\ 0 & 1 & 1 & 7 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Sometimes the columns of a reduced echelon can be ordered such that each leading one is immediately to the right of the leading one above it. This will ensure that the leading 1's form an identity matrix at the front of the matrix. The reduced echelon form will therefore have the following general block structure:

$$\begin{bmatrix} I & A \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$$

It is always possible to reduce any matrix to its echelon or reduced echelon form by an appropriate choice of elementary operations. The function `rref()` implemented in many math software applications will generate a reduced row echelon.

**1. Stoichiometry matrix on the left and identity matrix on the right.**

$$\begin{bmatrix} 1 & 0 & -1 \\ 0 & -1 & 1 \\ -1 & 1 & 0 \\ 0 & 1 & -1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

**2. Add the 1st row to the third row to yield:**

$$\begin{bmatrix} 1 & 0 & -1 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \\ 0 & 1 & -1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

**3. Add the 2nd row to the third and forth rows to yield:**

$$\begin{bmatrix} 1 & 0 & -1 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$$

**4. Multiply the second row by -1 to yield the final echelon form:**

$$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$$

The final operation achieves the goal of reducing the stoichiometry matrix to an echelon form (in this case it happens to be a reduced echelon form). Note that the operation has resulted in two zero rows appearing in the reduced stoichiometry matrix. These two rows correspond to the  $Y$  partition in equation (15.6). The lower two rows can be extracted from the right-hand matrix (what was once the identity matrix) to construct equation (15.7), thus:

$$\begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} ds_2/dt \\ des/dt \\ ds_1/dt \\ de/dt \end{bmatrix} = 0$$

Or:

$$\frac{ds_2}{dt} + \frac{des}{dt} + \frac{ds_1}{dt} = 0$$

$$\frac{des}{dt} + \frac{de}{dt} = 0$$

From the above equations the following conservation laws should be evident:

$$\begin{aligned} s_2 + es + s_1 &= T_1 \\ es + e &= T_2 \end{aligned} \quad (15.8)$$

In summary, the algorithm for deriving the conservation laws is as follows:

1. Apply elementary operations to the stoichiometry matrix until the matrix is reduced to its row echelon form. Simultaneously apply the elementary operations to an identity matrix. The size of the identity matrix should be equal to the number of rows in the stoichiometry matrix.
2. If there are zero rows at the bottom of the reduced stoichiometry matrix then there are conservation laws in the network; otherwise there are not. The number of conservation laws will be equal to the number of zero rows.
3. Extract the rows in the transformed identity matrix that correspond to the position of the zero rows in the reduced stoichiometry matrix. The extracted rows represent the conservation laws.

There are two points worth making when applying this algorithm. The first is that any row swaps made using the row reduction in the stoichiometry matrix will not translate to swaps in the names of the species on the right-hand side of the equation. This means that when reading the conservation rows, the names on the columns are not changed by any row exchanges in the stoichiometry matrix. The second point to make is that when carrying out the elementary row operations, it is recommended to eliminate, whenever possible, terms below a leading entry by adding rather than subtracting. This will ensure that entries in the transforming identity matrix remain positive and that the resulting conservation laws will be made up of positive terms. Sometimes the ability to add will not be possible and subtractions will be necessary. This results in negative terms appearing in the conservation laws which may make them more difficult to interpret physically.

A useful strategy to avoid negative terms in the conservation equations is to order the rows of the stoichiometry matrix such that any species that is likely to appear in more than one conservation relationship is placed at the bottom of the stoichiometry matrix. In the case of the previous example we would make sure that ES becomes the bottom row of the stoichiometry matrix. This ordering ensures that the independent species (top rows) are represented by the so-called free variables and the dependent species (bottom rows) by the shared variables. This means that the shared or dependent variables (i.e. complexes) will then be a function of the free variables which is more likely to result in positive terms [111]. A more brute force method is to try all permutations of the matrix rows until a positive set of conservation laws is found. For small models (< 10 species) this approach is a viable option.

Although it is possible to manually reduce a stoichiometry matrix, it is far easier to use specialized math software such Scilab, Octave, Python, Matlab and Mathematica or even

advanced modern pocket calculators. All these tools offer a `rref()` command for generating a reduced row echelon. The following examples will illustrate the use of the freely available Scilab application ([www.scilab.org](http://www.scilab.org)) to compute the conservation laws.

### Example 15.3

---

Row reduction using Scilab/Matlab. Given the following stoichiometry matrix, use Scilab functions to row reduce and extract the conservation laws.

$$\mathbf{N} = \begin{matrix} S_2 \\ ES \\ S_1 \\ E \end{matrix} \left[ \begin{array}{ccc} 1 & 0 & -1 \\ 0 & -1 & 1 \\ -1 & 1 & 0 \\ 0 & 1 & -1 \end{array} \right]$$

Enter the stoichiometry matrix into the software:

```
-->n = [1 0 -1; 0 -1 1; -1 1 0; 0 1 -1];
```

Augment the matrix with the identity matrix, this will allow us to record row reduction operations in the identity matrix part of the augmented matrix.

```
-->ni = [n, eye(4,4)]
ni =
 1. 0. - 1. 1. 0. 0. 0.
 0. - 1. 1. 0. 1. 0. 0.
 - 1. 1. 0. 0. 0. 1. 0.
 0. 1. - 1. 0. 0. 0. 1.
-->
```

Row reduce the augmented matrix:

```
-->rni = rref (ni)
rni =
 1. 0. - 1. 0. 0. - 1. 1.
 0. 1. - 1. 0. 0. 0. 1.
 0. 0. 0. 1. 0. 1. - 1.
 0. 0. 0. 0. 1. 0. 1.
```

The left partition of the reduced matrix contains two zero rows, therefore there are two conservation laws. These laws correspond to the two bottom rows in the right partition. We extract the rows in the right partition to yield:

```
-->c = rni(3:4,4:7)
c =
 1. 0. 1. - 1.
 0. 1. 0. 1.
```

The species column order is the same as the species row order in the original matrix, that is  $S_2$ ,  $ES$ ,  $S_1$  and  $E$ , therefore:

$$\begin{aligned}s_2 + s_1 - e &= T_1 \\ es + e &= T_2\end{aligned}$$

Note the negative  $E$  term in the first conservation law. At first glance this does not appear to be the same set of conservation laws that were derived earlier. However, if we substitute  $E$  from the second equation into the first, we will get the same set of conservation laws:  $S_1 + S_2 + ES = T$ , showing us that the two sets are identical. To avoid negative terms appearing in the conservation laws, we can use the rule that all complex species (that is shared species) such as  $ES$ , be moved to the bottom of the matrix (See next example).

The previous example can be also accomplished in Python using the sympy package, although its not as straight-forward <sup>1</sup>:

```
import sympy
import numpy as np

n = sympy.Matrix ([[1,0,-1],[0,-1,1],[-1,1,0],[0,1,-1]])

fs = sympy.Matrix (np.hstack ((n, sympy.eye (4))))
rni = fs.rref()[0]
print rni

c = rni[2:4,3:7]
print c
```

#### *Example 15.4*

Row reduction using Scilab/Matlab. Given the following stoichiometry matrix, use Scilab functions to row reduce and extract the conservation laws. In this example the shared species  $ES$  has been moved to the bottom of the matrix.

$$\mathbf{N} = \begin{matrix} S_2 \\ S_1 \\ E \\ ES \end{matrix} \left[ \begin{array}{ccc} 1 & 0 & -1 \\ -1 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{array} \right]$$

The reduced augmented matrix is:

```
-->rni = rref (ni)
rni =
```

<sup>1</sup>Although Python is touted as being one of the easiest languages to learn, it certainly does make some operations unnecessarily complex.

1.	0.	-	1.	0.	-	1.	0.	-	1.
0.	1.	-	1.	0.	0.	0.	0.	-	1.
0.	0.	0.	1.	1.	1.	0.	1.	1.	1.
0.	0.	0.	0.	0.	0.	1.	1.	1.	1.

Once again there are two zero rows, but this time the corresponding conservation laws all have positive entries, yielding the following equations:

$$\begin{aligned}s_2 + s_1 + es &= T_1 \\ es + e &= T_2\end{aligned}$$


---

The Scilab/Matlab code shown in Figure 15.10 will find the conservation laws for any stoichiometry matrix.

```
// Compute Conservation Laws
// -----
// Enter the stoichiometry matrix first

n = [1 0 -1; 0 -1 1; -1 1 0; 0 1 -1];
nRows = size(n, 1);
// Create the augmented matrix
ni = [n, eye(nRows, nRows)];
// Carry out row reduction
rni = rref (ni);
r = rank (n);
// Extract the conservation rows
c = rni(r+1:nRows, size(n, 2)+1:size(ni, 2));
// Display result
c
```

**Figure 15.10** General purpose Scilab/Matlab code to determine conservation laws using row reduction.

Similar calculations can be done using Python. However in Python one should import the SymPy library rather than the `scipy.linalg` library. For whatever reason the creators of SciPy made the explicitly decision not to include support for functions such as `rref` or even for computing the null space. Users of these libraries are expected to write their own functions to accomplish these tasks.

The equivalent code using SymPy in Python to compute the conservation laws is given in Figure 15.11. If we use only functions from the `sympy` library the code is quite compact. Two helper functions, `getNumColumns` and `getNumRows` were defined to make the code easier to read for those not familiar with the `shape` syntax. Many of the these computations can be done automatically using Tellurium.

```
# Compute Conservation Laws

import numpy as np
import sympy

def getNumColumns (m): return m.shape[1]
def getNumRows (m):    return m.shape[0]

n = sympy.Matrix ([[1, 0, -1], [0, -1, 1], [-1, 1, 0], [0, 1, -1]]);
nRows = getNumRows (n)
# Create the augmented matrix
ni = sympy.Matrix.hstack (n, sympy.Matrix.eye(nRows));
# Carry out the row reduction
rni = ni.rref ();
# Compute the rank
r = sympy.Matrix (n).rank();
# Extract the conservation rows
c = rni[0][r:nRows, getNumColumns (n):getNumColumns (ni)];
print c
```

**Figure 15.11** General purpose Python code to determine conservation laws using row reduction.

Row reduction of the augmented stoichiometry is probably the easiest way to derive the conservation laws. The main advantage of this method includes simplicity and significantly the ability to direct the calculation by setting the order of rows in the initial stoichiometry. However, it has one disadvantage which is the potential numerical instability for large systems. In particular, for large genomic style stoichiometry models [94] that involve many hundreds or even thousands of reactions and species, the method can suffer dramatic failures due to rounding errors during row reduction. There are more robust methods that rely on QR factorization [141] and Singular Value Decomposition (SVD). The main disadvantage of these other methods is that sometimes, depending on the particular algorithm, the row order can not be easily prescribed. Further details can be found in the last section of this chapter as well as the companion text book “Linear Algebra for Systems Biology” [110]. In any event, there are some simple tests one can do to check that the computed conservation laws are correct, one such test will be described next.

## Null Space of $N^T$

To complete this section let us consider in more detail the algebraic nature of the  $Y$  partition in equation (15.7). The elementary matrix,  $E$ , reduced the stoichiometry matrix to a row

echelon form, that is to:

$$\mathbf{E}\mathbf{N} = \begin{bmatrix} \mathbf{M} \\ \mathbf{0} \end{bmatrix} \quad (15.9)$$

The  $\mathbf{E}$  matrix corresponds to the same  $\mathbf{E}$  matrix in equation (15.6), so that we can partition the elementary matrix,  $\mathbf{E}$  row-wise into  $\mathbf{X}$  and  $\mathbf{Y}$  partitions (equation (15.6)).

$$\begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \end{bmatrix} \mathbf{N} = \begin{bmatrix} \mathbf{M} \\ \mathbf{0} \end{bmatrix}$$

From which we can immediately see that:

$$\mathbf{Y}\mathbf{N} = \mathbf{0}$$

Taking the transpose we obtain:

$$\mathbf{N}^T \mathbf{Y}^T = \mathbf{0}$$

The  $\mathbf{Y}$  partition is therefore the null space of the transpose of the stoichiometry matrix <sup>2</sup>. This is a significant result for a number of reasons. It gives a very concise definition of the conservation matrix, but more importantly, it opens up the possibility of using other computational approaches.

The other point of interest is that this result can be used to test whether a set of conservation laws were correctly derived or not. To do this we multiply the transpose of  $\mathbf{N}$  by the transpose of the conservation matrix  $\mathbf{Y}$  and make sure the product equals zero.

Many software packages such as Matlab, Scilab or Mathematica supply commands to compute the null space. This makes it easy to compute the conservation laws by computing the null space of the transpose of the stoichiometry matrix. For example, the following session shows how we can use Scilab to compute the conservation laws for the example matrix we used in previous examples.

```
-->N = [1 0 -1; -1 1 0; 0 1 -1; 0 -1 1]
N =
 1.    0.   - 1.
 0.   - 1.    1.
 - 1.    1.    0.
 0.    1.   - 1.
--> ns = kernel (N')
ans =
```

<sup>2</sup>cf. Chapter 4, Section Computing the Null Space in Introduction to Linear Algebra for Systems Biology (2015), Sauro

```

0.          0.6324555
0.          0.6324555
0.7071068 - 0.3162278
0.7071068  0.3162278
--> // Convert the orthonormal set
--> // into a rational basis using rref
-->rref (ns') '
ans =
1.    0.
1.    0.
0.    1.
1.    1.

```

The null space command in Scilab is `kernel`, in Matlab it is `null`, and in Mathematical it is `NullSpace`. In Python the null space command can be found in the SymPy library by using the `nullspace` command. Like many null space commands implemented in mathematical software, the `kernel` command in Scilab has the drawback of generating an orthonormal set<sup>3</sup>. In order to generate a rational basis we must row reduce the kernel, this results in a more interpretable set of conservation laws. In Matlab it is possible to use the modified null space command, `null (N, 'r')`, which will automatically generate a rational basis (neither Octave or Scilab support this format). Interestingly, Mathematica's (v7.0) null space function does generate a rational basis, however, the algorithm that Mathematica uses is unknown, raising its own issues. More recent version of Mathematica (v11) appear to generate orthonormal sets. The Python implementation based on SymPy does generate a normalized rational basis which is useful, for example:

```

sympy.Matrix.nullspace (sympy.Matrix.transpose (n))
[Matrix([[1], [1], [1], [0]]), Matrix([[0], [1], [0], [1]])]

```

Given that we can now compute the conservation laws for arbitrary networks, one question to consider is whether conservation laws have any behavioral consequences. The answer to this question will be considered in Chapter 16.

## 15.5 Additional Theory – Advanced

In this section we will look at further aspects of conservation laws analysis using a more formal approach. In a later section we will also consider more advanced numerical methods for computing conservation laws.

Let us begin by assuming that the rows of the stoichiometry matrix have been arranged so that the top rows,  $m_o$  include the independent rows and the bottom  $m - m_o$  rows the

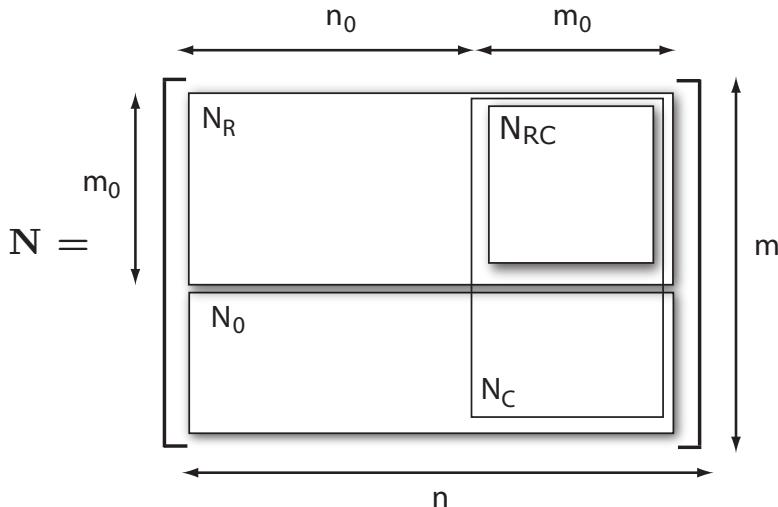
---

<sup>3</sup>An orthonormal set is one where the vectors are orthogonal to each other and of magnitude one.

dependent rows. If we designate the top rows with the symbol  $N_R$  and the bottom rows by  $N_0$  we can write the stoichiometry matrix as:

$$\mathbf{N} = \begin{bmatrix} N_R \\ N_0 \end{bmatrix}$$

where the submatrix  $N_R$  is full rank, and each row of the submatrix  $N_0$  can be derived by is a linear combination of the rows of  $N_R$ . We can also reorder the columns of the stoichiometry matrix of which there will also be  $m_o$  independent columns (column and rows ranks are equal). We will denote the partition of  $\mathbf{N}$  that contains the last  $m_o$  columns, the  $N_C$  matrix. Finally we will designate the partition of  $\mathbf{N}$  that includes only the independent rows and columns the  $N_{RC}$  matrix. The  $N_{RC}$  matrix will be a  $m_o \times m_o$  square invertible matrix.  $N_{RC}$  must be invertible because all rows and columns are independent. The graphical depiction of this partitioning is given in Figure 15.12.



**Figure 15.12** Partitioning of the Stoichiometry Matrix into Four Fundamental Partitions.

If there are no conserved cycles in the network, then the rank ( $\mathbf{N}$ ) =  $m$  (i.e. full rank) and  $\mathbf{N}$  equals  $N_R$ . Following Reder [102], Ehlde [30] and Hofmeyr [55], we make the following construction. Since the rows of  $N_0$  are linear combinations of the rows of  $N_R$  we can define a **link-zero matrix**,  $L_0$  which satisfies:

$$N_0 = L_0 N_R. \quad (15.10)$$

$L_0$  will have dimensions  $(m - m_o) \times m_o$ . We can combine  $L_0$  with the identity matrix – of dimension  $\text{rank}(\mathbf{N})$  – to form the  $m \times m_o$  link matrix,  $L$ , thus:

$$L = \begin{bmatrix} I \\ L_0 \end{bmatrix}$$

When  $\mathbf{N}$  has full rank,  $\mathbf{L}$  equals the identity matrix. Using equation (15.10) and the link matrix we can write:

$$\mathbf{N} = \begin{bmatrix} \mathbf{N}_R \\ \mathbf{N}_0 \end{bmatrix} = \begin{bmatrix} \mathbf{I} \\ \mathbf{L}_0 \end{bmatrix} \mathbf{N}_R = \mathbf{L} \mathbf{N}_R$$

For networks without conserved moieties the  $\mathbf{L}$  matrix reduces to the identity matrix,  $\mathbf{I}$ . If we delete the dependent columns of  $\mathbf{N}$  and  $\mathbf{N}_R$  we obtain:

$$\mathbf{N}_C = \mathbf{L}_0 \mathbf{N}_{RC} \quad \text{or} \quad \mathbf{L} = \mathbf{N}_C \mathbf{N}_{RC}^{-1}$$

By partitioning the stoichiometry matrix into a dependent and independent set we also partition the system equation. The full system equation which describes the dynamics of the network is thus:

$$\begin{bmatrix} \mathbf{I} \\ \mathbf{L}_0 \end{bmatrix} \mathbf{N}_R \mathbf{v} = \frac{d\mathbf{s}}{dt} = \begin{bmatrix} ds_i/dt \\ ds_d/dt \end{bmatrix}$$

where the terms  $ds_i/dt$  and  $ds_d/dt$  refer to the independent and dependent rates of change respectively. From the above equation, we see that

$$\frac{ds_d}{dt} = \mathbf{L}_0 \frac{ds_i}{dt}.$$

Integrating this last equation, we find

$$s_d(t) - s_d(0) = \mathbf{L}_0 [s_i(t) - s_i(0)]$$

for all time  $t$ . Introducing the constant vector  $\mathbf{T} = s_d(0) - \mathbf{L}_0 s_i(0)$ , we can write the above equation as

$$[-\mathbf{L}_0 \quad \mathbf{I}] \begin{bmatrix} s_i \\ s_d \end{bmatrix} = \mathbf{T} \quad (15.11)$$

If  $\mathbf{S} = (s_i, s_d)$ , we can introduce  $\boldsymbol{\Gamma} = [-\mathbf{L}_0 \quad \mathbf{I}]$ , and write this concisely as

$$\boldsymbol{\Gamma} \mathbf{S} = \mathbf{T}$$

We will call  $\boldsymbol{\Gamma}$  the **conservation matrix** and is equivalent to the  $\mathbf{Y}$  matrix in equation (15.7). Each row of the conservation matrix relates to a particular conserved cycle and thus the number of rows indicates the number of conserved cycles in the network. The elements in a particular row indicate which metabolite species contribute to a particular cycle.

The relationship,  $\mathbf{N}_0 = \mathbf{L}_0 \mathbf{N}_R$  (15.10) can be reexpressed in the following form:

$$[-\mathbf{L}_0 \quad \mathbf{I}] \begin{bmatrix} \mathbf{N}_R \\ \mathbf{N}_0 \end{bmatrix} = \mathbf{0} \quad (15.12)$$

However since the conservation matrix,  $\boldsymbol{\Gamma} = [-\mathbf{L}_0 \quad \mathbf{I}]$ , the above relation can be rewritten as:  $\boldsymbol{\Gamma} \mathbf{N} = \mathbf{0}$ . Taking the transpose of this gives us

$$\mathbf{N}^\top \boldsymbol{\Gamma}^\top = \mathbf{0} \quad (15.13)$$

We have already seen this equation in a previous section (15.4) and tells us that the conservation matrix is the null space of the transpose of the stoichiometry matrix. An equivalent way to state this is that the conservation matrix is the *left null space* of the stoichiometry matrix ( $\Gamma \mathbf{N} = \mathbf{0}$ ).

The significance of equation (15.13) is that there are many software tools that allow one to compute the null space very easily. For example Matlab, Mathematica, Maple, O-Matrix, Jarnac or Scilab can easily compute the null space of a matrix and thus derive the conservation laws. Some of these tools however, for example Scilab and Matlab, do not normalize the null space so that a second stage is required, but this is easily accomplished with the command `rref`. Matlab has a variant on the null command, `null (A, 'r')` which generates what is called a rational basis. In Scilab one would enter, `cm = rref (kernel (N')')`. The final transpose that is applied is simply to reorientate the conservation matrix for better viewing. In Jarnac one would enter, `cm = tr (ns (tr (N)))` and so on. One advantage to using Jarnac is that matrices are labeled with the reaction and species names which allows the conservation matrix to be easily interpreted without having to manually identify the columns. The Python based Tellurium [tellurium.analogmachine.org](http://tellurium.analogmachine.org) offers an integrated simulation environment that includes routines to automatically compute these matrices.

Returning once again to the network shown in Figure 15.9, equation (15.11) can be rearranged so that the dependent species can be computed from the independent species, that is:

$$\mathbf{s}_d = \mathbf{L}_0 \mathbf{s}_i + \mathbf{T} \quad (15.14)$$

The complete set of conservation law equations for this model is therefore, equation (15.14):

$$\begin{aligned} \begin{bmatrix} S_1 \\ E \end{bmatrix} &= \begin{bmatrix} -1 & -1 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} S_2 \\ ES \end{bmatrix} + \begin{bmatrix} T_1 \\ T_2 \end{bmatrix} \\ \begin{bmatrix} dS_2/dt \\ dES/dt \end{bmatrix} &= \begin{bmatrix} 1 & 0 & -1 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \end{aligned} \quad (15.15)$$

Note that even though there appears to be four variables in this system, there are in fact only two independent variables,  $\{ES, S_1\}$ , and thus two differential equations and two linear constraints. When solving the system in time, only two differential equations need to be explicitly integrated.

### Scaled $\mathbf{L}$

In metabolic control analysis [63, 102, 37] the link matrix,  $\mathbf{L}$  plays a central role in formulating the sensitivities. In such cases the scaled version of  $\mathbf{L}$ , denoted,  $\mathcal{L}$  is often used.

$\mathcal{L}$  is defined as:

$$\mathcal{L} = (\mathbf{D}^s)^{-1} \mathbf{L} \mathbf{D}^{S_I}$$

where  $\mathbf{D}$  represents a diagonal matrix of either the reciprocals of species,  $\mathbf{D}^s$  or a diagonal of the independent species,  $\mathbf{D}^{S_I}$ . For the previous example,  $\mathcal{L}$  would be given by:

$$\mathcal{L} = \begin{bmatrix} 1/S_2 & 0 & 0 & 0 \\ 0 & 1/ES & 0 & 0 \\ 0 & 0 & 1/S_1 & 0 \\ 0 & 0 & 0 & 1/E \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & -1 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} S_2 & 0 \\ 0 & ES \end{bmatrix}$$

## 15.6 Numerical Methods

---

In a previous section 15.4, a simple method based on forward elimination was described that could be used to derive the conservation laws. This method has a number of advantages but for large matrices can be numerically unstable. In this section we will review alternative methods that, although not always as flexible as forward elimination, are however well suited for the analysis of large matrices.

These methods fall into two groups, three methods based on QR factorization and one method based on Singular Value Decomposition (SVD). The method based on SVD is the simplest and will be described first.

### SVD

Singular Value Decomposition, or SVD is a very useful method for decompiling a matrix into the four orthonormal fundamental subspaces. These subspaces include the range, null space and its transpose. SVD is based on the following factorization:

$$\mathbf{A} = \mathbf{U} \mathbf{S} \mathbf{V}^\top$$

where  $\mathbf{A}$  be a  $m \times n$  matrix of real numbers,  $\mathbf{U}$  is a  $m \times m$  orthonormal matrix,  $\mathbf{V}$  is an  $n \times n$  orthonormal matrix and  $\mathbf{S}$  a  $m \times n$  diagonal matrix with entries  $\sigma_1 \geq \sigma_2 \geq \dots \sigma_p$  where  $p$  is either  $m$  or  $n$ , which ever is the smallest ( $p = \min\{m, n\}$ ). The numbers,  $\sigma_i$  are called the singular values and are positive. The columns of  $\mathbf{U}$  and  $\mathbf{V}$  form the left and right-hand singular vectors.

Of more interest here is the fact that the rows of  $\mathbf{V}$  which correspond to the zero singular values of  $\mathbf{A}$  form an orthonormal basis for the null space of  $\mathbf{A}$ . Therefore one way to obtain the null space of a given matrix is to extract these lower rows from the  $\mathbf{V}$  matrix. The number of rows in  $\mathbf{V}$  that correspond to the null space vectors will equal  $n - r$  where  $r$  is the rank and  $n$  the number of columns of  $\mathbf{A}$ . If there are no zero rows in the  $\mathbf{S}$  matrix then the null space is empty.

**Example 15.5**

Obtain an estimate for the null space of the *transpose* of the following stoichiometry matrix using SVD. Since we will be working on the transpose, the null space vectors will represent the conservation laws.

$$\mathbf{N} = \begin{bmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix}$$

Many math applications such as Scilab or Matlab have svd functions. Here we will use the svd function from Scilab.

```
-->[U, S, V] = svd (N')
V =
-0.316229  -0.707107  0.632456  0.
-0.316229  0.707107  0.632456  0.
-0.632456  0.          -0.316229  0.707107
 0.632456  0.          0.316229  0.707107
S =
 2.236068  0.          0.          0.
 0.          1.7320508  0.          0.
 0.          0.          1.587D-16  0.
U =
-1.886D-16 -0.8164966  0.5773503
-0.7071068  0.4082483  0.5773503
 0.7071068  0.4082483  0.5773503
```

We can extract the null space from  $V^T$ . The number of zero rows in the  $S$  matrix is two, therefore we must extract the bottom two rows of  $V^T$ . This gives us:

```
-->Vt = V'
-->Vt(3:4,1:4)
 0.6324555  0.6324555 -0.3162278  0.3162278
 0.          0.          0.7071068  0.7071068
```

SVD returns an orthonormal basis, to generate a rational basis we apply row reduction to these two rows to yield:

```
-->rref (kk)
ans =
 1.      1.      0.      1.
 0.      0.      1.      1.
```

The transpose of these two vectors is the null space of  $\mathbf{N}^T$ . This can be confirmed by computing the product  $\mathbf{N}^T \mathcal{N}(\mathbf{N}^T)$  and showing that the product equals zero:

$$\begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & 1 & -1 \\ -1 & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

where  $\mathcal{N}(\mathbf{N}^T)$  is the null space. Because there are no row or column exchanges during SVD, the rows in the null space vectors correspond to the same rows in the original matrix,  $\mathbf{N}$ . This makes it easy to identify the individual conservation entries in the conservation law vectors.

We can formalize the SVD algorithm using the following Scilab/Matlab code.

```
// Use SVD to estimate conservation laws
// Operate on the transpose of n

[u, s, v] = svd (n');
vt = v';
nRows = size(vt, 1);
nCols = size(vt, 2);
// Extract bottom nCols(n')-rank orthonormal rows
orthogns = Vt(r+1:nRows,1:nCols);
// Row reduce the transpose to get rational basis
ratns = rref (orthogns)';
// Display Result
ratns'
// Confirm it is the null space, ns should equal 0
ns = n'*ratns
```

Since there are no column or row exchanges during SVD, the order of the rows in the stoichiometry matrix can be used to influence the form of final conservation laws. Just like the row reduction technique, the order of rows in the stoichiometry matrix should be such that any shared species (i.e species containing more than one moiety) be located as close to the bottom of the matrix as possible. This will ensure that negative terms will tend not appear in the final conservation equations.

## QR Factorization

The SVD method given in the last section is an excellent choice for determining the conservation laws. However, it has two downsides, the first is that it is far more computationally intensive than the simple row reduction technique described in 15.4. The second problem with the SVD approach is the need to carry out a final Gauss-Jordan elimination to obtain a rational basis for the conservation laws. Depending on the size of the stoichiometry matrix Gauss-Jordan elimination can be numerically unstable.

Methods that have both excellent stability properties and are less computationally intense than SVD are methods based on QR factorization.

The first QR method to describe is based on computing  $\mathbf{L}_0$ . Any  $m \times n$  matrix can be factored into a product of two matrices  $\mathbf{Q}$  and  $\mathbf{R}$  and a permutation matrix  $\mathbf{P}$ :

$$\mathbf{A}\mathbf{P} = \mathbf{Q}\mathbf{R}$$

$\mathbf{Q}$  is an  $m \times m$  orthogonal matrix, that is  $\mathbf{Q}^T\mathbf{Q} = \mathbf{I}$ ,  $\mathbf{R}$  is a  $m \times n$  upper trapezoidal matrix and  $\mathbf{P}$  a permutation matrix. If  $\mathbf{A}$  is the transpose of the stoichiometry matrix  $\mathbf{N}^T$ , the permutation matrix will also reorder the columns of  $\mathbf{N}^T$  such that the independent columns are on the left and the dependent rows on the right. This is equivalent to reordering the rows in  $\mathbf{N}$ . This partitioning can be written as follows where  $\mathbf{R}$  has been partitioned to match the left side:

$$\mathbf{Q}^T [\mathbf{N}_{\mathbf{R}}^T \quad \mathbf{N}_{\mathbf{0}}^T] = \begin{bmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$$

Note that the partitioned matrix has been absorbed into the reordered  $\mathbf{N}^T$  matrix during the reordering. If we multiply out the terms we obtain:

$$\begin{bmatrix} \mathbf{R}_{11} \\ \mathbf{0} \end{bmatrix} = \mathbf{Q}^T \mathbf{N}_{\mathbf{R}}^T$$

$$\begin{bmatrix} \mathbf{R}_{12} \\ \mathbf{0} \end{bmatrix} = \mathbf{Q}^T \mathbf{N}_{\mathbf{0}}^T$$

Given that  $\mathbf{N}_0 = \mathbf{L}_0 \mathbf{N}_R$ ,  $\mathbf{R}_{12}$  can be rewritten as:

$$\begin{bmatrix} \mathbf{R}_{12} \\ \mathbf{0} \end{bmatrix} = \mathbf{Q}^T \mathbf{N}_{\mathbf{R}}^T \mathbf{L}_0^T$$

so that

$$\begin{bmatrix} \mathbf{R}_{12} \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{R}_{11} \\ \mathbf{0} \end{bmatrix} \mathbf{L}_0^T$$

That is

$$\mathbf{R}_{12} = \mathbf{R}_{11} \mathbf{L}_0^T$$

Since the permutation matrix post-multiplies  $\mathbf{N}^T$ , it means that the columns are reordered, this is reflected in column reordering in the  $\mathbf{R}$  matrix such that all independent columns are moved to the left and dependent columns to the right. Row reduction of the  $\mathbf{R}$  matrix to a reduced echelon form will therefore result in the left partition being transformed into the identity matrix, that is  $\mathbf{R}_{11} = \mathbf{I}$ . From this it follows that the reduced left partition,  $\mathbf{R}_{12} = \mathbf{L}_0^T$ , which is the result we seek:

$$\mathbf{L}_0 = \mathbf{R}_{12}^T$$

By augmenting the  $\mathbf{L}_0$  matrix with an appropriately sized identity matrix we can use this method to generate conservation laws in the standard form, that is in the form  $[-\mathbf{L}_0 \quad \mathbf{I}]$ . This also means that the rows of the stoichiometry matrix will also have been reordered in the process as determined by the permutation matrix obtained from the QR factorization.

Therefore, unlike the row reduction technique or SVD, it is not possible to greatly influence the kind of conservation laws generated by presetting the row order of the stoichiometry matrix although some flexibility still exists. It is still advantageous to make sure that all the shared species are in the bottom rows. The one potential problem with the method is the final Gauss-Jordan elimination, however the reordering of the columns will make this less of an issue.

The Scilab/Matlab code below illustrates an implementation of this method. It is very important to note that the species labels attached to the columns of the conservation matrix is determined by the permutation matrix. This part of the calculation is not shown in the following code.

```
// Use QR to estimate conservation laws via Lo
// Operate on the transpose of n
[qm, rm, p] = qr (n');
nRows = size(n, 1);
nCols = size(n, 2);
mo = rank (n);
m = size(n, 1);
mmo = m - mo;
// Extract bottom nCols-rank orthonormal rows
rt = rm(1:r,1:nRows);
// Row reduce the transpose to get a rational basis
rrt = rref (rt);
Lo = rrt(1:mo,mo+1:nRows)';
// Display Lo
Lo
// Construct the conservation vectors and display
cm = [-Lo eye(mo,mo)];
cm
```

### **Example 15.6**

---

Compute the  $L_0$  matrix of the following stoichiometry matrix using QR factorization.

$$\mathbf{N} = \begin{bmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix}$$

Many software tools offer standard QR factorization. In this example we use Scilab. Applying QR factorization yields the following  $R$  matrix:

```
R =
1.414217 -0.707107 -1.414217 -0.707107
0.          1.224745  0.          -1.224745
0.          0.          0.          0.
```

Since the rank of the stoichiometry matrix is 2, we extract the top two rows from  $\mathbf{R}$  and carry out a complete row reduction (for example by using the `rref()` function) to yield:

$$\begin{aligned} \text{ans} &= \\ 1. &\quad 0. \quad - 1. \quad - 1. \\ 0. &\quad 1. \quad 0. \quad - 1. \end{aligned}$$

The transpose of the  $\mathbf{L}_0$  matrix can be found in the top right corner starting at column  $m_o + 1$  where  $m_o$  equals the number of independent rows in the original stoichiometry matrix. In this case  $m_o$  equals 2, therefore the  $\mathbf{L}_0$  matrix (after transposition) is given by:

$$\begin{matrix} -1 & 0 \\ -1 & -1 \end{matrix}$$

We now combine the negative of this with the identity matrix to obtain the conservation vectors:

$$\begin{matrix} 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{matrix}$$

The only thing that remains is the species labeling for the conservation columns. These can be obtained from the original stoichiometry matrix and the permutation matrix,  $\mathbf{P}$ . As returned by the QR factorization,  $\mathbf{P}$  is given by:

$$\begin{aligned} \mathbf{P} &= \\ 1. &\quad 0. \quad 0. \quad 0. \\ 0. &\quad 0. \quad 1. \quad 0. \\ 0. &\quad 1. \quad 0. \quad 0. \\ 0. &\quad 0. \quad 0. \quad 1. \end{aligned}$$

and the original species order was  $ES, E, S_1, S_2$ . The permutation matrix shows that the new species order should be:  $ES, S_1, E, S_2$ .

---

The final QR method to consider is one based on rank revealing methods, sometimes called RRQR [19]. The specific algebra is described in more detail in the companion book [110] but the method uses the following formula to estimate the null space:

$$\mathbf{A} \mathbf{P} \begin{bmatrix} -\mathbf{R}_{11}^{-1} \mathbf{R}_{12} \\ \mathbf{I} \end{bmatrix} = \mathbf{0} \tag{15.16}$$

This approach is of particular interest because it generates a rational basis for the null space because of the identity matrix in the lower partition. The downside is that it requires an inversion of  $\mathbf{R}_{11}$  but since  $\mathbf{R}_{11}$  is triangular it is possible to exploit widely available and efficient routines for inverting such matrices.

**Example 15.7**

Use the RRQR based method to compute the null space for the transpose of the stoichiometry matrix:

$$\mathbf{N} = \begin{bmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix}$$

From the last example we saw that QR factorization yielded the following  $\mathbf{R}$  matrix:

$$\begin{aligned} \mathbf{R} = \\ 1.414217 & -0.707107 & -1.414217 & -0.707107 \\ 0. & 1.224745 & 0. & -1.224745 \\ 0. & 0. & 0. & 0. \end{aligned}$$

Since the rank of the stoichiometry matrix is 2, we can partition  $\mathbf{R}$  into the following submatrices:

$$R_{11} = \begin{bmatrix} 1.414217 & -0.707107 \\ 0. & 1.224745 \end{bmatrix} \quad R_{12} = \begin{bmatrix} -1.414217 & -0.707107 \\ 0. & -1.224745 \end{bmatrix}$$

We now compute  $-R_{11}^{-1} R_{12}$  to obtain:

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

Combining this with an appropriately sized identity matrix gives the null space:

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Like the previous method we need to be aware of the permutation matrix as this will determine the labels that are associated with the rows of the null space.

There are also ways to obtain the conservation vectors via the  $\mathbf{Q}$  matrix and these are discussed in [141].

For a completely different approach to computing the conservation laws, the reader is referred to the work by Schuster and colleagues. In this work, convex analysis [121] is used to determine the conservation laws and is used primarily to generate conservation laws that only contain (where possible) positive entries.

Most modern simulation applications either use the simpler row reduction technique or more commonly in recent years, they use the QR factorization technique based on estimating the  $\mathbf{L}_0$  matrix [141].

Method	Advantages	Disadvantages
Row Reduction	a) Simple b) Fast c) Row Order	Potential numerical instabilities
SVD	a) Robust b) Expensive on large systems	Requires one final Gauss-Jordan step
QR by $L_0$	a) Robust b) Faster than SVD	Requires one final Gauss-Jordan step
QR by RRQR	a) Robust b) Row order	No Gauss-Jordan step required

**Table 15.1** Comparison of different approaches to computing conservation laws.

## 15.7 Design of Simulation Software

One practical implication of moiety conservation concerns the design of software for simulation and analysis. Two issues arise, one concerns increasing simulation efficiency by reducing the number of differential equations and the second concerns numerical stability by removing the dependent species from a model.

### Reduced Systems

The first concern is straight forward, instead of solving the full set of systems equations many simulators instead solve the following reduced set:

$$\begin{aligned} s_d &= L_0 s_i + T \\ \frac{dS_i}{dt} &= N_R v(s_i, s_d) \end{aligned} \tag{15.17}$$

In these equations,  $s_i$  is the vector of independent species,  $s_d$ , the vector of dependent species,  $L_0$  the link matrix,  $T$  the total mass vector,  $N_R$  the reduced stoichiometry matrix and  $v$  the rate vector. This modified equation (15.17) constitutes the most general expression for a differential equation based temporal model [55, 53]. Equations (15.5) shows a typical reduced system. Note that in these equations the dependent species are first computed from the dependent species. This is followed by the evaluation of the reduced set of differential equations. The order is crucial. The total amounts,  $T$ , can be computed at the start of a simulation by using equation (15.14) and the initial conditions. In multi-compartmental

systems where the size of compartments may differ, it is important to sum the amounts not concentrations.

One obvious advantage reducing the model is that it lessens the computationally burden to solving the full set of differential equations. Many biochemical simulation packages will automatically check for moiety conservations and perform this simplification before performing any analysis of the system equations. This is especially important for large models. For example, in one the *E. coli* model obtained from the BiGG repository <http://bigg.ucsd.edu/>, approximately five percent of the differential equations are redundant, that is they can be safely eliminated from the model by using moiety conservation constraints.

## Multicompartment Systems

Up to now we have not mentioned the fact that many models may include multiple compartments, that is separate volume spaces where the movement of mass between volumes is via specific transporter proteins. The literature is not very clear or extensive in discussing the modeling of multicompartment systems however one crucial point to bear in mind when considering conservation laws that cross compartments is that the sum must be with respect to the total mass. For convenience models will often assume a unit volume for a compartment such that any conserved cycles within the compartment are expressed as the sum of concentrations. In such situations it is easy to forget that what is actually conserved is in fact mass not concentration. In general a conservation law is therefore expressed in the form:

$$\sum V_i s_i = T$$

where  $V_i$  is the volume that the concentration of species  $S_i$  resides.

## Numerical Stability

Although simplifying a model by eliminating the dependent species can offer speed improvements to simulations, the most important reason for model reduction is the gain in numerical stability. One of the most important metrics that arises often in the analysis of pathways (or any dynamical system for that matter) is the Jacobian matrix.

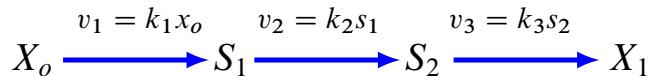
The Jacobian matrix is an  $m \times m$  matrix of partial derivatives of the rates of change with respect to the species, that is:

$$\mathbf{J} = \frac{\partial}{\partial \mathbf{s}} \left( \frac{ds}{dt} \right)$$

For example, for a simple linear chain shown in Figure 15.13 the differential equations are given by:

$$\frac{ds_1}{dt} = v_1 - v_2$$

$$\frac{ds_2}{dt} = v_2 - v_3$$

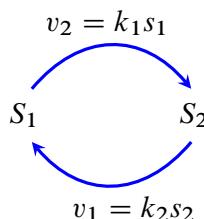


**Figure 15.13** Linear Chain of Three Reactions.

The Jacobian matrix is then given by:

$$\mathbf{J} = \begin{bmatrix} \frac{\partial(v_1 - v_2)}{\partial s_1} & \frac{\partial(v_1 - v_2)}{\partial s_2} \\ \frac{\partial(v_2 - v_3)}{\partial s_1} & \frac{\partial(v_2 - v_3)}{\partial s_2} \end{bmatrix} = \begin{bmatrix} -k_2 & 0 \\ k_2 & -k_3 \end{bmatrix}$$

The Jacobian is used in many ancillary calculations, for example, solving differential equations (particularly stiff equations), solving for the steady state, calculating sensitivities, frequency analysis, certain optimization algorithms and others. In many of these cases the calculation involves the inversion of the Jacobian. In the case of the linear pathway, there will always be an inverse so long as the rate constants are non-zero. However if we consider a simple cycle such as the one shown below:



then the Jacobian matrix is given by:

$$\mathbf{J} = \begin{bmatrix} \frac{\partial(v_1 - v_2)}{\partial s_1} & \frac{\partial(v_1 - v_2)}{\partial s_2} \\ \frac{\partial(v_2 - v_1)}{\partial s_1} & \frac{\partial(v_2 - v_1)}{\partial s_2} \end{bmatrix} = \begin{bmatrix} -k_1 & k_2 \\ k_1 & -k_2 \end{bmatrix}$$

This shows that the row dependencies in the stoichiometry matrix reappear as dependencies in the Jacobian. This means that the Jacobian cannot be inverted and any calculations that require the inversion of the Jacobian will fail. The solution is to work with the reduced model, this eliminates the dependent species from the stoichiometry matrix which in turn makes sure that the Jacobian is once again invertible.

## Further Reading

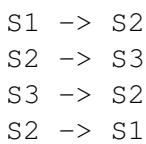
---

1. Reich, JG, and Selkov EE (1981) Energy metabolism of the cell. Academic Press, London.
2. Hofmeyr JH, Kacser H, van der Merwe KJ. (1986) Metabolic control analysis of moiety-conserved cycles. Eur J Biochem. 155(3):631-41.
3. Sauro HM and Ingalls B (2004) Conservation analysis in biochemical networks: computational issues for software writers, Biophys Chem, 109, 1–15.
4. Sauro HM (2015) Introduction to Linear Algebra for Systems Biology. Ambrosius Publishing
5. Vallabhajosyula RR, Chickarmane V and Sauro HM (2006) Conservation analysis of large biochemical networks, Bioinformatics, 22(3), 346-353.
6. Cornish-Bowden and Hofmeyr J-HS (2002) The Role of Stoichiometric Analysis in Studies of Metabolism: An Example J. theor. Biol 216, 179-191.
7. Ehlde M, and Zacchi G (1997) Chemical Engineering Science 52:2599–2606(8).

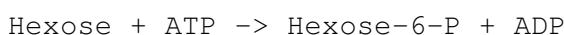
## Exercises

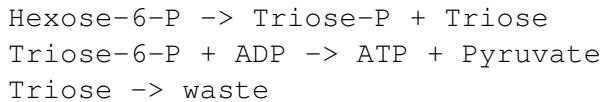
---

1. Define the terms *moiety* and *conserved moiety*.
2. Write the differential equations for the following model and show that some of the rates of change are dependent on others:



3. The following paper: Eisenthal R, and Cornish-Bowden A. 1998. Journal of Biological Chemistry 273 (10):5500-5505, describes a model of Trypanosome energy metabolism. Read the paper and identify the conserved moiety cycles described in the model.
4. Analyse the following model using matrix reduction techniques to determine the conserved cycle. Assume Hexose is a fixed species.





5. Use Matlab, Scilab or Octave to determine the conserved cycles in the model by Eisenthal R, and Cornish-Bowden A. 1998. Journal of Biological Chemistry 273 (10):5500-5505.

## Proofs

---

A general non-symmetric,  $m \times n$  matrix  $\mathbf{A}$  can be as a product of two matrices  $\mathbf{Q}$  and  $\mathbf{R}$  as:

$$\mathbf{A} \mathbf{P} = \mathbf{Q} \mathbf{R}$$

where  $\mathbf{Q}$  is an  $m \times m$  orthogonal matrix such that  $\mathbf{Q}^T \mathbf{Q} = \mathbf{I}$ ,  $\mathbf{P}$  is an  $m \times n$  permutation matrix that indicates column changes in  $\mathbf{A}$  and  $\mathbf{R}$  is an  $m \times n$  upper trapezoidal matrix.

# 16

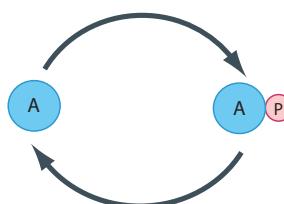
## *Moiety Conserved Cycles*

### 16.1 Moiety Conserved Cycles

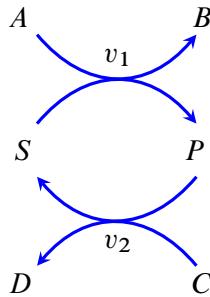
---

In this chapter the topic of moiety conserved cycles and their impact on behavior will be examined. There are some more advanced concepts in this chapter which require some prerequisite knowledge of bistability.

Chapter 15 introduced the idea of moiety conceived cycles. To recap, Figure 16.1 shows a typical moiety-conserved cycle. In this cycle there are two species, one might be a protein and the other, the phosphorylated protein. If we assume that the synthesis and degradation rates for the protein is low in comparison to the dynamics of phosphorylation and dephosphorylation, then we can assume that the total amount of protein, phosphorylated plus unphosphorylated is constant over the period of study. Figure 16.2 shows a more abstract form of a moiety-conserved cycle.



**Figure 16.1** Phosphorylation and Dephosphorylation Cycle forming a Moiety Conservation Cycle between Unphosphorylated (left species) and Phosphorylated protein (right species).



**Figure 16.2** Simple Conserved cycle where  $s + p = \text{constant}$ .

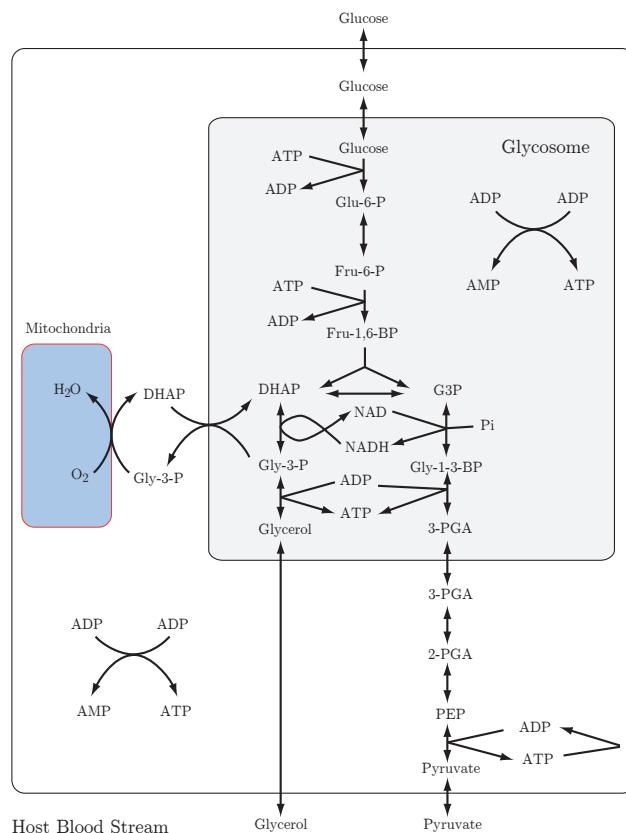
### Constraining Species Levels

One of the simplest effects of a moiety-conserved cycle is that it puts upper limits on the concentrations of the participants. In the simple cycle shown in Figure 16.2 where the total number of moles in the cycle is fixed at  $s + p$ , the upper limit that either  $S$  or  $P$  can reach is  $s + p$ . This effect was made very clear in a study of glycolysis in *Trypanosoma brucei* (Figure 16.3). What is unusual about the pathway is that much of glycolysis resides in a single membrane organelle called the glycosome. Many of the metabolites in the glycosome are phosphorylated, for example, glucose-6-phosphate, glyceraldehyde-3-phosphate and it is this that creates a constraint on the level of phosphate. In addition to the glycosome, the mitochondrion of *Trypanosoma brucei* appears to do very little other than oxidize glycerol 3-phosphate via oxygen utilization.

What is of interest is that an analysis of the network using the techniques described in Chapter 15 indicates the presence of four conservation laws, these include:

- 1:  $\text{ATP}_c + \text{ADP}_c + \text{AMP}_c$
- 2:  $\text{ATP}_g + \text{ADP}_g + \text{AMP}_g$
- 3:  $\text{NAD}_g + \text{NADH}_g$
- 4:  $\text{glycerol 3-phosphate}_c + \text{dihydroxyacetone phosphate}_c + \text{glycerol 3-phosphate}_g + \text{dihydroxyacetone phosphate}_g + \text{glucose 6-phosphate}_g + \text{fructose 6-phosphate}_g + \text{fructose 1,6-bisphosphate}_g + \text{glyceraldehyde 3-phosphate} + 1,3\text{-bisphosphoglycerate} + \text{ATP}_g + \text{ADP}_g$

where the subscript  $c$  means cytoplasm and  $g$  means glycosome. Figure 16.4 shows the same metabolic map as Figure 16.3 but with the conserved moieties highlighted with ellipses. The fact that phosphate is a conserved moiety means that any species that includes the moiety will be constrained by the total amount of phosphate. As pointed out by Eisenthal and

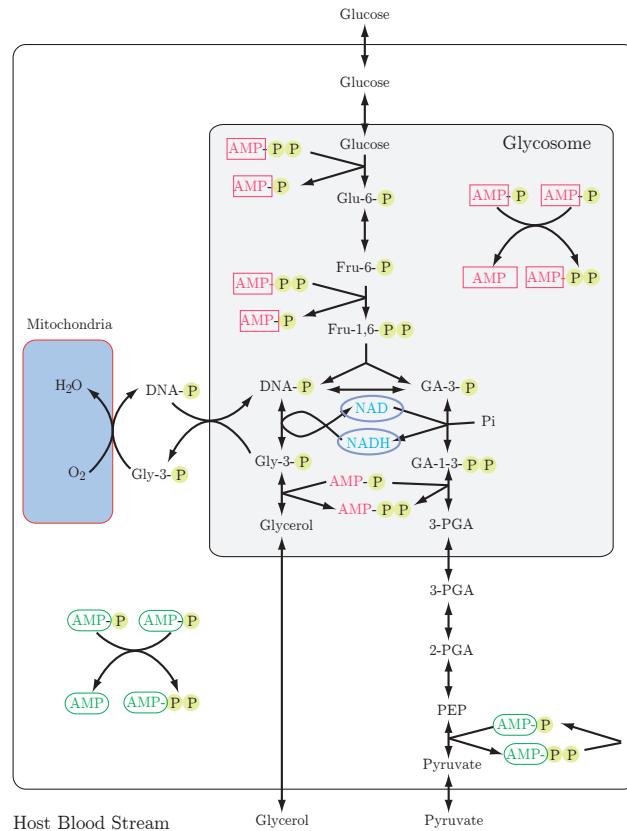


**Figure 16.3** Energy Metabolism of *Trypanosoma brucei*

Cornish-Bowden [31] in relation to the work of Bakker et al [5], there are two possible ways to disrupt an organism metabolically. One can either reduce a flux to a very low level or increase one or more metabolite levels to such high levels that they become toxic. Bakker's analysis [5] of *Trypanosoma* metabolism showed that much of the flux control was on glucose transport. This limits the number of potential sites for flux disruption. As for disrupting concentrations, only one step had a significant concentration control coefficient, Pyruvate transport [31] again limiting the choice for drug targets. The reason why pyruvate transport is a susceptible target is because it is one of the few steps not involved in the conservation laws.

## 16.2 MCA of Conserved Cycles

The MAPK (mitogen-activated protein kinase) pathways, are highly conserved and common components in signal transduction pathways (Chang & Karin, 2001). Virtually all eukaryotic cells that have been examined (ranging from yeast to man) possess multiple MAPK

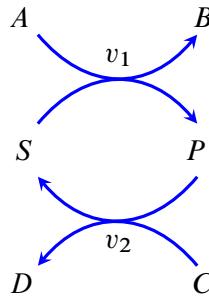


**Figure 16.4** Energy Metabolism of *Trypanosoma brucei*

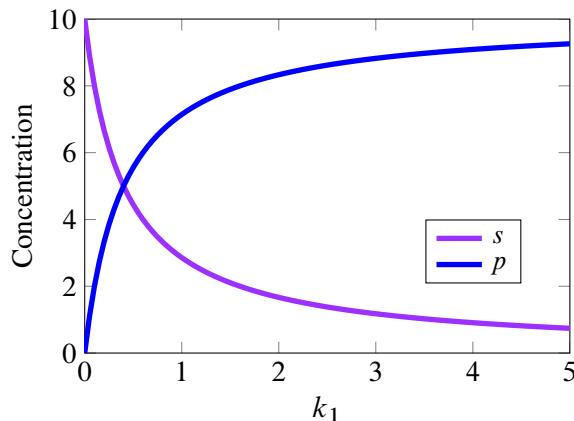
pathways each of which responds to multiple inputs. In mammalian systems MAPK pathways are activated by a wide range of input signals including a variety of growth factors and environmental stresses such as osmotic shock and ischemic injury (Kyriakis & Avruch, 2002; Gomperts et al., 2002). Once the MAPK pathways have integrated these signals, they coordinately activate gene transcription with resulting changes in protein expression leading to cell cycling, cell death and cell differentiation.

Consider the simple conserved cycle shown in Figure 16.5. As discussed in Chapter 15, the two species,  $S$  and  $P$  are conserved because the total  $s + p$  remains constant over time (at least over a time scale shorter than protein synthesis and degradation). Let us assume that the kinetics governing each cycle arm is simple first order mass-action kinetics.

If we plot the steady state concentration of  $S$  and  $P$  versus the first-order kinetic constant  $k_1$  we get the response curves shown in Figure 16.6. The response curves are in fact hyperbolic. For example,  $P$  rises linearly then levels off to 10 concentration units in the limit. As  $k_1$  increases more and more  $S_1$  is converted to  $P$  leading to a rise in  $P$  and a fall in  $S$ . The limit is reached because there is only a fixed amount of mass in the cycle.



**Figure 16.5** Simple Conserved cycle where  $s + p = \text{constant} = T$ .

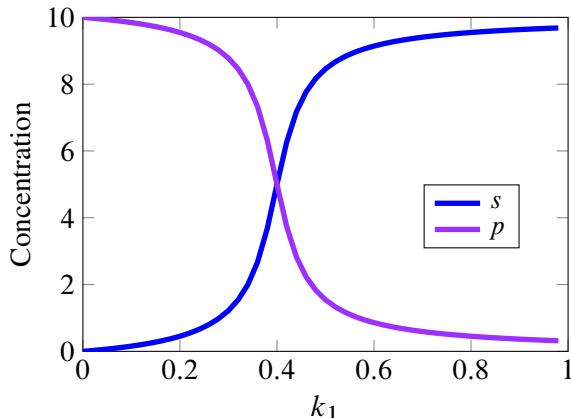


**Figure 16.6** Simulation of the simple cycle with linear kinetics. Plot shows the steady state concentration of each species as a function of  $k_1$ . Model:  $S \rightarrow P; k_1 \cdot S; P \rightarrow S; k_2 \cdot P; S=10; k_1=0.1; k_2=0.4$

### Simple Cycle with Non-Linear Kinetics

If we now take the simple cycle model from the last section and instead of linear kinetics we now use non-linear kinetics, for example Michaelis-Menten kinetics on the forward and reverse arms then additional changes in behavior will be observed.

The response is now sigmoidal rather than hyperbolic. The reason for this is explained in Figure 16.8. The intersection points marked by a grey marker represents the corresponding steady state point ( $v_1 = v_2$ ). A perpendicular dropped from these indicates the corresponding steady state concentration of  $S_1$ . If the activity of  $v_1$  is increased by increasing  $k_1$  by 20% then the  $v_1$  curve moves up. The left intersection point indicates how much the steady state concentration moves as a result, shown by  $\Delta s$ . The closer the steady state point is to the saturated point of the curve, the more the steady state will move. This shows that the response in  $S_1$  can be very sensitive in changes in  $k_1$ . Because  $k_1$  is a linear term in the



**Figure 16.7** Simulation of the simple cycle with non-linear kinetics illustrating sigmoid or ultrasensitive behavior. Model:  $S \rightarrow P; k_1 \cdot S / (K_m + S); P \rightarrow S; k_2 \cdot P / (K_m + P); S = 10; k_1 = 0.1; K_m = 0.5; k_2 = 0.4; K_m = 0.5$

rate law we could replace it with the concentration of the enzyme implied in the Michaelis-Menten law. In practice such a cycle could represent a phosphorylation/dephosphorylation cycle where the implied enzyme is now a kinase. The kinase in turn could be controlled by other processes so that changes in the kinase activity result in sigmoid (or switch like) behavior in the cycle dynamics. In the literature such behavior was studied by [45, 46] and has been observed experimentally [58].

The sigmoid behavior observed in a moiety conserved cycle has been termed **ultrasensitivity** and is defined in terms of the gain between an input and output. Traditionally the gain of a biological system, especially in the context of a cooperative system, has been the fold change in ligand required to change the response from 10% to 90% of maximum, termed the response coefficient [45]<sup>1</sup>. It is defined as:

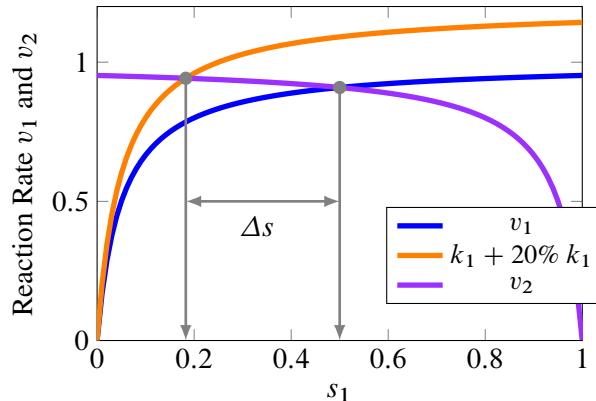
$$R = \frac{S_{0.9}}{S_{0.1}}$$

where  $S_{0.1}$ , for example, is the ligand concentration required to reach 10% of the maximum response. Using this definition, it is possible [109] to relate  $R$  to the standard Hill equation (2.13) such that if:

$$v = V_m \frac{s^n}{K_H^n + s^n} \quad \text{then} \quad R = 81^{1/n}$$

If  $n = 1$  then the system is hyperbolic and is considered not to be ultrasensitive. Qualitatively it means that the ligand concentration must change 81 fold in order to go from 10% to 90% of the response, a much more sensitive change. A system with a Hill coefficient of 4 will

<sup>1</sup>Not to be confused with the response coefficient defined in MCA.



**Figure 16.8** Plots the two cycle rates,  $v_1$  and  $v_2$  for the simple cycle with non-linear kinetics. Model:  $S_1 \rightarrow S_2; k_1 \cdot S_1 / (K_m + S_1); S_2 \rightarrow S_1; k_2 \cdot S_2 / (K_m + S_2); S_1 = 1; k_1 = 1; K_m = 0.05; k_2 = 1; K_m = 0.05$ . The intersection points marked by a grey marker represents the steady state point ( $v_1 = v_2$ ). See main text for explanation.

have a response,  $R$ , equal to 3. That it a ligand must change 3 fold in order to go from 10% to 90% of the response. A response,  $R$ , less than 81 indicates that the system is ultrasensitive and tends to one as the sigmoidicity reaches its asymptotic limit. Although experimentally accessible, the definition of  $R$  is a little ad-hoc and instead it is more useful from the point of view of theory to define the gain as the ratio of the fractional change between an output,  $Y$ , and input,  $X$ . That is:

$$R_X^Y = \frac{d \ln Y}{d \ln X}$$

For example, if the input,  $X$ , is changed by 1% and this results in an output change in  $Y$  of 2% then the input is amplified two fold. In engineering this effect is often referred to as the **gain** of the system. Amplification of signals is a fundamental operation in both biology and engineering. Many inputs in biology are small magnitude that they must be amplified before they can be acted upon. Examples include changes in hormonal levels to and amplification of sensory inputs from external stimuli such as nutrient gradients or light levels in the retina.

It should be apparent that  $R_X^Y$  is clearly related to the response coefficient introduced in section 8.3. In contrast to the ratio,  $S_{0.9}/S_{0.1}$ , a system is ultrasensitive if  $R_X^Y > 1$ . In the next section the response coefficient,  $R_X^Y$ , will be determined for a cycle in terms of the cycle elasticities.

## 16.3 Using MCA to Understand Ultrasensitivity

It is possible to use the machinery of metabolic control analysis derive the conditions for ultrasensitivity [123] in a covalent modification cycle without recourse to specific kinetic laws [45, 46].

Consider again the simple cycle shown in Figure 16.2. To investigate ultrasensitivity, the concentration control coefficient,  $C_{e_1}^p$  will be evaluated, that is how sensitivity  $P$  is to changes in the activity of the enzyme (often a kinase) that catalyzes the conversion of  $S$  to  $P$ . There are different but related ways to approach this derivation. The first is to look at the effect of perturbations on the cycle.

To make matters simpler, assume that both  $v_1$  and  $v_2$  are irreversible and not product inhibited. Let us make a small change,  $\delta e_1$ , to  $e_1$  that catalyzes the reaction rate  $v_1$ . We can write down the two local equations as follows:

$$\frac{\delta v_1}{v_1} = \frac{\delta e_1}{e_1} + \varepsilon_s^1 \frac{\delta s}{s} \quad \frac{\delta v_2}{v_2} = \varepsilon_p^2 \frac{\delta p}{p}$$

At steady state  $\delta v_1/v_1 = \delta v_2/v_2$ , in addition, the changes in  $s$  and  $p$  must be constrained by  $s + p = T$ . This means that it must be true that  $\delta s = -\delta p$ . Substituting  $\delta s$  with  $\delta s$  and equating the two local equations yields:

$$\frac{\delta e_1}{e_1} + \varepsilon_s^1 \frac{(-)\delta p}{s_1} = \varepsilon_p^2 \frac{\delta p}{p}$$

This equation can be rearranged to:

$$\frac{\delta e_1}{e_1} = \varepsilon_s^1 \frac{\delta p}{p} \frac{p}{s} + \varepsilon_p^2 \frac{\delta p}{p} = 0$$

Collecting  $\delta p/p$

$$\frac{\delta e_1}{e_1} = \frac{\delta p}{p} \left( \varepsilon_s^1 \frac{p}{s} + \varepsilon_p^2 \right)$$

Dividing both sides by  $\delta e_1/e_1$  and noting that  $(\delta p/p)/(\delta e_1/e_1)$  is equal to  $C_{e_1}^p$ , we obtain after rearrangement:

$$C_{e_1}^p = \frac{s}{p\varepsilon_s^1 + s\varepsilon_p^2} \tag{16.1}$$

Instead of working with actual concentrations, divide by the total,  $T$  to use molar fractions,  $M_s = s/T$  and  $M_p = p/T$ :

$$C_{e_1}^p = \frac{M_s}{M_p\varepsilon_s^1 + M_s\varepsilon_p^2} \tag{16.2}$$

The equation can be modified to indicate how an arbitrary effector of  $v_1$  influences  $P$ . Recalling the response coefficient relationship (8.14), the response of an effector  $X$  on the concentration,  $p$  is given by:

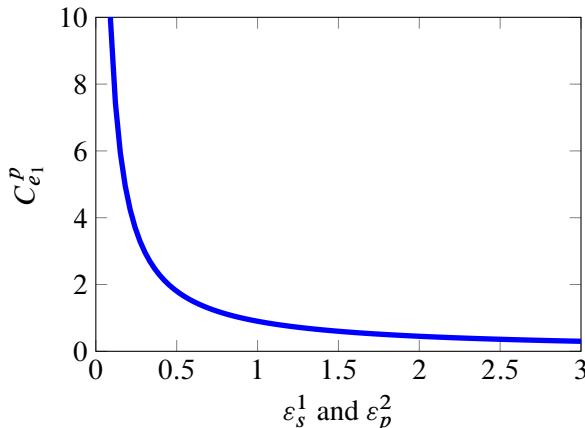
$$R_x^p = C_{e_1}^p \varepsilon_x^{v_1} = \varepsilon_x^{v_1} \frac{M_s}{M_p\varepsilon_s^1 + M_s\varepsilon_p^2} \tag{16.3}$$

If it is assumed that  $v_1$  and  $v_2$  operate far below saturation than both elasticities, then  $\varepsilon_s^1$  and  $\varepsilon_p^2$  are approximately one, that is:

$$C_{e_1}^p = \frac{M_s}{M_p + M_s}$$

This value will always be less than or equal to one. Therefore there is no possibility of ultrasensitivity when both enzymes are operating in their first-order regime. However if the enzymes are operating near saturation, then  $\varepsilon_s^1 < 1$  and  $\varepsilon_p^2 < 1$ . For example, if  $s = 9$  and  $p = 1$  and both elasticities are equal to 0.5 then the control coefficient,  $C_{e_1}^p = 1.8$ . If we reduce the elasticities further  $\varepsilon_s^1 = 0.2$  and  $\varepsilon_p^2 = 0.2$ , the response of the system rises to 4.5. That is a 1% increase in  $E_1$  will result in a 4.5% increase in  $P$  (Figure 16.9) In most cases like this,  $P$  is itself a protein, often a kinase. So that a small change in one protein,  $E_1$  can have a large effect on another,  $P$ .

The advantage of this approach is that it is kinetic mechanism independent. That is, in formulating (16.2) nothing was assumed about the details of the mechanism by which the cycle steps operated. The only thing that was assumed was the local sensitivity of the reaction rate to their respective substrates.



**Figure 16.9** Plot of  $C_{e_1}^p$  as a function of the two cycle elasticities,  $\varepsilon_s^1$  and  $\varepsilon_p^2$ . Both elasticities have the same value on the  $x$  axis. The plot shows the rapid rise in sensitivity as the elasticities fall.

To complete the discussion we can also derive the effect  $e_1$  has on the cycling flux,  $J$ . To do this we recall that:

$$\frac{\delta v_2}{v_2} = \varepsilon_p^2 \frac{\delta p}{p}$$

At steady state,  $\delta v_2/v_2$  is equal to the change in the steady state cycling flux,  $\delta J/J$ . Dividing both sides by  $\delta e_1/e_1$  we obtain:

$$C_{e_1}^J = \varepsilon_p^2 C_{e_1}^p = \varepsilon_p^2 \frac{M_s}{M_p \varepsilon_s^1 + M_s \varepsilon_p^2}$$

This relationship shows that the cycling flux is less sensitive to  $e_1$  than  $p$ . This is because the elasticity,  $\varepsilon_p^2$  will be small when the cycle is operating in the ultrasensitive mode.

$C_{e_1}^s$ ,  $C_{e_2}^s$  and  $C_{e_1}^p$  can be derived in a similar manner. A more formal approach to deriving  $C_{e_1}^p$ , is to write down the dynamical equations for the system and implicitly differentiate with respect to the parameter of interest.

**Derivation by Implicit Differentiation** We can also derive the  $C_{e_1}^p$  term using implicit differentiation. We start by defining the differential equation for  $P$ :

$$\frac{ds}{dt} = v_1(e_1, s(e_1)) - v_2(p(e_1))$$

We assume the rate of reaction for  $v_1$  is a function of  $E_1$  which is the kinase catalysing  $v_1$  and the unphosphorylated protein  $S$ . We assume that  $P$  has no effect on  $v_1$ . The key observation here is that because  $S$  is the dependent variable we must replace it with the function  $s_1(e_1)$ , indicating that the concentration of  $S_1$  is also a function of  $E_1$ . Likewise for  $v_2$  we assume it is only a function of  $P$  which in turn is also a function of  $E_1$ . Once in this form we set the equation to zero to indicate steady state and differentiate the equation with respect to our chosen parameter  $E_1$ :

$$0 = \frac{\partial v_1}{\partial e_1} + \frac{\partial v_1}{\partial s} \frac{ds}{de_1} - \frac{\partial v_2}{\partial p} \frac{dp}{de_1}$$

Noting the conservation law  $s + p = T$  and differentiating this with respect to  $e_1$  we can state that:

$$\frac{ds}{de_1} = -\frac{dp}{de_1}$$

Substituting  $ds/de_1$  with  $-dp/de_1$  we obtain:

$$0 = \frac{\partial v_1}{\partial e_1} - \frac{\partial v_1}{\partial s} \frac{dp}{de_1} - \frac{\partial v_2}{\partial p} \frac{dp}{de_1}$$

Multiplying both sides by  $e_1$ , dividing both sides by  $p$  and noting that at steady state  $v_1 = v_2$  and dividing both sides by  $v_1$  or  $v_2$  as appropriate we can rewrite the above as:

$$0 = \frac{\partial v_1}{\partial e_1} \frac{e_1}{v_1} - \frac{\partial v_1}{\partial s} \frac{s}{v_1} \frac{p}{s} \frac{dp}{de_1} \frac{e_1}{p} - \frac{\partial v_2}{\partial p} \frac{p}{v_2} \frac{dp}{de_1} \frac{e_1}{p}$$

Translating the terms into elasticities and  $C_{e_1}^p$  we obtain:

$$0 = \varepsilon_{e_1}^{v_1} - \varepsilon_s^{v_1} \frac{p}{s_1} C_{s_1}^p - \varepsilon_p^{v_2} C_{e_1}^p$$

Letting  $\varepsilon_{e_1}^{v_1} = 1$  and solving for  $C_{e_1}^p$  we obtain:

$$C_{e_1}^p = \frac{1}{\varepsilon_s^{v_1} \frac{p}{s_1} - \varepsilon_p^{v_2}}$$

Multiplying top and bottom by  $s_1$  we arrive at:

$$C_{e_1}^p = \frac{s}{\varepsilon_s^{v_1} p - \varepsilon_p^{v_2} s}$$

This is the same result as equation (16.1).

### Sensitivity of the Cycle to Changes in $T$

Let the total amount of species in the cycle by  $T = s + p$ . Make a change in  $T$ ,  $\delta T$ , by adding externally some  $S$  or  $P$ . This will cause both rates,  $v_1$  and  $v_2$  to change as well as changes to the steady state levels of  $S$  and  $P$ . To make matters simpler, assume as before that both  $v_1$  and  $v_2$  are irreversible and not product inhibited, this allows us to write the following statements:

$$\frac{\delta v_1}{v_1} = \varepsilon_s^1 \frac{\delta s}{s}, \quad \frac{\delta v_2}{v_2} = \varepsilon_p^2 \frac{\delta p}{p}$$

Note that there have been no changes to  $e_1$  or  $e_2$ . At steady state the changes  $\delta v_1/v_1$  and  $\delta v_2/v_2$  will equal each other. Therefore:

$$\varepsilon_s^1 \frac{\delta s}{s} - \varepsilon_p^2 \frac{\delta p}{p} = 0$$

Given that  $T$  has been changed by  $\delta T$ , it must be that case that:  $\delta T = \delta s + \delta p$ . That is  $\delta s = \delta T - \delta p$ . Substituting this into the previous equation to eliminate  $\delta s$ , we obtain:

$$\varepsilon_s^1 \frac{\delta T - \delta p}{s} - \varepsilon_p^2 \frac{\delta p}{p} = 0$$

Dividing both sides by  $\delta T/T$  and noting that in the limit  $(\delta p/p)/(\delta T/T) = C_T^p$  we obtain after rearrangement:

$$C_T^p = \frac{T \varepsilon_s^1}{\varepsilon_s^1 p + \varepsilon_p^2 s}$$

As before the absolute concentrations in  $s$  and  $p$  can be converted to molar fractions by dividing top and bottom by  $T$  to yield:

$$C_T^p = \frac{\varepsilon_s^1}{\varepsilon_s^1 M_p + \varepsilon_p^2 M_s} \quad (16.4)$$

The sensitivity of the cycling flux with respect to  $T$  can be obtained as follows. The change in the cycling rate as a result of increasing  $T$  by  $\delta T$  is given by:

$$\frac{\delta v_2}{v_2} = \varepsilon_p^2 \frac{\delta p}{p}$$

However  $\delta v_2/v_2$  is also the change in the cycling flux,  $J$ . Dividing both sides of the equation by  $\delta T/T$  and substituting  $C_T^p$  that was derived above yields:

$$C_T^J = \varepsilon_p^2 C_T^p = \varepsilon_p^2 \frac{\varepsilon_s^1}{\varepsilon_s^1 M_p + \varepsilon_p^2 M_s} \quad (16.5)$$

## 16.4 Cycle Connectivity Theorems

---

Moiety conserved cycles possess modified connectivity theorems. This is because the species in the cycle cannot have arbitrary values but are constrained by the total number of moles in the cycle [39, 57].

Consider again the simple cycle shown in Figure 16.2. Assume that a small change,  $\delta s_1$  is made in  $s_1$ . In order to maintain a constant amount of  $s + p$ , we also make a compensating change in  $\delta p$  equal to  $-\delta s$ . These changes will cause both  $v_1$  and  $v_2$  to change, however we can make changes to  $E_1$  and  $E_2$  such that the reaction rates are unchanged. We can express this thought experiment using the following local equations:

$$\frac{\delta v_1}{v_1} = \frac{\delta e_1}{e_1} + \varepsilon_s^1 \frac{\delta s}{s} + \varepsilon_p^1 \frac{\delta p}{p} = 0$$

$$\frac{\delta v_2}{v_2} = \frac{\delta e_2}{e_2} + \varepsilon_p^2 \frac{\delta p}{p} + \varepsilon_s^2 \frac{\delta s}{s} = 0$$

To make things simpler let us assume that each forward reaction is product insensitive, that is  $\varepsilon_1^2 = 0$  and  $\varepsilon_2^1 = 0$  so that the local equations are now reduced to:

$$\frac{\delta e_1}{e_1} = -\varepsilon_s^1 \frac{\delta s}{s}$$

$$\frac{\delta e_2}{e_2} = -\varepsilon_p^2 \frac{\delta p}{p}$$

We can also express the thought experiment in terms of the systems equation:

$$\frac{\delta p}{p} = C_{e_1}^p \frac{\delta e_1}{e_1} + C_{e_2}^p \frac{\delta e_2}{e_2}$$

We can now substitute the local equations into the systems equation:

$$-\frac{\delta p}{p} = C_{e_1}^p \varepsilon_s^1 \frac{\delta s}{s} + C_{e_2}^p \varepsilon_p^2 \frac{\delta p}{p} \quad (16.6)$$

The thought experiment added the constraints that  $\delta s_1 = -\delta s_2$ , so that:

$$-\frac{\delta p}{p} = -C_s^p \varepsilon_s^1 \frac{\delta p}{p} \frac{s}{s} + C_{e_2}^p \varepsilon_p^2 \frac{\delta p}{p}$$

Cancelling the  $\delta s_2/s_2$  terms yields the covalent modification connectivity theorem:

$$-1 = -C_{e_1}^p \varepsilon_s^1 \frac{p}{s} + C_{e_2}^p \varepsilon_p^2$$

One way to reexpress this is to divide both sides by  $p$  and rearrange so that:

$$\frac{1}{p} = C_{e_1}^p \varepsilon_s^1 \frac{1}{s} - C_{e_2}^p \varepsilon_p^2 \frac{1}{p} \quad (16.7)$$

The above equation is the modified connectivity theorem for the moiety conserved cycle. It is possible to combine this equation with the summation theorem (which is unmodified):

$$C_{e_1}^p + C_{e_2}^p = 0$$

and solving for  $C_{e_1}^p$ . This is an alternative way to determine  $C_{e_1}^p$ .

### Other Relationships

There are a number of other relationships related to moiety conserved cycles that are worth mentioning. Consider a conserved moiety cycle at steady state of  $m$  species,  $S_1, S_2, \dots, S_m$  constrained by the conservation law,  $s_1 + s_2 + \dots + s_m = T$ . Let us make a small change to an enzyme that catalyzes one of the reaction steps in the cycle. For example, change  $e_i$  by an amount  $\delta e_i$ . This will result in the system changing to a new steady state where changes in  $S_1, S_2$ , etc are observed. Since the total number of moles in the cycle has not changed, it must be true that:

$$\delta s_1 + \delta s_2 + \dots + \delta s_m = 0$$

Scaling by each species yields:

$$s_1 \left( \frac{\delta s}{s} \right) + s_2 \frac{\delta s_2}{s_2} + \dots + s_m \frac{\delta s_m}{s_m} = 0$$

Finally dividing each term by the relative change in  $e_i$ :

$$s_1 \left( \frac{\delta s_1 / s_1}{\delta e_i / e_i} \right) + s_2 \left( \frac{\delta s_2 / s_2}{\delta e_i / e_i} \right) + \dots + s_m \left( \frac{\delta s_m / s_m}{\delta e_i / e_i} \right) = 0$$

which can be rewritten as:

$$s_1 C_{e_i}^{s_1} + s_2 C_{e_i}^{s_2} + \dots + s_m C_{e_i}^{s_m} = 0$$

Note the difference with the usual concentration summation theorem where the summation is over a single species. Here the summation is over all cycle species. The relationship can be easily modified to use fractional molar amounts instead:

$$M_1 C_{e_i}^{s_1} + M_2 C_{e_i}^{s_2} + \dots + M_m C_{e_i}^{s_m} = 0$$

where  $M_i = s_i / T$ .

Similar to the previous relationship there is also a summation with respect to  $R_T^{s_i}$ . Consider making a change to the total number of moles in a cycle, by an amount  $\delta T$ . This will result in a change to the steady state such that:

$$\delta s_1 + \delta s_2 + \dots + \delta s_m = \delta T$$

As before, each term can be scaled by the corresponding  $s_i$  amount and each term divided by the relative change in  $T$ :

$$\frac{s_1}{T} \left( \frac{\delta s_1 / s_1}{\delta T / T} \right) + \frac{p}{T} \left( \frac{\delta s_2 / s_2}{\delta T / T} \right) + \dots + \frac{s_m}{T} \left( \frac{\delta s_m / s_m}{\delta T / T} \right) = 1$$

which can be simplified to:

$$M_1 C_T^{s_1} + M_2 C_T^{s_2} + \dots + M_m C_T^{s_m} = 1$$

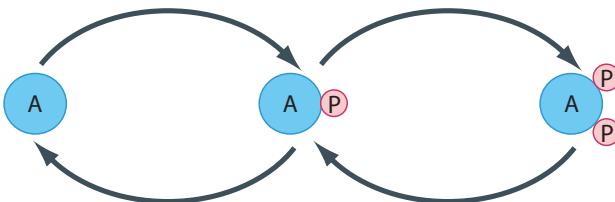
The two results are summarised below:

$$M_1 C_{e_i}^{s_1} + M_2 C_{e_i}^{s_2} + \dots + M_m C_{e_i}^{s_m} = 0 \quad (16.8)$$

$$M_1 C_T^{s_1} + M_2 C_T^{s_2} + \dots + M_m C_T^{s_m} = 1 \quad (16.9)$$

## Dual Cycles

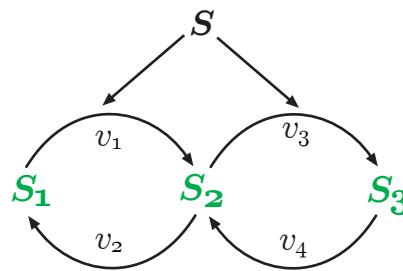
In many signaling pathways, for example the MAPK pathway, one will often find dual phosphorylation cycles. That is, cycles where a protein is double phosphorylated, often in a strict order.



**Figure 16.10** A Dual Phosphorylation and Dephosphorylation Cycle.

Consider a dual cycle shown in Figure 16.11 where  $S$  is a signal (eg a kinase) that catalyzes  $v_1$  and  $v_3$ . The stoichiometry matrix for the dual cycle is given by:

$$\mathbf{N} = \begin{bmatrix} -1 & 1 & 0 & 0 \\ 1 & -1 & -1 & 1 \\ 0 & 0 & 1 & -1 \end{bmatrix}$$



**Figure 16.11** Two cycles connected by a common intermediate,  $S_2$ . The rate laws for each step is given by  $v_1 = k_1 s_1$ ,  $v_2 = k_2 s_2$ ,  $v_3 = k_3 s_2$ ,  $v_4 = k_4 s_3$ .  $S$  is the stimulus signal which acts by increasing  $k_1$  and  $k_3$  by the same factor.

For a dual cycle it is possible to show that there is one conservation law given by the relation:

$$s_1 + s_2 + s_3 = T$$

If we assume simple linear mass-action kinetics for each of the reactions, simulation will reveal that the concentration of  $S_3$  shows sigmoid behavior with respect to the stimulus signal  $S$ . We can assume that the stimulus signal,  $S$ , operates on the rate constants,  $k_1$  and  $k_3$  by the same factor, that is an increase in  $S$  by  $x\%$  results in a change in  $k_1$  and  $k_3$  by  $x\%$ . What is interesting about this case is that a sigmoid response can be generated from first-order reactions. This can be called first-order ultrasensitivity to distinguish it from zero-order ultrasensitivity that is generated from a single cycle.

The control coefficient,  $C_s^{s_3}$  can be derived as follows. To keep things simple it will be assumed that the reactions are irreversible and not product inhibited. The strategy is to first derive the terms  $C_{e_1}^{s_3}$  and  $C_{e_3}^{s_3}$ . The response to signal  $S$  is the sum of  $C_{e_1}^{s_3}$  and  $C_{e_3}^{s_3}$ . It is assumed that the elasticity of  $S$  towards each reaction is one, which is reasonable if the signal is a protein catalyst such as a kinase.

The derivation will be illustrated for  $C_{e_1}^{s_3}$ .  $C_{e_3}^{s_3}$  can be derived in a similar manner by perturbing  $e_3$ . To derive  $C_{e_1}^{s_3}$ , perturb  $e_1$  by an amount  $\delta e_1$ . This will change the steady state from which the following local equations can be obtained:

$$\begin{aligned}\frac{\delta v_1}{v_1} &= \varepsilon_p^1 \frac{\delta s}{s} + \frac{\delta e_1}{e_1}, & \frac{\delta v_2}{v_2} &= \varepsilon_2^2 \frac{\delta s_2}{s_2} \\ \frac{\delta v_3}{v_3} &= \varepsilon_2^3 \frac{\delta s_2}{s_2}, & \frac{\delta v_4}{v_4} &= \varepsilon_3^4 \frac{\delta s_3}{s_3}\end{aligned}$$

At steady state  $v_1 = v_2$  and  $v_3 = v_4$ , though it is not necessarily the case that  $v_1 = v_3$ . This means that when the steady state changes  $\delta v_1 = \delta v_2$  and  $\delta v_3 = \delta v_4$ . In relative terms we state that:  $\delta v_1/v_1 = \delta v_2/v_2$  and  $\delta v_3/v_3 = \delta v_4/v_4$ . For example, by equating the local

equations  $\delta v_1/v_1$  and  $\delta v_2/v_2$  we obtain:

$$\varepsilon_{s_2}^1 \frac{\delta s}{s} + \frac{\delta e_1}{e_1} = \varepsilon_2^2 \frac{\delta s_2}{s_2}$$

Both sides of the equation can be divided by  $\delta e_1/e_1$  to give:

$$\varepsilon_{s_2}^1 C_{e_1}^{s_1} + 1 = \varepsilon_2^2 C_{e_1}^{s_2}$$

The same can be applied to the  $v_3, v_4$  pair of local equations. As a result we have two equations in three unknowns  $C_{e_1}^{s_1}, C_{e_1}^{s_3}$ , and  $C_{e_1}^{s_3}$ . To solve for the three unknowns, a third equation is necessary. The dual cycle has a single conservation equation,  $s_1 + s_2 + s_3 = T$ . Perturbing  $e_1$  by  $\delta e_1$  does not disturb the total  $T$  but will change the distribution of species such that the change in species must be constrained by  $\delta s_1 + \delta s_2 + \delta s_3 = 0$ . Scaling each term:

$$s_1 \frac{\delta s}{s} + s_2 \frac{\delta s_2}{s_2} + s_3 \frac{\delta s_3}{s_3} = 0$$

and diving throughout by  $\delta e_1/e_1$  yields:

$$s_1 C_{e_1}^{s_1} + s_2 C_{e_1}^{s_2} + s_3 C_{e_1}^{s_3} = 0$$

Note that this is the same equation as (16.9). We now have three equations in three unknowns which can be solved. For example,  $C_{e_1}^{s_3}$  is given by:

$$C_{e_1}^{s_3} = \frac{s_1 \varepsilon_2^3}{s_1 \varepsilon_3^4 \varepsilon_2^2 + s_2 \varepsilon_3^4 \varepsilon_1^1 + s_3 \varepsilon_1^1 \varepsilon_2^3}$$

Using the same technique, a solution to  $C_{e_3}^{s_3}$  can also be found. The sum,  $C_{e_1}^{s_3} + C_{e_3}^{s_3}$ , as indicated previously, will yield the total response of  $s_3$  due to changes in the signal  $s$ . The sum is given by the equation below (16.10):

$$C_s^{s_3} = \frac{s_1(\varepsilon_2^3 + \varepsilon_2^2) + s_2 \varepsilon_1^1}{s_1 \varepsilon_2^2 \varepsilon_3^4 + s_2 \varepsilon_1^1 \varepsilon_3^4 + s_3 \varepsilon_1^1 \varepsilon_2^3} \quad (16.10)$$

Equation (16.10) looks a little complicated but can be simplified by assuming all reactions are first-order. Under these conditions all the elasticities equal one so that the equation reduces to something much more manageable::

$$C_s^{s_3} = \frac{2s_1 + s_2}{s_1 + s_2 + s_3}$$

This shows that given the right ratios for  $s_1, s_2$  and  $s_3$ , it is possible for  $C_s^{s_3} > 1$ . The maximum value the equation can reach is when  $s_2$  and  $s_3$  are zero, at this point  $C_{e_1}^{s_3} = 2$ .

Unlike the case of a single cycle where near saturation is required to achieve ultrasensitivity, multiple cycles can achieve ultrasensitivity with simple linear kinetics

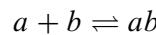
## 16.5 Sequestration

The cyclic models considered so far assume negligible sequestration of the cycle species by the catalyzing kinase and phosphatase. In reality this is not likely to be the case because experimental evidence indicates that the concentrations of the catalyzing enzymes and cycle species are comparable (See [10] for a range of illustrative data). In such situations additional effects are manifest [40, 107], of particular interest is the emergence of new regulatory feedback loops which can alter the behavior quite markedly (See [81] and [93]).

Interlocked cycles that also incorporate sequestration effects are likely to be able to display an extremely wide range of behaviors. This is an area of theoretical analysis has not received much attention in the literature [145].

### Ultrasensitivity via Sequestration

The simplest mechanism that can generate an ultrasensitivity response involves the dimerization of two dissimilar species. The effect depends on a conservation law between the active participant,  $A$  and a sequester molecule  $B$ . The reaction is simply:

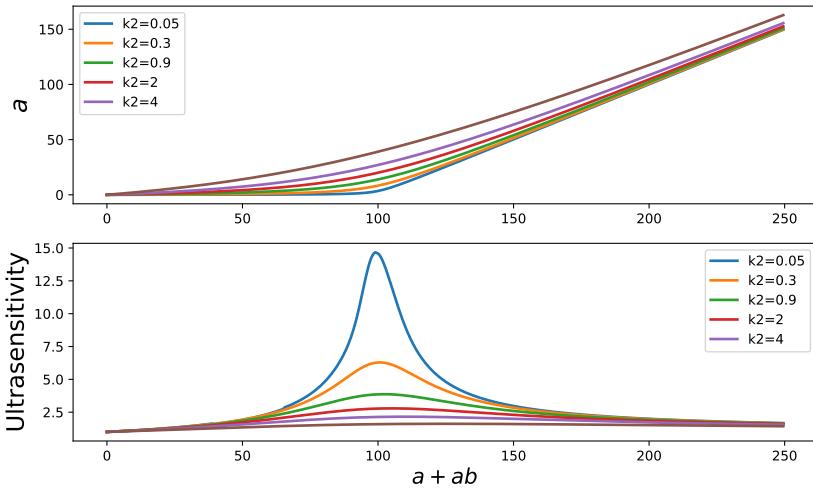


The response begins with a low level of active molecule  $A$ . As  $A$  is increased, most of the added  $A$  is sequestered by  $B$ . However eventually enough  $A$  is added so that there remains very little of  $B$  left that can sequester  $A$ . At this point the concentration of  $A$  rises rapidly. This is the point where ultrasensitivity is seen. Figure 16.12 illustrates a plot that shows ultrasensitivity near 80 and 120 on the  $x$  axis [15, 14].

To illustrate another example where conservation laws contribute to new behavior, we will look at a very simple linear pathway where there is a dead-end leak caused by complex formation (see Figure 16.13). The observed ultrasensitivity is in response to a change in the stimulus signal and originates from a combination of kinetic and conservation factors. Sigmoid behavior can be observed in both the free species,  $X$  and the complex  $XI$  forms. Saturation in the level of  $XI$  is due to a conservation law involving the  $I$  moiety. To achieve a saturating effect in  $X$ , the second step,  $v_2$  should be modeled using a Michaelis-Menten rate law (itself based on a conservation law between free enzyme and enzyme substrate complex) and the first step,  $v_1$  should be reversible to ensure that a steady state exists at high stimulus levels ( $X$  would go to infinity otherwise). Figure 16.14 shows an example simulation that illustrates ultrasensitivity in this more complex sequestration model.

### The Markevich Switch

The next example will illustrate a fairly complex set of interlinked conservation laws that leads to quite elaborate behavior. This system, first discovered by Kholodenko and co-



**Figure 16.12** Ultrasensitivity seen in a simple sequestration model.  $a + b \rightleftharpoons ab$ ;  $k_1 \cdot a \cdot b - k_2 \cdot ab$ ;  $a = 0.001$ ;  $b = 100$ ;  $ab = 0$ ;  $k_1 = 0.4$ ;  $k_2 = 1$ . The Tellurium script that generated this plot can be found at 16.1 for details.

workers *et al.* will be referred to as the **Markevich Switch** after the first author on the original paper [81].

The system involves a double cycle but with secondary sequestration effects occurring on the limbs. Figure 16.15 illustrates the full pathway. The model describes the catalysis of the conversion of  $S_1$  through two enzyme catalyzed reactions,  $v_1$  and  $v_2$ . The individual catalytic cycles are made explicit in this model, that is, the binding of  $S_1$  to enzyme  $E_1$  to form complex and dissociation to form product,  $S_2$  is explicitly modeled. In addition there is the reverse conversion of  $S_3$  back to  $S_1$ , again by a sequence of two enzyme catalyzed reactions,  $v_3$  and  $v_4$  again in explicit form. The stimulus,  $S$ , acts by adding more total  $E_1$  to the upper limbs.

This pathway has multiple conservation laws stemming from the two different enzymes and a separate substrate cycle. These conservation laws include:

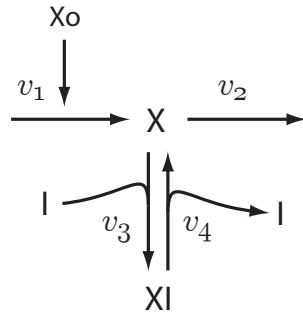
$$S_1 + S_2 + S_3 + ES_1 + ES_2 + ES_3 + ES_4 = T_1 \quad (16.11)$$

$$E_1 + ES_1 + ES_2 = T_2 \quad (16.12)$$

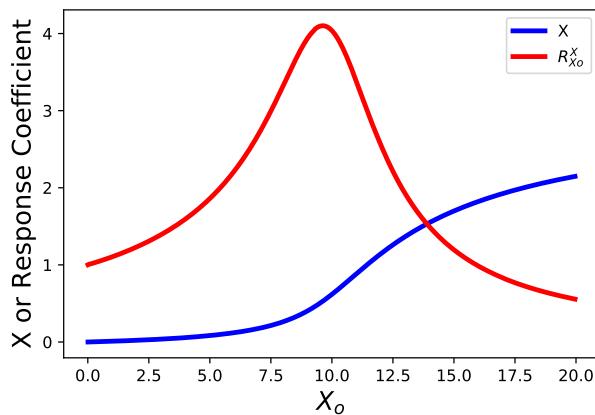
$$E_2 + ES_3 + ES_4 = T_3 \quad (16.13)$$

Figure 16.17 illustrates graphically the three conservation laws. The presence of the conservation laws leads to the emergence of bistable behavior.

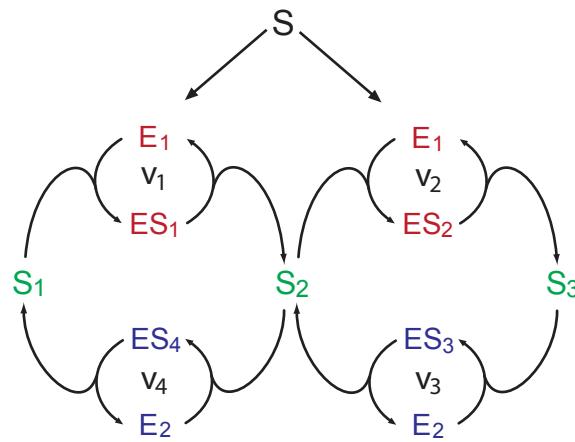
That is, given a particular set of parameters, there exists three possible steady states, two stable and one unstable (sometimes called metastable). We can see this depicted in the steady



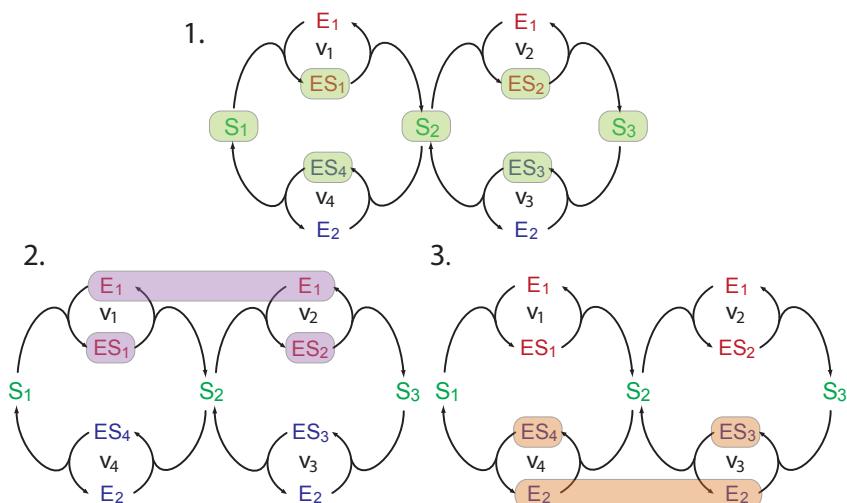
**Figure 16.13** Sequestration steps attached to a pathway with a signal input  $X_o$  to  $v_1$ .



**Figure 16.14** Simulation of model shown in Figure 16.13. The degree of ultrasensitivity is measured by the response coefficient which achieved a maximum of four (upper curve). The Tellurium script that generated this plot can be found at 16.2.

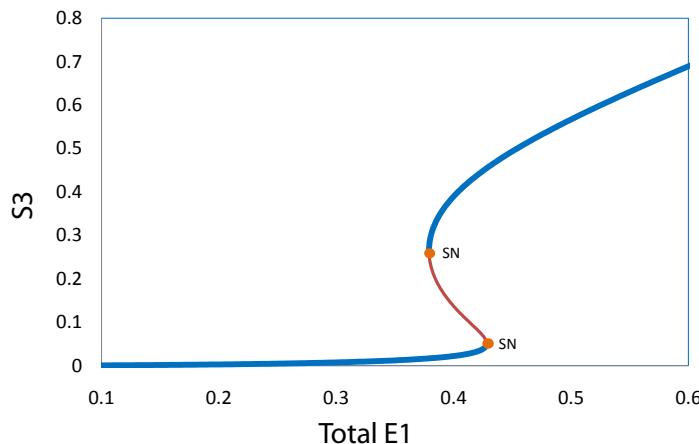


**Figure 16.15** A complex interlinked set of conserved cycles that describes the Markovitch switch [81].  $S$  controls the activity of the pathway by controlling the amount of total  $E_1$ .



**Figure 16.16** Three conservation laws in the extended dual cycle network with the conserved species highlighted in ellipses.

state plot (Figure 16.17) that shows the concentration of  $S_3$  versus total  $E_1$  ( $E_1 + ES_1$ ). At a certain range of total  $E_1$ , the curve shows three possible steady states. A high stable state, a low stable state and an intermediate unstable state (thin line in the graph). In principle the unstable state could be achieved and maintained indefinitely but random fluctuations at the molecular level would move the network to one of the two stable steady state. The question is how does this come about since there is no obvious positive feedback in the network?

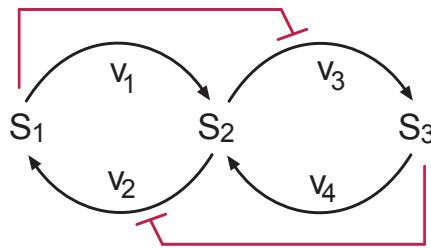


**Figure 16.17** Bifurcation plot illustrating bistability in the concentration of  $S_3$  as a function of  $E_1$ . The symbol SN indicates a turning point, i.e. a change in stability. Thick lines represent stable branches and the thinner central line an unstable branch. Simulations were carried out by the Oscill8 Tool ([oscill8.sourceforge.net](http://oscill8.sourceforge.net)), the model was obtained from [99] as a SBML file via the BioModels Database (<http://www.ebi.ac.uk/biomodels-main/>).

A major part of the answer lies in the constraints imposed by the conservation laws. Consider the following scenario. If the activity of the two forward limbs,  $v_1$  and  $v_2$  is increased, this will cause more  $S_2$  and  $S_3$  to be made. These changes have a number of consequences. To begin with, the additional  $S_3$  will bind to more  $E_2$  to form complex  $ES_3$ . However because  $ES_3$  is linked by way of a conservation law (16.13) to the levels of  $ES_4$  and  $E_2$ , these concentrations will therefore decline. This effectively makes  $S_3$  compete with  $S_2$  for  $E_2$ . The result is that there is less  $E_2$  to catalyze  $v_4$  resulting in an effective inhibition of  $v_4$  by  $S_3$ . This kind of inhibition has been called apparent regulation because there is no direct molecular mechanism involved, it is simply an effect brought about by competitive sequestration. There are other factors in play here as well, for example the degree of saturation (see [81] for details), however the constraints imposed by the conservation laws

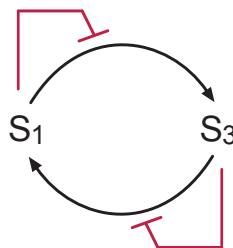
are critical to the observed bistability.

To continue, given that  $S_2$  and  $S_3$  have both increased then  $S_1$  is likely to have decreased (16.12). If this is the case then there is less binding of  $S_1$  to  $E_1$ . This results in a greater availability of  $E_1$  which can be used to increase  $v_2$ . If we invert the logic here then we see that **increases** in  $S_1$  will lead to **decreases** in  $v_2$ . This is another example of apparent regulation due to conservation law constraints, in this case due to equation (16.12). We can therefore redraw the pathway in a more simplified way as depicted in Figure 16.18).



**Figure 16.18** Two apparent regulatory loops in the Markevich pathway.

We can simplify this diagram even further by removing the central link,  $S_2$  to give the diagram shown in Figure 16.19. This shows more clearly the opposing repression loops that surround the pathway.

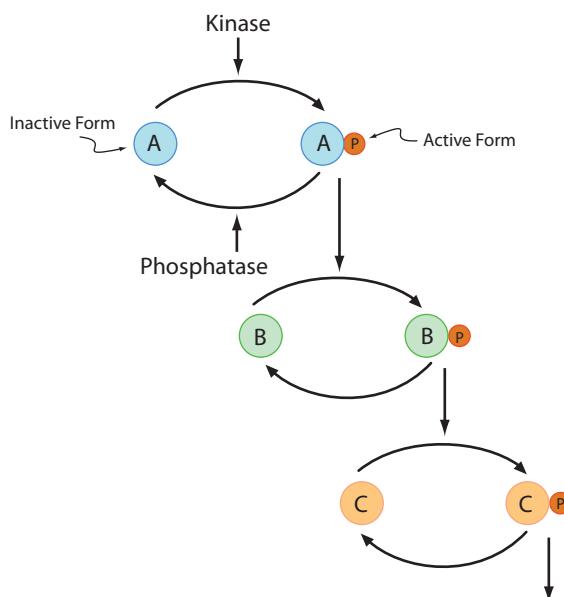


**Figure 16.19** A highly simplified version of the Markevich pathway showing the opposing repression loops that surround the pathway.

In essence, what we have here is a toggle switch. Consider the states that can possibly exist in the pathway shown in Figure 16.19. If the concentration of  $S_1$  is low then this relieves the inhibition on the forward limb this converting  $S_1$  into  $S_3$  and thus maintaining  $S_1$  in the low state.  $S_3$  is now at a higher concentration and its effect is to repress the low limb. This state of affairs is therefore stable. If on the other hand we start  $S_1$  at a high concentration, the reverse logic applies. The forward limb is now repressed this stabilizing  $S_1$  at its high state. In contrast  $S_3$  must now be at a low concentration where the repression it apply to the lower limb is now released thus stabilizing it's low level.

## 16.6 Cascades

It is common in eukaryotic organisms to find cascades of covalent modification cycles. A single cycle will comprise of two or more interconvertible forms of a signaling protein, often tagged with phosphate groups to distinguish the forms. Phosphatases and kinases are responsible for dephosphorylating or phosphorylating the different forms. One or more of the forms within a cycle will be active meaning that they are themselves capable of phosphorylating proteins in other cycles (Figure 16.20). A sequence of such cycles is referred as a cascade. Although Figure 16.20 shows only a single phosphorylation event per cycle, in many cases there will be multiple phosphorylation events.



**Figure 16.20** A cascade of phosphorylation/dephosphorylation cycles.

For the purpose of this study we will focus only on cycles that involve a single phosphorylation event. In order to keep track of the various entities and terms in the subsequent analysis a specific notation will be used. A layer will be defined as a single cycle with a cascade. For example Figure 16.21 illustrates a cascade with two layers. Layers will be numbered from top to bottom. The top layer will be layer one, the second layer, layer two and so on.

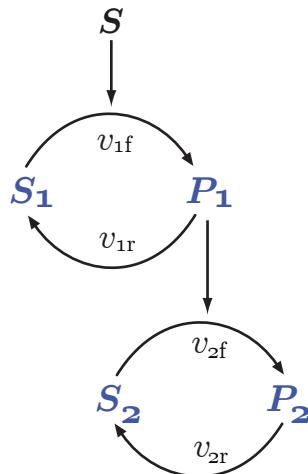
Within a layer there is a cycle comprising two protein forms termed  $S$  and  $P$ . The  $P$  form refers to the phosphorylated and active form (active in the sense that it can phosphorylate other proteins). To two forms will have subscripts to denote which layer they belong to. For example,  $S_2$  and  $P_2$  refers to the protein forms in layer two.

The phosphorylation and dephosphorylation steps are also indicated using a specific notation. The rate of the phosphorylation step, or forward step, will be designated  $v_{if}$  where  $i$  will

refer to the particular layer it belongs to. The rate of the dephosphorylation step, or reverse step, will be designated  $v_{ir}$  where  $i$  refers to the layer it belongs to.

For example,  $v_{2f}$  and  $v_{2r}$  refer to the forward and reverse rates for the steps in layer two. Figure 16.21 shows a cascade made from two cycles that uses this notation.

Finally, the relative molar amounts of protein in a given layer will be designated  $M_{si}$  or  $M_{pi}$  where  $i$  is the corresponding layer.  $M_{si} = s_i/T_i$  and  $M_{pi} = p_i/T_i$  where  $T_i$  is the total amount of protein in the  $i^{\text{th}}$  layer.



**Figure 16.21** Two cycles in a cascade.

### Effect of Modulating the Input Signal

The signal  $S$  alters the activity of  $v_{1f}$ , in this case by phosphorylating  $S_1$ . The actual mechanism of action is unimportant at this stage. The effectiveness of the signal on  $v_{1f}$  can be summarised by the elasticity  $\varepsilon_s^{v_{1f}}$ . As before, sequestration effect of signal binding are ignored in this model.

One question to ask is what is the sensitivity of  $P_2$  to changes in signal  $S$ ? To answer this question,  $R_s^{P_1}$  and  $R_s^{P_2}$  will be evaluated in turn. We already know from a previous section (16.2), that:

$$C_{e_1}^{P_1} = \frac{M_{s_1} \varepsilon_{e_1}^{P_1}}{M_{p_1} \varepsilon_{s_1}^1 + M_{s_1} \varepsilon_{p_1}^2} \quad (16.14)$$

Consider now evaluating  $R_s^{P_2}$ .  $v_3$  is a function of  $p_1$  and  $s_2$  and  $v_4$  a function of  $p_2$ . We assume irreversibility in all reaction steps to make matters simpler and also assume negligible sequestration of  $P_1$  by binding with  $S_2$ , that is:

$$v_{2f} = f(s_2, p_1) \quad \text{and} \quad v_{2r} = g(p_2)$$

Making a perturbation to the signal,  $S$  by  $\delta s$ , the relative change in  $v_{2f}$  and  $v_{2r}$  is given by:

$$\frac{\delta v_{2f}}{v_{2f}} = \frac{\partial v_{2f}}{\partial s_2} \frac{p_1}{v_{2f}} \frac{\delta s_2}{p_1} + \frac{\partial v_{2f}}{\partial p_1} \frac{p_1}{v_{2f}} \frac{\delta p_1}{p_1}$$

$$\frac{\delta v_{2r}}{v_{2r}} = \frac{\partial v_{2r}}{\partial p_2} \frac{p_2}{v_{2r}} \frac{\delta p_2}{p_2}$$

At steady state the change in rates will be equal,  $\delta v_{2f}/v_{2f} = \delta v_{2r}/v_{2r}$ . In addition, due to conservation of  $S_2$  and  $P_2$ ,  $\delta s_2 = -\delta p_2$ . Therefore:

$$\frac{\delta v_{2r}}{\delta p_2} \frac{p_2}{v_{2r}} \frac{\delta p_2}{p_2} = \frac{\partial v_{2f}}{\partial p_1} \frac{p_1}{v_{2f}} \frac{\delta p_1}{p_1} + \frac{\partial v_{2f}}{\partial s_2} \frac{s_2}{v_{2f}} \frac{(-)\delta p_2}{p_2} \frac{p_2}{s_2}$$

Divide both sides by  $\delta s/s$ :

$$\varepsilon_{p_2}^{2r} R_s^{p_2} = \varepsilon_{p_1}^{2f} R_s^{p_1} - \varepsilon_{s_2}^{2f} R_s^{p_2} \frac{p_2}{s_2}$$

Solving for  $R_s^{p_2}$  are replacing the absolute amounts by relative amounts yields:

$$R_s^{p_2} = \frac{M_{s_2} \varepsilon_{p_1}^{2f} R_s^{p_1}}{M_{s_2} \varepsilon_{p_2}^{2r} + M_{p_2} \varepsilon_{s_2}^{2f}} \quad (16.15)$$

The term  $(M_{s_2} \varepsilon_{p_1}^{2f})/(M_{s_2} \varepsilon_{p_2}^{2r} + M_{p_2} \varepsilon_{s_2}^{2f})$  will be recognized from (16.3) as the sensitivity of  $p_2$  to changes in  $p_1$ . Following [67] we will refer to these ‘local’ sensitivities using the expression:

$$r_{p_1}^{p_2} \quad (16.16)$$

For a cascade with multiple layers, the local sensitivity of the  $i^{\text{th}}$  layer with respect to the proceeding signal will be denoted by:

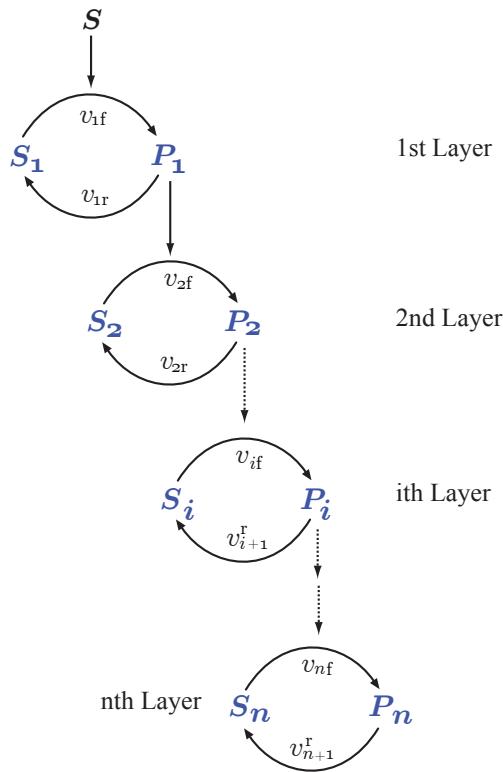
$$r_{p_{i-1}}^{p_i}$$

Equation (16.15) can be generalized for any length cascade so that the sensitivity of the  $i^{\text{th}}$  layer to changes in the proceeding signal is given by:

$$r_{p_{i-1}}^{p_i} = \frac{M_{s_i} R_{s_{i-1}}^{p_{i-1}} \varepsilon_{p_{i-1}}^{vir}}{M_{s_i} \varepsilon_{p_i}^{vir} + M_{p_i} \varepsilon_{s_i}^{vir}}$$

To illustrate this formalism consider the two cycle layer in Figure 16.21. Let the overall sensitivity of  $P_2$  to changes in the signal equal  $R_s^{p_2}$ . Since there are two layers there will be two local sensitivities,  $r_s^{p_1}$  and  $r_{p_1}^{p_2}$ . To obtain the overall sensitivity, we form the product of the local sensitivities:

$$R_s^{p_2} = r_s^{p_1} r_{p_1}^{p_2} \quad (16.17)$$



**Figure 16.22** Arbitrary number of cycles or layers in a cascade.

For a cascade of  $n$  layers (Figure 16.22), the overall sensitivity will be:

$$R_s^{p_n} = r_s^{p_1} r_{p_1}^{p_2} r_{p_2}^{p_3} \dots r_{p_{n-1}}^{p_n}$$

This derivation assumes that there are no regulatory feedback or feedforward loops and that there is no sequestration between cycles. It should be apparent that if each cycle is operating in an ultrasensitivity mode where an individual  $r$  will be high, the overall sensitivity will be significantly higher. For example, if we consider cascade of three cycles, where each individual cycle sensitivity,  $r$  is of the order of 4.0, then the overall sensitivity will be  $4 \times 4 \times 4 = 64$ .

### Effect of Modulating the Cycle Totals

Another perturbation that can be made to the cascade is to change the total mass in a given cycle and investigate how that affects the output signal. Previously it was shown (16.5) that for a single cycle, the effect of changing the total on the concentration of cycle species was

given by:

$$C_T^p = \frac{\varepsilon_s^1}{\varepsilon_s^1 M_p + \varepsilon_p^2 M_s}$$

Likewise it was shown that the cycling flux as a result of changes in the total cycle mass was given by:

$$C_T^J = \varepsilon_p^2 C_T^p = \varepsilon_p^2 \frac{\varepsilon_s^1}{\varepsilon_s^1 M_p + \varepsilon_p^2 M_s}$$

Consider the two layer cascade in Figure 16.21. The response of  $p_1$  to changes in  $T_1$  is given by:

$$C_{T_1}^{p_1} = \frac{\varepsilon_{s_1}^1}{\varepsilon_{s_1}^1 M_{p_1} + \varepsilon_{p_1}^2 M_{s_1}}$$

The local response of  $p_1$  on  $p_2$  is given by the term in equation (16.15):

$$r_{p_1}^{p_2} = \frac{M_{s_2} \varepsilon_{p_1}^{2f}}{M_{s_2} \varepsilon_{p_2}^{2r} + M_{p_2} \varepsilon_{p_2}^{2r}}$$

However equation (16.4) indicates how  $T_1$  influences  $p_1$ , therefore  $R_{T_1}^{p_2}$  can be obtained as the product:

$$R_{T_1}^{p_2} = R_{T_1}^{p_1} r_{p_1}^{p_2}$$

$$R_{T_1}^{p_2} = \frac{\varepsilon_{s_1}^{1f}}{(\varepsilon_{s_1}^{1f} M_2 + \varepsilon_{p_1}^{1r} M_1)} \cdot \frac{M_{s_2} \varepsilon_{p_1}^{2f}}{(M_{s_2} \varepsilon_{p_2}^{2r} + M_{p_2} \varepsilon_{p_2}^{2f})}$$

In general for a cascade with  $n$  layers:

$$R_{T_1}^{p_n} = R_{T_1}^{p_1} r_{p_1}^{p_2} r_{p_2}^{p_3} r_{p_3}^{p_4} \dots r_{p_{n-1}}^{p_n}$$

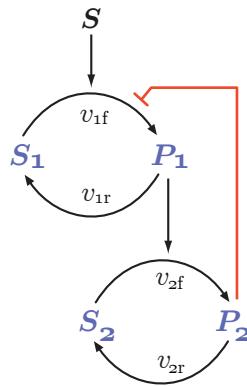
When considering the influence of other cycle totals, for example  $R_{T_2}^{p_n}$ , the relationship is similar but shorter:

$$R_{T_2}^{p_2} = R_{T_2}^{p_2} r_{p_2}^{p_3} r_{p_3}^{p_4} \dots r_{p_{n-1}}^{p_n}$$

As expected, perturbations of the cycle totals that are farther away from the output,  $p_n$ , lead to higher responses because each layer has the potential to amplify.

### Cascades with Negative Feedback

Many signalling pathways include feedback loops which can have marked effects on the behavior of the pathways [66, 71]. Figure 16.23 illustrates a two layer protein cascade with a single negative feedback loop from the output signal  $P_2$  to the reaction step where the input signal,  $S$  enters.



**Figure 16.23** Two layer protein cascade with negative feedback loop from  $P_2$  to  $v_{1f}$ .

The usual approach to deriving the response coefficient,  $R_s^{P_2}$  is quite tedious when dealing with negative feedback. Instead, following Bruggeman [13, 12] we will write out the function equation in the following way:

$$\begin{aligned} p_1 &= p_1(p_2, s, T_1) \\ p_2 &= p_2(p_1, T_2) \end{aligned}$$

In these equations  $s_1$  and  $s_2$  are absent, this is because we invoke the conservation laws,  $s_1 + p_1 = T_1$  and  $s_2 + p_2 = T_2$ . For example,  $s_1$  is a function of  $p_1$  and  $T_1$  which allows us to omit  $s_1$  from the function and replace it with  $p_1$  and  $T_1$ . Likewise for the second equation. If we assume that the total amount of mass in each cycle is constant then we can also omit  $T_1$  and  $T_2$ , yielding:

$$\begin{aligned} p_1 &= p_1(p_2, s) \\ p_2 &= p_2(p_1) \end{aligned}$$

Taking the total derivative of each equation yields:

$$\begin{aligned} dp_1 &= \frac{\partial p_1}{\partial p_2} dp_2 + \frac{\partial p_1}{\partial s} ds \\ dp_2 &= \frac{\partial p_2}{\partial p_1} dp_1 \end{aligned}$$

Dividing both sides by  $p_1$  and  $p_2$  respectively:

$$\begin{aligned} \frac{dp_1}{p_1} &= \frac{\partial p_1}{\partial p_2} \frac{1}{p_1} dp_2 + \frac{\partial p_1}{\partial s} \frac{1}{p_1} ds \\ \frac{dp_2}{p_2} &= \frac{\partial p_2}{\partial p_1} \frac{1}{p_2} dp_1 \end{aligned}$$

Multiplying top and bottom by the appropriate term  $p_1$ ,  $p_2$  and  $s$ :

$$\frac{dp_1}{p_1} = \frac{\partial p_1}{\partial p_2} \frac{p_2}{p_1} \frac{dp_2}{p_2} + \frac{\partial p_1}{\partial s} \frac{s}{p_1} \frac{ds}{s}$$

$$\frac{dp_2}{p_2} = \frac{\partial p_2}{\partial p_1} \frac{p_1}{p_2} \frac{dp_1}{p_1}$$

Finally dividing both sides by  $ds/s$  yields:

$$R_s^{p_1} = \frac{\partial p_1}{\partial p_2} \frac{p_2}{p_1} R_s^{p_2} + \frac{\partial p_1}{\partial s} \frac{s}{p_1}$$

$$R_s^{p_2} = \frac{\partial p_2}{\partial p_1} \frac{p_1}{p_2} R_s^{p_1}$$

The partial derivative terms as the local sensitivities,  $r$  (16.16). The equations can therefore be rewritten as:

$$R_s^{p_1} = r_{p_2}^{p_1} R_s^{p_2} + r_s^{p_1}$$

$$R_s^{p_2} = r_{p_1}^{p_2} R_s^{p_1}$$

Given the two equations we can solve for  $R_s^{p_2}$ :

$$R_s^{p_2} = \frac{r_1^2 r_s^1}{1 - r_1^2 r_2^1} \quad (16.18)$$

The feedback term is given by  $r_2^1$ . Note the similarity to the generic negative feedback response to equation (12.1). If the feedback term is set to zero, meaning the feedback loop is absent, the equation reduces to equation (16.17):

$$R_s^{p_2} = r_1^2 r_s^1$$

The sign for the negative feedback term will be negative, this means that the presence of negative feedback will *reduce* the sensitivity of  $P_2$  to signal,  $S$ . That is negative feedback will lock  $P_2$  into a narrow range. If the gain of the cascade is large, that is  $r_2^1 \gg 0$  then the response equation can be simplified further to:

$$R_s^{p_2} = \frac{r_s^1}{r_1^2}$$

In this mode, the cascade acts as a negative feedback amplifier [112, 131] although with a gain reduced by  $r_1^2$  but partly compensated by  $r_s^1$ . If the two gains remains relatively constant over the operating range then the cascade acts as a **tracking device**, tracking  $p_2$  in response to changes to  $s$ .

## Further Reading

---

1. Reich, J.G. and Selkov, E.E., 1981. Energy metabolism of the cell: a theoretical treatise. Academic Pr.
2. Hofmeyr, J.H.S., Kacser, H. and Merwe, K.J., 1986. Metabolic control analysis of moiety conserved cycles. The FEBS Journal, 155(3), pp.631-640.
3. Bluthgen, N., Bruggeman, F.J., Legewie, S., Herzel, H., Westerhoff, H.V. and Khodenko, B.N., 2006. Effects of sequestration on signal transduction cascades. The FEBS journal, 273(5), pp.895-906.

## Exercises

---

1. Create a model of single protein phosphorylation cycle and investigate using simulation the effects of changing the Kinase and phosphatase  $K_m$ s on the steady state response to changes in Kinase concentrations.
2. Repeat the first exercise but add additional cycles forming a multi-layered phosphorylation cascade. Show by simulation how the response coefficient of the entire cascade increases with the number of layers.
3. Create a two layered cascade and add a negative feedback loop from the last species in the cascade to the first cycle. Investigate by simulation the response of the output of the cascade to changes in the inputs as a function of the strength of the feedback.

## 16.A Python/Tellurium Scripts

---

```
import tellurium as te
import roadrunner
import pylab

r = te.loada'''
    var AB, A, B
    A + B -> AB; k1*A*B - k2*AB;

    A = 0.001; B = 100; AB = 0;
    k1 = 0.4; k2 = 1;
'''

r.conservatedMoietyAnalysis = True;

print "A, B, AB = ", r.A, r.B, r.AB
```

```

print "_CSUM0 = ", r._CSUM0
print "_CSUM1 = ", r._CSUM1

# _CSUM0 = AB + A
# _CSUM1 = AB + B

pylab.figure(figsize=(10,5))
pylab.xticks(fontsize=14)
pylab.yticks(fontsize=14)

abSumX = []; aY = [];
n = 400
for i in range (n):
    r.steadyState()
    abSumX.append (r._CSUM0);
    aY.append (r.A);
    r._CSUM0 = r._CSUM0 + 0.5

pylab.plot (abSumX, aY, linewidth=3)
pylab.xlabel ("$a+ab$", fontsize=16)
pylab.ylabel ("$a$", fontsize=18)
pylab.savefig("sequestrationUltraPlotJustA.pdf")

pylab.show()

```

**Listing 16.1** Ultrasensitivity by sequestration using  $a + b \rightleftharpoons ab$  model

```

import tellurium as te
import roadrunner
import pylab

r = te.loada("""
J1: $Xo -> X; kcat*(k1*Xo - k11*X);
J2: X -> $X1; k2*X/(X + Km);
X + II -> XI; k3*X*II - k4*XI;

Xo = 1; k1 = 0.1; k2 = 1;
k3 = 3; k4 = 0.1; II = 3;
Km = 0.001; k11 = 12; kcat = 1;
""")

r.conervedMoietyAnalysis = True;

r.Xo = 0.01;
xv = []; yv = []; yc = []
n = 400;

```

```
for i in range (n):
    r.steadyState()
    xv.append (r.Xo)
    yc.append (r.getCC ('XI', 'Xo'))
    yv.append (r.XI)
    r.Xo = r.Xo + 0.05

pylab.xlabel ("$X_o$", fontsize=16)
pylab.ylabel ("X or Response Coefficient", fontsize=16)

pylab.plot (xv, yv, color='b', label='X', linewidth=2)
pylab.plot (xv, yc, color='r', label='$R^X_{\{Xo\}}$', linewidth=2)
pylab.legend()

pylab.savefig ('UltrasensitivityMoreComplex.pdf')
```

**Listing 16.2** Ultrasensitivity by sequestration showing sigmoid behavior

# 17

## *Flux Constraints*

### 17.1 Flux Constraints

---

Figure 17.1 shows a simplified metabolic map from *Corynebacterium glutamicum* [98]. The numbers next to the reaction steps indicate the flux through each step and shows how the flow of mass through the different metabolic pathway are distributed. Stoichiometry has a significant effect on the possible space of flux distributions and in chapter and the next the focus will be on stoichiometry and flux balance.

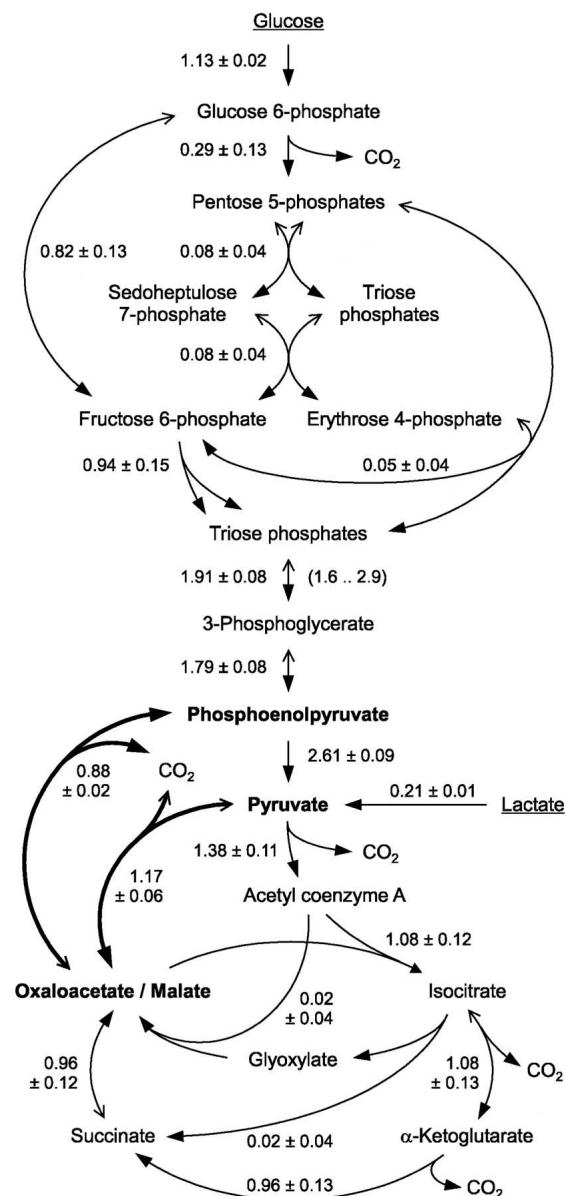
### 17.2 Flux Balance Laws

---

While the rows of the stoichiometry matrix,  $\mathbf{N}$ , indicate possible conservation relationships among the molecular species, the columns provide information on the constraints among the reaction rates at steady state. As we've already seen, the steady state of a system is defined when the rates of change are zero, that is when:

$$\mathbf{N}\mathbf{v}(\mathbf{s}, \mathbf{p}) = \mathbf{0}$$

In future we will shorten the notation to  $\mathbf{N}\mathbf{v} = \mathbf{0}$ . Note that although the rates of change may be zero at steady state, the net flows (or fluxes) through individual reactions will not be zero. By illustration, let us look at the very simple branched pathway shown in Figure 17.2. The stoichiometry matrix for this pathway is:  $\mathbf{N} = [1 \ -1 \ -1]$  and the balance equation at

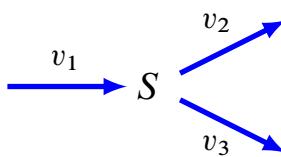


**Figure 17.1** Metabolic Map of *Corynebacterium glutamicum* central metabolism adapted from [98].

steady state is given by:

$$\begin{bmatrix} 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = 0$$

Or, more simply,  $v_1 - v_2 - v_3 = 0$ .



**Figure 17.2** Simple branched pathway.

A common need by metabolic engineers is to know the flux distribution throughout a reaction network. One approach to obtain this information is to measure every individual flux in the network. This can be done, at least in principle, by measuring the consumption or turnover rates of all the metabolites in the network. The easiest rates to measure are on the reaction steps that connect directly to the external environment, such steps might be involved in nutrient and oxygen consumption, carbon dioxide, ethanol or biomass production, quantities that can be measured experimentally. However, the internal fluxes that are deep inside the metabolic networks are much more difficult to measure, although the use of  $^{13}\text{C}$ -labeled substrates has made such measurements more accessible.

In practice it is extremely difficult to measure every reaction rate directly, instead the balance equations can be exploited to reduce the number of necessary flux measurements. To illustrate, the balance equation for the simple branched pathway (Figure ??) shows us that only two rates actually need be measured because the third can be computed. For example, if  $v_2$  and  $v_1$  were measured, the third rate,  $v_3$ , could be calculated from the balance equation  $v_3 = v_2 - v_1$ , taking note that the pathway must be in steady state. For an experimentalist this is a great benefit because it reduces the number of measurements that need to be made.

One of the practical aims of flux balance analysis is to devise methods that allow all the fluxes in a pathway to be determined with the minimum effort. To devise such methods however, a number of questions need to be answered. For example, what are the minimum number of fluxes that should be measured experimentally to fully determine all fluxes in a pathway? In the simple branch pathway (Figure 17.2) a minimum of two fluxes were required. Alternatively it may not be possible to measure even the minimum number, in such cases can a best estimate for the flux distribution in a pathway be computed? The following sections will consider some of the approaches to answering these questions, particularly for arbitrary networks where systematic approaches are required.

### Box 17.1 The Null Space

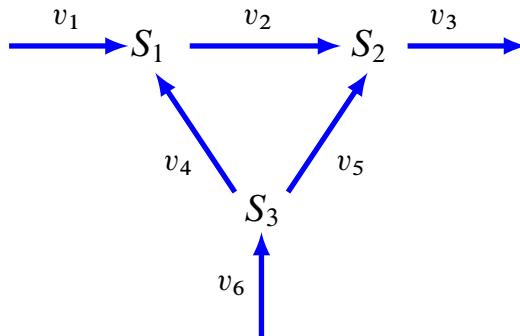
Given a matrix equation of the form  $A\mathbf{x} = \mathbf{0}$  where  $A$  is an  $m \times n$  matrix and  $\mathbf{x}$  is a column vector of  $n$  elements, the solution, that is the set of vectors  $\mathbf{x}$  that satisfy this equation, is called the **null space** of  $A$ .

The minimum number of vectors required to fully describe the null space is called the **dimension** and is equal to the rank of the matrix  $\text{rank}(A)$  minus the number of columns,  $n$ . These vectors form what is called a **basis** for the space and linear combinations of these vectors can generate any other vector in the null space. In order to form a basis, the vectors must also be linearly independent.

Many tools can compute the basis for the null space, for example `null (A, 'r')` will compute the basis in Matlab, while `NullSpace [A]` can be used to compute the basis in Mathematica.

## 17.3 Determined Systems

Consider the more complicated pathway shown in Figure 17.3.



**Figure 17.3** Complex branched pathway.

The stoichiometry matrix for this pathway is:

$$\mathbf{N} = \begin{matrix} & v_1 & v_2 & v_3 & v_4 & v_5 & v_6 \\ S_1 & 1 & -1 & 0 & 1 & 0 & 0 \\ S_2 & 0 & 1 & -1 & 0 & 1 & 0 \\ S_3 & 0 & 0 & 0 & -1 & -1 & 1 \end{matrix} \quad (17.1)$$

which corresponds to the following three balance equations:

$$\begin{aligned} v_1 - v_2 + v_4 &= 0 \\ v_2 - v_3 + v_5 &= 0 \\ v_6 - v_4 - v_5 &= 0 \end{aligned}$$

The first question to ask is what is the minimum number of fluxes that need to be measured so that the remainder of the fluxes can be estimated? Since there are three equations and six unknowns, at least three of the fluxes must be measured so that number of unknowns can be reduced to three. However, of the six, which of the three fluxes should be measured? For example, measuring  $v_1$ ,  $v_2$  and  $v_4$ , will not help because it is not possible to compute the others from these measurements.

In order to answer this question let us divide the fluxes into two groups, call one the measured fluxes ( $\mathbf{J}_M$ ) and the other the computed fluxes ( $\mathbf{J}_C$ ). The computed fluxes will be calculated from some combination of the measured fluxes. Consider the system equation at steady state:

$$\mathbf{N} \mathbf{v} = \mathbf{0}$$

Apply row reduction to the system equation until  $\mathbf{N}$  is in reduced echelon form (See Box 15.3). Since the right-hand side is zero, it remains unchanged in the process. These operations lead to:

$$\begin{bmatrix} \mathbf{I} & \mathbf{M} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{v} = \mathbf{0} \quad (17.2)$$

The process is likely to result in column as well as row exchanges and as a result the linearly independent columns will move to the left partition forming the identity matrix and the linearly dependent columns will be found in the partition corresponding to  $\mathbf{M}$ . Let us partition the  $\mathbf{v}$  vector to correspond to the partitioning in the echelon matrix, so that:

$$\begin{bmatrix} \mathbf{I} & \mathbf{M} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{bmatrix} = \mathbf{0}$$

which when multiplied out gives  $\mathbf{v}_1 = -\mathbf{M} \mathbf{v}_2$ . This implies that the flux terms in the  $\mathbf{v}_1$  partition correspond to the computed fluxes,  $\mathbf{J}_C$ , and  $\mathbf{v}_2$  to the measured fluxes,  $\mathbf{J}_M$ , that is  $\mathbf{J}_C = -\mathbf{M} \mathbf{J}_M$ .

This relation describes a set of computed fluxes,  $\mathbf{J}_C$ , as a function of a set of measured fluxes,  $\mathbf{J}_M$  via a transformation matrix,  $\mathbf{M}$ . To follow conventional notation, the term  $-\mathbf{M}$  will be renamed to  $\mathbf{K}_0$  (that is  $\mathbf{M} = -\mathbf{K}_0$ ) so that

$$\mathbf{J}_C = \mathbf{K}_0 \mathbf{J}_M. \quad (17.3)$$

and equation (17.2) can be reexpressed as:

$$\begin{bmatrix} \mathbf{I} & -\mathbf{K}_0 \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{J}_C \\ \mathbf{J}_M \end{bmatrix} = \mathbf{0} \quad (17.4)$$

Returning to the example shown in Figure 17.3, let us apply a series of elementary operations to the stoichiometry matrix to reduce the stoichiometry to its reduced echelon form (17.4):

**1. Start with the stoichiometry matrix.**

$$\begin{bmatrix} 1 & -1 & 0 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & -1 & 1 \end{bmatrix}$$

**1. Multiply the 3rd row by -1.**

$$\begin{bmatrix} 1 & -1 & 0 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & -1 \end{bmatrix}$$

**2. Add the 2nd row to the 1st row.**

$$\begin{bmatrix} 1 & 0 & -1 & 1 & 1 & 0 \\ 0 & 1 & -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & -1 \end{bmatrix}$$

**3. Add the 3rd row times -1 to the 1st row.**

$$\begin{bmatrix} 1 & 0 & -1 & 0 & 0 & 1 \\ 0 & 1 & -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & -1 \end{bmatrix}$$

**4. And finally, exchange the 3rd and 4th columns.**

$$\begin{bmatrix} 1 & 0 & 0 & -1 & 0 & 1 \\ 0 & 1 & 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & -1 \end{bmatrix}$$

These operations lead to the following reduced echelon matrix (leading ones are shown in a darker gray):

$$\text{Reduced Echelon} = \left[ \begin{array}{cccccc} v_1 & v_2 & v_4 & v_3 & v_5 & v_6 \\ \textcolor{red}{1} & 0 & 0 & -1 & 0 & 1 \\ 0 & \textcolor{red}{1} & 0 & -1 & 1 & 0 \\ 0 & 0 & \textcolor{red}{1} & 0 & 1 & -1 \end{array} \right] \quad (17.5)$$

During the reduction, the third and forth columns were exchanged. The partition that holds the identity matrix marks the computed fluxes and the right-hand partition which holds the  $K_0$  matrix marks the measured fluxes. Thus the **computed fluxes** correspond to the

**independent columns** and the **measured fluxes** to the **dependent columns**. If we extract the  $K_0$  partition, equation (17.3) can be used to relate the computed to the measured fluxes as follows:

$$\begin{bmatrix} v_1 \\ v_2 \\ v_4 \end{bmatrix} = \odot \begin{bmatrix} -1 & 0 & 1 \\ -1 & 1 & 0 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} v_3 \\ v_5 \\ v_6 \end{bmatrix} \quad (17.6)$$

Or

$$v_1 = v_3 - v_6$$

$$v_2 = v_3 - v_5$$

$$v_4 = v_6 - v_5$$

This shows that in principle only  $v_3$ ,  $v_5$  and  $v_6$  need be measured from which all remaining rates can be calculated. A visual inspection of the pathway in Figure 17.3, will reveal this to be true, thus,  $v_4$  can be computed from  $v_5$  and  $v_6$ ;  $v_2$  can be computed from  $v_5$  and  $v_3$ ; and lastly,  $v_1$  can be computed from  $v_2$  and  $v_4$ .

Software tools such as PySCeS [92] or Tellurium can be used to automatically compute the  $K_0$  matrix along with an appropriately reordered stoichiometry matrix. In summary, the method outlined above enables us to derive the minimum set of fluxes to measure in order to determine all fluxes in an arbitrary pathway.

### Linear Algebra of Determined Systems

Let us assume we can reorder the columns of the stoichiometry matrix so that all the dependent columns are moved to the left-side of the matrix and the independent columns are moved to the right-side of the matrix. Note this is the opposite order to the columns in equations (17.5) and (17.2). Furthermore, let us also assume that the rows have also been reordered so that the independent rows are moved to the top and the dependent rows to the bottom of the matrix. These prerequisites means that the stoichiometry matrix has a partitioned structure shown in Figure 17.4.

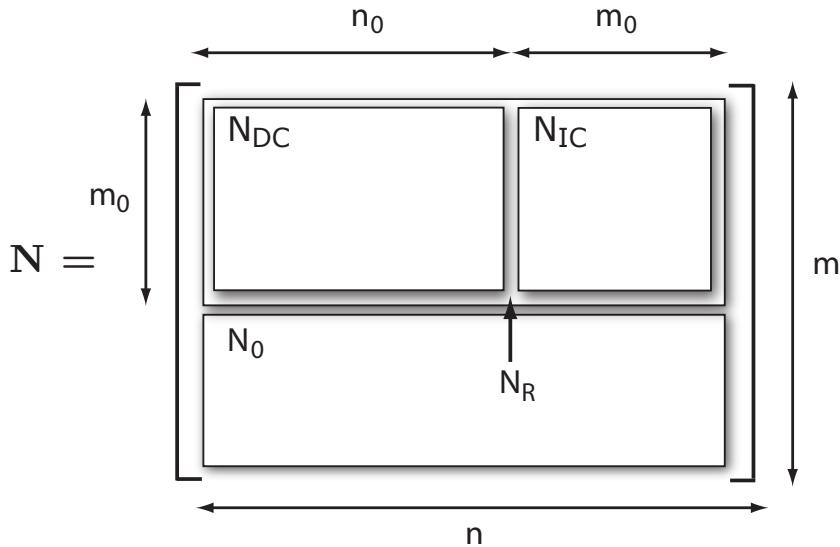
If we take  $N_r$  and partition it into the dependent and independent column we can state::

$$\begin{bmatrix} N_R \end{bmatrix} \begin{bmatrix} J_M \\ J_C \end{bmatrix} = 0$$

$N_R$  can be further partitioned as shown in Figure 17.4:

$$\begin{bmatrix} N_{DC} & N_{IC} \end{bmatrix} \begin{bmatrix} J_M \\ J_C \end{bmatrix} = 0$$

where  $N_{DC}$  represents the set of linearly dependent columns and  $N_{IC}$  the set of linearly independent columns. To reemphasize again, the order of the computed and measured fluxes are swapped compared to that shown in equation (17.4).



**Figure 17.4** Partitioned Stoichiometry Matrix:  $n$  = number of reactions;  $m$  = number of species;  $N_{DC}$  = partition of linearly dependent columns;  $N_{IC}$  = partition of linearly independent columns;  $N_R$  = reduced stoichiometry matrix;  $N_0$  partition of linearly dependent rows.

Expanding this equation gives  $N_{DC} J_M + N_{IC} J_C = 0$ . This equation can be rearranged and both sides multiplied by the inverse of  $N_{IC}$  to obtain:

$$J_C = -(N_{IC})^{-1} N_{DC} J_M \quad (17.7)$$

This result gives us a relationship between the computed and measured fluxes. The term  $-(N_{IC})^{-1} N_{DC}$  can be replaced by,  $K_0$ , so that  $J_C = K_0 J_M$ . This equation is identical to equation (17.3) but offers an alternative approach to computing  $K_0$  and is the method often cited in the literature [128, 30]. The inverse of  $N_{IC}$  is guaranteed to exist because the matrix is square and all rows and columns are guaranteed by construction to be linearly independent.

The equation,  $K_0 = -(N_{IC})^{-1} N_{DC}$  can be rearranged into the following form:

$$\begin{bmatrix} N_{DC} & N_{IC} \end{bmatrix} \begin{bmatrix} I \\ K_0 \end{bmatrix} = \mathbf{0} \quad (17.8)$$

or more simply:

$$N_R K = \mathbf{0} \quad (17.9)$$

This shows that the  $K_0$  matrix is related to the null space (Box 17.1) of the reordered stoichiometry matrix.

### Examples

The following examples illustrate the application of equation (17.7).

- a) Consider the branched pathway shown in Figure 17.3. The columns of the stoichiometry matrix can be reordered so that the linearly dependent columns ( $N_{DC}$ ) are first, followed by the linearly independent columns ( $N_{IC}$ ). Row reduction to the reduced echelon form (17.4) can be used to determine which are the linearly independent and dependent columns, equation (17.5). In the stoichiometry matrix below, the partitions have been exchanged so that the linearly independent columns are first, followed by the linearly dependent columns:

$$\mathbf{N} = \begin{bmatrix} v_3 & v_5 & v_6 & v_1 & v_2 & v_4 \\ 0 & 0 & 0 & 1 & -1 & 1 \\ -1 & 1 & 0 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 & 0 & -1 \end{bmatrix}$$

From the reordered matrix, the  $N_{DC}$  and  $N_{IC}$  partitions can be extracted from which the dependency relations can be derived by applying equation (17.7).

$$\mathbf{K}_0 = - \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 0 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 \\ 1 & -1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \quad (17.10)$$

The derived  $\mathbf{K}_0$  corresponds to the same result found in equation (17.6).

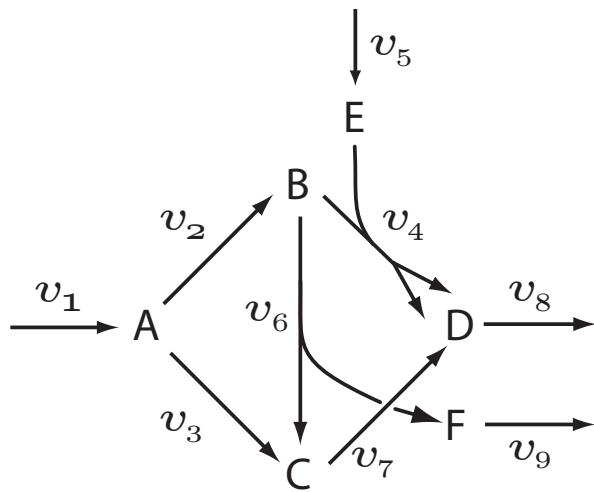
- b) A more complex example of a pathway is shown in Figure 17.5. The stoichiometry matrix for this network is given by:

$$\mathbf{N} = \begin{array}{c} \begin{array}{ccccccccc} v_1 & v_2 & v_3 & v_4 & v_5 & v_6 & v_7 & v_8 & v_9 \\ \hline A & 1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 \\ B & 0 & 1 & 0 & -1 & 0 & -1 & 0 & 0 \\ C & 0 & 0 & 1 & 0 & 0 & 1 & -1 & 0 \\ D & 0 & 0 & 0 & 2 & 0 & 0 & 1 & -1 \\ E & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 \\ F & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{array} \end{array}$$

and the balance equations by:

$$\begin{array}{ll} v_1 - v_2 - v_3 = 0 & v_2 - v_4 - v_6 = 0 \\ v_3 + v_6 - v_7 = 0 & 2v_4 + v_7 - v_8 = 0 \\ v_5 - v_4 = 0 & v_6 - v_9 = 0 \end{array}$$

Reorder the columns of the stoichiometry matrix so that the linearly dependent columns are on the left and linearly independent columns are on the right (Figure 17.4). Note that all



**Figure 17.5** Complex Network incorporating two input fluxes and two output fluxes, coupled internally by multiple branches and one reaction that exhibits non-unity stoichiometry ( $v_4$ ).

the rows are linearly independent so that there is no  $N_0$  partition in the reordered matrix. Reordering can be accomplished by carrying out a row reduction on the matrix to reduced echelon form (17.2) and recording the column changes in the stoichiometry matrix. Note that the partitions must be exchanged to match the structure shown in equation (18.1). The simplest reordering is given by the following stoichiometry matrix:

$$\mathbf{N} = \begin{array}{c|ccccc|cccc} & v_7 & v_8 & v_9 & v_1 & v_2 & v_3 & v_4 & v_5 & v_6 \\ \hline A & 0 & 0 & 0 & 1 & -1 & -1 & 0 & 0 & 0 \\ B & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 & -1 \\ C & -1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ D & 1 & -1 & 0 & 0 & 0 & 0 & 2 & 0 & 0 \\ E & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 \\ F & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}$$

The  $\mathbf{K}_0$  matrix can be computed from the null space (17.9) of this reordered matrix:

$$\mathbf{K} = \begin{array}{c|ccc} v_7 & 1 & 0 & 0 \\ v_8 & 0 & 1 & 0 \\ v_9 & 0 & 0 & 1 \\ \hline v_1 & 0.5 & 0.5 & 0 \\ v_2 & -0.5 & 0.5 & 1 \\ v_3 & 1 & 0 & -1 \\ v_4 & -0.5 & 0.5 & 0 \\ v_5 & -0.5 & 0.5 & 0 \\ v_6 & 0 & 0 & 1 \end{array} \quad \mathbf{K}_0 = \begin{array}{c|ccc} v_1 & 0.5 & 0.5 & 0 \\ v_2 & -0.5 & 0.5 & 1 \\ v_3 & 1 & 0 & -1 \\ v_4 & -0.5 & 0.5 & 0 \\ v_5 & -0.5 & 0.5 & 0 \\ v_6 & 0 & 0 & 1 \end{array}$$

From the  $\mathbf{K}_0$  matrix the relation between the measured and computed fluxes can be determined. From the reordering of the stoichiometry matrix it should be apparent that the measured fluxes are  $v_7$ ,  $v_8$ , and  $v_9$ , that is a minimum of three fluxes must be measured in order to fully determine the remainder. Of the three measured fluxes,  $v_7$  is the most problematic because it is an internal flux which experimentally would not be easy to determine. It is however possible to derive other combinations of measured and computed fluxes. Most notable is the following list of independent fluxes,  $v_5$ ,  $v_8$  and  $v_9$ . All three are edge fluxes which in principle are easier to measure.

The stoichiometry matrix can be reordered as follows:

$$\mathbf{N} = \begin{array}{c|ccccccccc} & v_5 & v_8 & v_9 & & v_6 & v_2 & v_3 & v_4 & v_1 & v_7 \\ A & 0 & 0 & 0 & | & 0 & -1 & -1 & 0 & 1 & 0 \\ B & 0 & 0 & 0 & | & -1 & 1 & 0 & -1 & 0 & 0 \\ C & 0 & 0 & 0 & | & 1 & 0 & 1 & 0 & 0 & -1 \\ D & 0 & -1 & 0 & | & 0 & 0 & 0 & 2 & 0 & 1 \\ E & 1 & 0 & 0 & | & 0 & 0 & 0 & -1 & 0 & 0 \\ F & 0 & 0 & -1 & | & 1 & 0 & 0 & 0 & 0 & 0 \end{array}$$

which yields the following  $\mathbf{K}_0$  matrix from the null space:

$$\mathbf{K}_0 = \begin{array}{c|ccc} v_6 & 0 & 0 & 1 \\ v_2 & 1 & 0 & 1 \\ v_3 & -2 & 1 & -1 \\ v_4 & 1 & 0 & 0 \\ v_1 & -1 & 1 & 0 \\ v_7 & -2 & 1 & 0 \end{array}$$

In turn this gives the dependency equations using equation (17.3):

$$\begin{aligned}v_6 &= v_9 \\v_2 &= v_5 + v_9 \\v_3 &= v_8 - v_9 - 2 v_5 \\v_4 &= v_5 \\v_1 &= v_8 - v_5 \\v_7 &= v_8 - 2 v_5\end{aligned}$$

In summary, measuring only  $v_5$ ,  $v_8$  and  $v_9$  allows us to completely determine all the fluxes in the network. Unfortunately in real systems the internal structure of the network will be much more complex and will include many more degrees of freedom. This means that in many cases there will be insufficient information to fully determine the internal fluxes. Such cases are called **underdetermined systems** and alternative strategies must be used to gain access to the unknown fluxes. Two common strategies to the study of underdetermined systems include **flux balance analysis** and **metabolic flux analysis**. Flux balance analysis relies on linear programming while metabolic flux analysis uses  $^{13}\text{C}$ -labeled substrates to estimate fluxes. Both these topics are however beyond the scope of this book but useful reading material is provided below.

---

## Further Reading

---

1. Bonarius HPJ, Schmidb G and Tramper J (1997) Flux analysis of underdetermined metabolic networks: the quest for the missing constraints. *Trends in Biotechnology*, 15(8), 308–314
2. Klamt S, Schuster S and Gilles ED (2002) Calculability analysis in underdetermined metabolic networks illustrated by a model of the central metabolism in purple non-sulfur bacteria. *Biotechnology and Bioengineering*, 77(7), 734–751
3. Klamt S, Schuster S. (2002) Calculating as many fluxes as possible in underdetermined metabolic networks. *Mol Biol Rep.* 29(1-2), 243-8.
4. Palsson BO (2006) Systems Biology Systems Biology: Properties of Reconstructed Networks. Cambridge University Press, ISBN: 978-0521859035
5. Wiechert W. (2001)  $^{13}\text{C}$  metabolic flux analysis. *Metabolic Engineering* Jul;3(3):195-206.
6. Stephanopoulos, Gregory (1998). "Chapter 9: Methods for the Experimental Determination of Metabolic Fluxes by Isotope Labeling". *Metabolic engineering: principles and methodologies*. San Diego: Academic Press. pp. 356–404. ISBN 0-12-666260-6.

# 18

## *Steady State Flux Patterns*

The previous chapter discussed some aspects of flux measurements and constraints in metabolic pathways. In this chapter we will take a closer look at some of the more theoretical tools available and the insights they provide. The first to consider is the Null space of the stoichiometry matrix.

### 18.1 The Null Space

---

The previous chapter considered the null space (Box 17.1) of the stoichiometry matrix. In this section the question of how to physical interpret the null space will be considered. Equations (18.1) and (18.1) indicate that the null space gives information on the relationships between fluxes in a network at steady state. However there is additional information that can be gleaned from the null space. First, recall that the null space of the stoichiometry matrix is given by the relation:

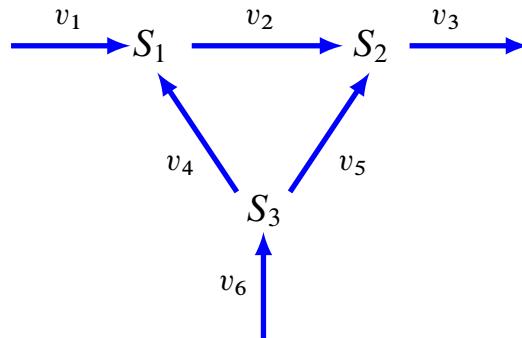
$$\begin{bmatrix} \mathbf{N}_{DC} & \mathbf{N}_{IC} \end{bmatrix} \begin{bmatrix} \mathbf{I} \\ \mathbf{K}_0 \end{bmatrix} = 0 \quad (18.1)$$

or more simply:

$$\mathbf{N}_R \mathbf{K} = 0 \quad (18.2)$$

For the pathway shown in Figure 18.1 the  $\mathbf{K}$  matrix was derived in the previous chapter (17.10).

The simplest interpretation of the  $\mathbf{K}$  matrix is that the vectors that make up  $\mathbf{K}$  represent possible steady state flow patterns in the network. In addition, any linear combination of the

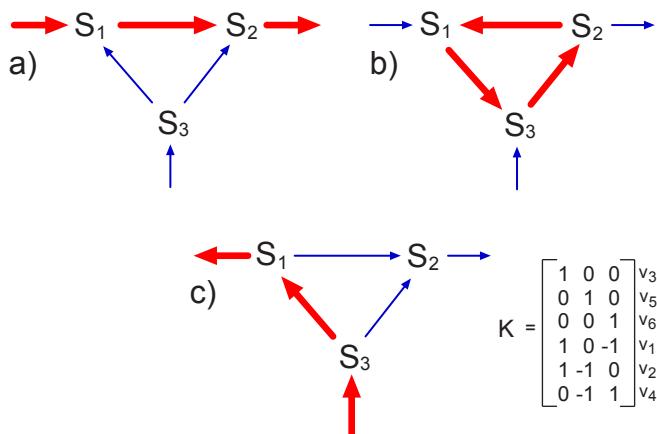


**Figure 18.1** Branched pathway.

vectors is also a valid steady state flow pattern. Thus for the network shown in Figure 18.1, the null space can be shown to be:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & -1 \\ 1 & -1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$$

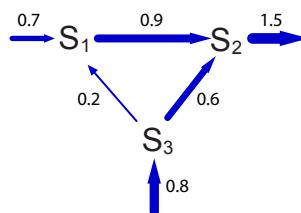
The null space contains three vectors which can be interpreted as flow patterns which satisfy the steady state condition. These flow patterns are shown in Figure 18.2. Any actual flow



**Figure 18.2** Flow Patterns Based on the Null Space of the Stoichiometry Matrix.

pattern is some linear combination of the three basic patterns indicated in the null space. For example, the following combination is a potential flow pattern:

$$\mathbf{J} = \begin{bmatrix} 1.5 \\ 0.6 \\ 0.8 \\ 0.7 \\ 0.9 \\ 0.2 \end{bmatrix} = 1.5 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} + 0.6 \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ -1 \\ -1 \end{bmatrix} + 0.8 \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \\ 0 \\ 1 \end{bmatrix}$$



**Figure 18.3** Potential Flow Patterns Based on the Null Space.

The one problem with this interpretation is that negative terms in the  $\mathbf{K}$  matrix indicate that the flow is in the opposite direction to that indicated in the network diagram. Such flows might be thermodynamically unlikely. For example, pattern c) shows the reaction  $v_1$  operating in the opposite direction to that indicated in the original Figure 17.3. It may be the case that  $v_1$  is reversible in which case pattern c) is a legitimate flow pattern. However  $v_1$  may be irreversible, in which case the flow pattern is not likely to occur *in vivo*. The use of elementary modes (see next section) eliminates this problem by forbidding patterns that include irreversible reactions.

In summary, the null space vectors, and combinations thereof, can be interpreted as possible steady state flows through a given network. There is however another interpretation which is also useful. Let us consider the system equation again:

$$\mathbf{N} \mathbf{v} = 0$$

Let us assume that it is possible, by some means, to change the rates through the reactions such that the species levels remain unchanged but the flux changes. In particular it can be shown [53] that the unscaled concentration control coefficients and the null space are related by the expression:

$$\mathbf{C}^s \mathbf{K} = \mathbf{0} \quad (18.3)$$

where elements of the  $\mathbf{C}^s$  matrix equal  $ds_i/dv_j$ , that is, how a given enzyme,  $E_j$  effects the steady state concentration of a given molecular species,  $S_i$ . The equation tell us that

perturbations in reaction rates that match entries in the  $\mathbf{K}$  vector results in no changes in concentrations. The same applies to linear combinations of vectors in the  $\mathbf{K}$  matrix. With this in mind we can state that there is a set of perturbations,  $\delta\mathbf{v}$ , that satisfies the following:

$$\mathbf{N}(\mathbf{v} + \delta\mathbf{v}) = 0$$

where the  $\delta\mathbf{v}$  term signifies a vector of changes that are made to the rates to satisfy the steady state condition. Because  $\mathbf{N}(\mathbf{v} + \delta\mathbf{v}) = 0$ , this equation can be simplified to:

$$\mathbf{N}\delta\mathbf{v} = 0$$

This equation tells that the null space of  $\mathbf{N}$  can be interpreted as the vector  $\delta\mathbf{v}$ . That is,  $\mathbf{K}$ , can be interpreted as a set of rate changes which leaves the steady state species levels unchanged. Such perturbations could be achieved by changing the level of gene expression at each reaction which has a non-zero entry in the  $\mathbf{K}$  matrix. For example the first column of the  $\mathbf{K}$  matrix is  $[100110]^T$ . This means that changing rates  $v_3$ ,  $v_1$  and  $v_2$  by a  $\delta v$  amount will leave the steady state concentration of  $S_1$ ,  $S_2$  and  $S_3$  unchanged but will increase the net flow from  $v_1$  to  $v_3$  by a  $\delta v$  amount. In practice such changes might not be realizable but in principle one could imagine changing the enzyme activities at  $v_1$ ,  $v_2$  and  $v_3$  through changes in gene expression. Since enzyme activity is proportional to the concentration of enzyme it must be true that proportional changes in an enzyme concentration,  $E_i$  will lead to proportional changes in the reaction rate  $v_i$ . To make a given relative change in  $v_i$ ,  $\delta v_i/v_i$ , we need only make the same proportional change in  $E_i$ , that is:

$$\frac{\delta e_i}{e_i} = \frac{\delta v_i}{v_i} \quad (18.4)$$

From the first column of  $\mathbf{K}$ , we can state that  $\delta v_1 = \delta v_2 = \delta v_3$  or equivalently:

$$\frac{\delta v_1}{v_1} = \frac{\delta v_2}{v_2} \frac{v_2}{v_1} = \frac{\delta v_3}{v_3} \frac{v_3}{v_1}$$

or

$$\frac{\delta e_1}{e_1} = \frac{\delta e_2}{e_2} \frac{v_2}{v_1} = \frac{\delta e_3}{e_3} \frac{v_3}{v_1}$$

In practice, if we change the activity of enzyme,  $E_1$  by a percentage,  $\alpha$ , then the percentage changes we must make in  $E_2$  and  $E_3$  will equal:

$$\begin{aligned} \frac{\delta e_1}{e_1} &= \alpha \\ \frac{\delta e_2}{e_2} &= \alpha \frac{v_2}{v_1} \\ \frac{\delta e_3}{e_3} &= \alpha \frac{v_3}{v_1} \end{aligned} \quad (18.5)$$

This result shows the changes in enzyme activity that are necessary in order to increase the flux through  $v_1$ ,  $v_2$  and  $v_3$  while keeping all other fluxes and metabolite levels the same. It shows that the relative changes in enzyme concentrations is related to the proportion of flux that the particular step carries. Note that this result applies to large as well as small changes in enzyme concentrations. The ability to alter fluxes independently of metabolite concentrations is a desirable goal in metabolic engineering.

## 18.2 Elementary Modes

A closely related concept to the null space of the stoichiometry matrix is the space of elementary modes. As previously discussed, the vectors in the null space of the stoichiometry matrix can be interpreted as flow patterns in a network. However, one criticism of this approach is that the vectors in the null space can admit patterns that are thermodynamically unlikely (See Figure 18.2). In addition, the set of null space vectors is not unique. Elementary modes are an extension to the null space that avoids these issues.

Elementary models may be considered minimal realizable flow patterns through a network that can sustain a steady state. This means that elementary modes **cannot** be decomposed further into simpler pathways. Like the null space vectors, all possible flows through a network can be constructed from linear combinations of the elementary modes. Whereas the null space vectors admit flow patterns which violate thermodynamic considerations, pattern b) and c) in Figure 18.2, elementary modes do not. This also means that there are likely to be more elementary mode vectors than null space vectors. Figure 18.7 illustrates four elementary modes when the first reaction is considered reversible. Figure 18.8 on the other hand shows only three elementary modes when the first reaction is assumed to be irreversible.

Unlike the null space, the set of elementary modes for a given network is unique (up to an arbitrary scaling factor). Given the fundamental nature of elementary modes in relation to uniqueness and non-decomposability, the term ‘pathway’ can be defined as an elementary mode. Note that the set of elementary modes will change as the set of expressed enzymes change during transitions from one cell state to another. Mathematically the set of elementary modes is defined as the set of flux vectors that satisfy the steady state condition,  $\mathbf{N}v = 0$ , and that the rate of each irreversible reaction must be non-negative,  $v_{irr} \geq 0$ .

Computing elementary models efficiently is a non-trivial calculation but a small number of tools are available. In particular METATOOL (4.3 series) developed by a number of authors including, Thomas Pfeiffer, Stefan Schuster, Juan Carlos Nuno and Ferdinand Moldenhauer is recommended.

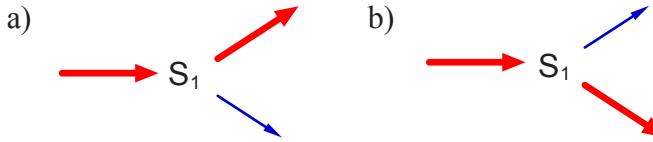
Elementary flux modes provide a comprehensive description of all metabolic routes for a group of enzymes that are stoichiometrically and thermodynamically feasible [120]. As a result, metabolic pathways can be defined in terms of their elementary modes.

1. The vector must satisfy:  $\mathbf{N}\mathbf{e}_i = \mathbf{0}$ , that is the steady state condition.
2. For all irreversible reactions:  $v_i \geq 0$ . This means that all flow patterns must use reactions that proceed in their most natural direction. This makes the pathway described by the elementary mode a thermodynamically feasible pathway.
3. The vector  $\mathbf{e}_i$  must be elementary, that is, it should not be possible to generate  $\mathbf{e}_i$  by combining two other vectors that satisfy the first and second requirements using the same set of enzymes that appear as non-zero entries in  $\mathbf{e}_i$ . In other words it should not be possible to decompose  $\mathbf{e}_i$  into two other pathways that can themselves sustain a steady state.

**Table 18.1** Conditions necessary to define an elementary mode where  $\mathbf{e}_i$  is an elementary mode.

Mathematically elementary modes are defined as follows. An elementary mode,  $\mathbf{e}_i$ , is defined as a vector of fluxes,  $v_1, v_2, \dots$ , such that the three conditions listed in Table 18.1 are satisfied.

In the following examples, all elementary models were computed using Metatool. Figure 18.4 shows the two elementary modes for a simple branched pathway.



**Figure 18.4** Elementary mode patterns in a simple branched pathway assuming irreversibility at all reaction steps. Highlighted reactions in bold signify steps that belong to the elementary mode.

All steps in Figure 18.4 are assumed to be irreversible. Let us show that each mode in this system satisfies the three conditions (Table 18.1). The first condition is steady state, that is for each mode  $\mathbf{e}_i$ ,  $\mathbf{N}\mathbf{e}_i = \mathbf{0}$

The two modes are given by:

$$\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \quad (18.6)$$

By substituting each of these vectors into  $\mathbf{N}\mathbf{e}_i = \mathbf{0}$ , it is easy to show that condition one is satisfied. For condition two we must ensure that all reactions that are irreversible have

positive entries in the corresponding elements of the elementary modes. Since all three reactions in the branch are irreversible and all entries in the elementary modes are positive then condition two is satisfied.

Finally to satisfy condition three we must ask whether we can decompose the two elementary modes into other paths that can sustain a steady state while using the same non-zero entries in the elementary mode. In this example it is impossible to decompose the elementary modes any further without disrupting the ability to sustain a steady state. Therefore with all three conditions satisfied we can conclude that the two vectors given previously are elementary modes.

Like the basis for the null space, all possible flows through a network can be constructed from linear combinations of the elementary modes, that is:

$$\mathbf{v} = \sum \lambda_i \mathbf{e}_i \quad (18.7)$$

where

$$\lambda \geq 0$$

such that the entire space of flows through a network can be described.  $\lambda_i$  must be greater than or equal to zero to ensure that irreversible steps aren't inadvertently made to go in the reverse direction. For example, the following is a possible flow in the branched pathway:

$$v = 2.5 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + 0.5 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 3.0 \\ 2.5 \\ 0.5 \end{bmatrix}$$

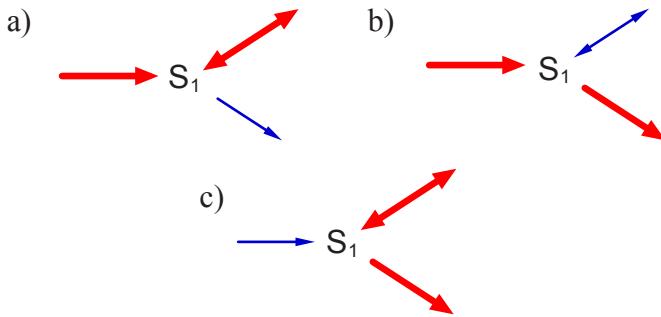
If one of the outflow steps in the simple branched pathway is made reversible an additional elementary mode becomes available that represents the flow between the two outflow branches (Figure 18.5). An additional mode emerges because with only the first two modes it is impossible to represent a flow between the two branches because the scaling factor,  $\lambda_i$ , cannot be negative which would be required to reverse the flow.

Equation (18.3) indicates that if specific perturbations are made along the route indicated by a vector in the null space, then all species remain unchanged while the net flux increases. This equation can be extended to also include elementary modes, so that if  $\mathbf{E}$  is the matrix of elementary modes in column order, then since an elementary mode can be generated from a suitable combination of null space vectors (personal communication: Stefan Schuster), it must be true that:

$$\mathbf{C}^s \mathbf{E} = \mathbf{0} \quad (18.8)$$

This is an important results because it indicates that pathways represented by individual

elementary modes can also be perturbed such that species levels remain unchanged which should have a significant bearing on metabolic engineering strategies.



**Figure 18.5** Elementary mode patterns in a simple branched pathway assuming reversibility at the upper outflow branch.

### Cyclic Branched Model

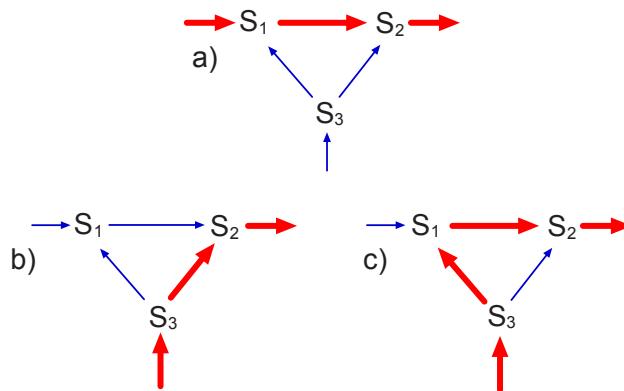
Figure 18.7 lists the elementary modes for a cyclic branched model. Whereas the null space vectors admit flow patterns which violate thermodynamic considerations, elementary modes do not. For example, pattern (b) and (c) in Figure 18.2, the reactions  $v_1$  and  $v_4$  are going in the reverse direction. This also means that there are likely to be more elementary mode vectors than the dimension of the null space. Figure 18.7 illustrates four elementary modes when the first reaction is considered reversible. Figure 18.8 on the other hand shows only three elementary modes when the first reaction is assumed to be irreversible.

### Comment on Condition Three

Condition three in Table 18.1 requires further explanation. Condition three relates to the non-decomposability of an elementary mode and is partly what makes elementary modes interesting, the two other important features include uniqueness and thermodynamic plausibility. Decomposition implies that it is possible to represent a mode as a combination of two or more other modes. For example, a mode  $\mathbf{e}_1$  might be composed from two other modes,  $\mathbf{e}_2$  and  $\mathbf{e}_3$ :

$$\mathbf{e}_1 = \lambda_1 \mathbf{e}_2 + \lambda_2 \mathbf{e}_3$$

If a mode can be decomposed does it mean that the mode is **not** an elementary mode? Condition three provides a rule to determine whether a decomposition means that a given mode is an elementary mode or not. If it is only possible to decompose a given mode by introducing enzymes that are not used in the mode, then the mode is elementary. That is, is there more than one way to generate a pathway (i.e something that can sustain a steady



**Figure 18.6** Elementary mode patterns in a multi-branched pathway assuming **irreversibility** at each reaction step.

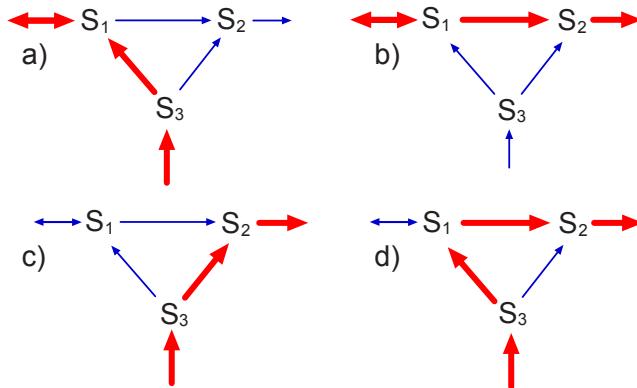
state) with the enzymes currently used in the mode? If so, then the mode is not elementary. To illustrate this subtle condition consider the pathway shown in Figure 18.9.

This pathway represents a stylized rendition of glycolysis. Step three and six are reversible and correspond to triose phosphate isomerase and glycerol 3-phosphate dehydrogenase respectively.

The network has four elementary flux modes which are shown in Figure 18.10. The elementary flux mode vectors are shown below:

$$\begin{array}{cccc}
 \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 & \mathbf{e}_4 \\
 \left[ \begin{array}{cccc}
 0 & 1 & 1 & 1 \\
 0 & 1 & 1 & 1 \\
 1 & -1 & 1 & 0 \\
 1 & 0 & 2 & 1 \\
 1 & 0 & 2 & 1 \\
 -1 & 2 & 0 & 1
 \end{array} \right] & & & (18.9)
 \end{array}$$

Note that it is possible to have negative entries in the set of elementary modes because they will correspond to the reversible steps. Of interest is the observation that the fourth vector,  $e_4 = [1 \ 1 \ 0 \ 1 \ 1 \ 1]^T$  (where  $T$  represents the transpose) can be formed from the sum of the first and second vectors (18.10). This suggests that the fourth vector is not an elementary



**Figure 18.7** Elementary mode patterns in a multi-branched pathway assuming **reversibility** at the first reaction step.

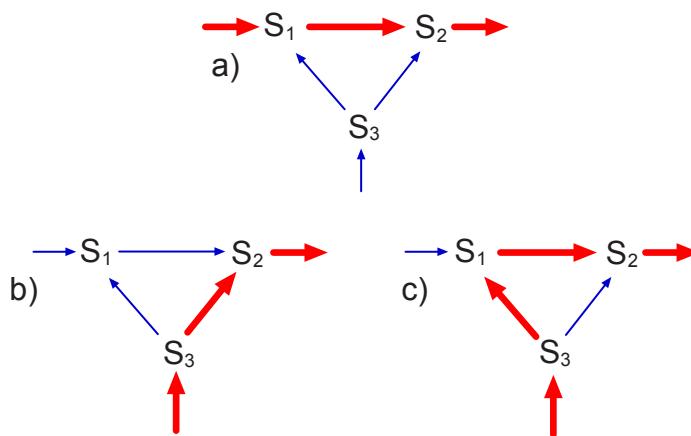
mode.

$$\begin{bmatrix} \mathbf{e}_4 \\ 1 \\ 1 \\ 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{e}_1 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \\ -1 \end{bmatrix} + \begin{bmatrix} \mathbf{e}_2 \\ 1 \\ 1 \\ -1 \\ 0 \\ 0 \\ 2 \end{bmatrix} \quad (18.10)$$

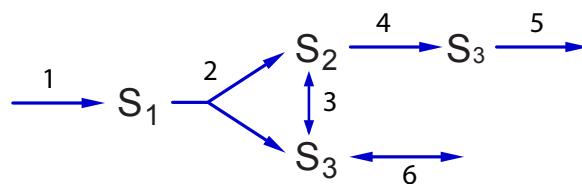
However, this decomposition only works because we have introduced a new enzyme,  $E_4$  (triose phosphate isomerase) which is not used in the second vector. It is in fact impossible to decompose  $\mathbf{e}_4$  into pathways that can sustain the steady state with only the five steps,  $E_3, E_4, E_5$  and  $E_6$ , used in the elementary mode. We conclude therefore that  $\mathbf{e}_4$  is an elementary mode.

### 18.3 Definition of a Pathway

Unlike the basis of the null space, the set of elementary modes for a given network is unique (up to an arbitrary positive scaling factor). Given the fundamental nature of elementary modes, particularly their uniqueness and non-decomposability, they are a vehicle with which to define the notion of a **pathway**. That is every elementary mode and every positive linear combination of elementary modes is by definition, a pathway. A single elementary mode can therefore be thought of as an elementary pathway. Note that the set of elementary modes will change as the set of expressed enzymes change during transitions from one cell state to another.



**Figure 18.8** Elementary Flow patterns in a simple branched pathway assuming irreversibility in first reaction step.



**Figure 18.9** Stylized Glycolytic Pathway

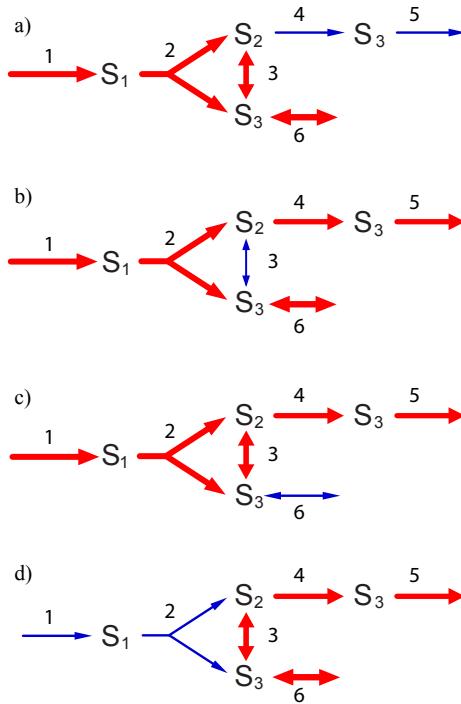
An elementary pathway is defined as a route taken by an elementary mode. A pathway is defined a linear combination of elementary modes.

## 18.4 Maximum Yield Predictions

An important application of elementary modes is finding pathways that give the maximum molar yield, that is the largest product/substrate rate ratio:

$$\text{Yield} = \frac{\text{Synthesis Rate of Product}}{\text{Consumption Rate of Substrate}} \quad (18.11)$$

In many situations the biosynthesis of a product can be achieved by a number of different pathway routes and the question then arises what are the routes that achieve the maximum



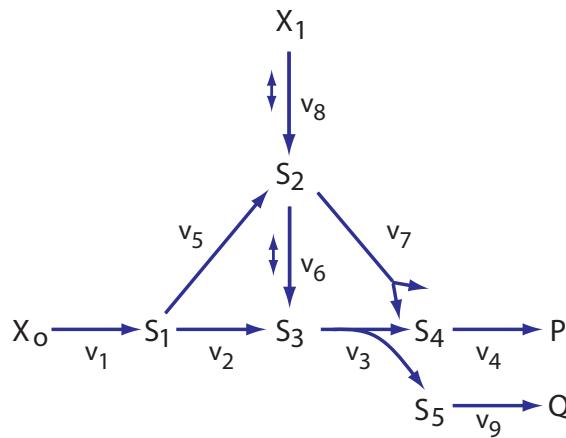
**Figure 18.10** Stylized Glycolytic Pathway illustrating the four elementary flux modes. Elementary modes are shown as bold arrows.

yield of product relative to a given starting material.

A very interesting property of elementary modes is that the set of elementary modes in a particular pathway represent the highest yielding pathways.

The argument for this is as follows. Any flux distribution can be described as a non-negative linear combination of elementary modes (18.7), for example,  $\lambda_1 \mathbf{e}_1 + \lambda_2 \mathbf{e}_2$ . The yield of a given product and substrate is the weighted average of the yields of each of the elementary modes that make up the pathway. However, the average yield will always be smaller than the elementary mode in the set that has the highest yield. The average of two numbers is always smaller than the highest of the two individual numbers. Hence, given that elementary modes cannot be decomposed, the elementary modes must represent the highest yielding pathways.

Consider the network (from [137]) shown in Figure 18.11. The stoichiometry for the network is given by equation (??):

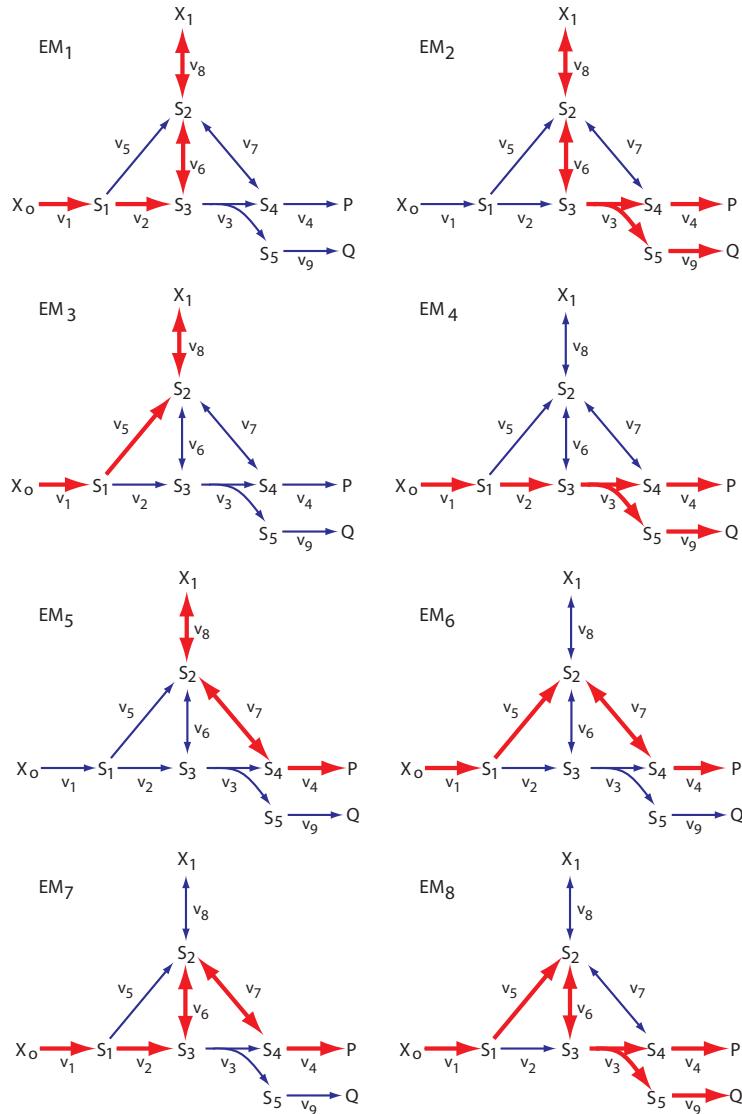


**Figure 18.11** Example Network to Illustrate Computation of Maximum Yields.  $X_o$ ,  $X_o$ ,  $P$  and  $Q$  are boundary species. Reactions  $v_6$  and  $v_8$  are reversible.

$$\mathbf{N} = \begin{matrix} & v_1 & v_2 & v_3 & v_4 & v_5 & v_6 & v_7 & v_8 & v_9 \\ S_1 & \left[ \begin{array}{ccccccccc} 1 & -1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & -1 & -1 & 0 \\ 0 & 1 & -1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 2 & 0 \end{array} \right] \\ S_2 & \\ S_3 & \\ S_4 & \\ S_5 & \end{matrix} \quad (18.12)$$

It is straight forward to show using software such as Metatool that the network in Figure 18.11 has eight elementary modes, labeled  $EM_1$ ,  $EM_2$ ,  $EM_3$ ,  $EM_4$ ,  $EM_5$ ,  $EM_6$ ,  $EM_7$  and  $EM_8$  (See Figure 18.12).

Let us suppose that we are interested in maximizing the production of product  $P$  from the feed substrate,  $X_o$ . Of the eight elementary modes, only six start with  $X_o$ . However, only four of these six result in the production of product  $P$ , that is  $EM_4$ ,  $EM_6$ ,  $EM_7$  and  $EM_8$ .



**Figure 18.12** Example Network to Illustrate Computation of Maximum Yields.  $X_o$ ,  $X_o$ ,  $P$  and  $Q$  are boundary species. Reactions  $v_6$  and  $v_8$  are reversible. The network admits eight elementary mode. Each mode is indicated in red (thickened reactions).

The four elementary modes that connect  $X_o$  to  $P$  are given below:

$$\mathbf{N} = \begin{bmatrix} & \mathbf{EM}_4 & \mathbf{EM}_6 & \mathbf{EM}_7 & \mathbf{EM}_8 \\ & 1 & 1 & 1 & 1 \\ & 1 & 0 & 1 & 0 \\ & 1 & 0 & 0 & 1 \\ & 1 & 2 & 2 & 1 \\ & 0 & 1 & 0 & 1 \\ & 0 & 0 & -1 & 1 \\ & 0 & 1 & 1 & 0 \\ & 0 & 0 & 0 & 0 \\ & 1 & 0 & 0 & 1 \end{bmatrix} \quad (18.13)$$

The question is which of these elementary modes achieves the highest yield? Equation (18.11) will allow us to compute the yields for the elementary modes. Recall that the entries in the elementary mode vectors represent relative flux values and since the yield equation is a ratio of fluxes we can use the entries in the elementary modes to compute the yield for each elementary mode. For example, consider EM<sub>4</sub>. The yield for this mode is given by  $v_4/v_1 = 1/1 = 1$ . Table 18.2 summarizes the yields for each of the four elementary modes.

EM	Yield
4	1
6	2
7	2
8	1

**Table 18.2** Yield for each elementary mode that consumes input  $X_o$  and produces product  $P$ .

From the table (Table 18.2) it should be clear that two of the modes, EM<sub>6</sub> and EM<sub>7</sub> produce twice the yield as EM<sub>4</sub> and EM<sub>8</sub>. From this information it would be logical therefore to over express the enzymes along EM<sub>6</sub> and EM<sub>7</sub> pathways. However examination of EM<sub>6</sub> and EM<sub>7</sub> shows that EM<sub>6</sub> includes four enzymatic steps whereas EM<sub>7</sub> includes five enzymatic steps. We can therefore narrow down the choice further and suggest that EM<sub>6</sub>, which has fewer steps, would be the initial target for engineering.

Having chosen the pathway to engineer we now need to determine by how much each enzyme should be over expressed. From equation (18.5) we know that in a branched system not every enzyme must be over expressed by the same amount. Instead we must compute the relative over expression in each enzyme from the known fluxes through the pathway (possibly computed using Flux Balance Analysis). We are also assured from equation (18.8) that during this engineering, none of the metabolites will change.

Flux balance analysis using linear programming can also be used to compute pathways with the highest yields by suitable adjustment of the objective function. However, elementary modes provides a systematic approach to uncovering all high yielding pathways [119, 120]. Linear programming will sometimes inadvertently uncover pathways that represent elementary modes and the work by Varma and Palsson [143, 144] on biomass yields in *E. coli* did just that.

## 18.5 Pathway Engineering

---

Most approaches used to engineer metabolic pathways stem from an intuitive understanding of how metabolism operates. For example, to increase the output of some product it seems logical to first increase the level of enzymes that are involved directly in the production of the product and secondly to reduce enzyme activities of those pathways that may divert flux away from the product pathway. This approach has been shown to work in certain cases [104]. However the method can fail due to inadvertent changes in metabolite levels that cause metabolites to increase to toxic levels. In addition changes to enzyme levels can also disrupt cofactor levels such as NAD that have a global and disruptive impact on cellular metabolism. What is needed is a more systematic approach to engineering pathways. One possible approach is described here.

The strategy is as follows:

1. Enumerate all elementary modes in the metabolic network.
2. Find all modes which end at the desired product.
3. Select one of the modes in step two for engineering. The choice of mode will depend on the number of steps (which should be minimized), the costs involved in genetically engineering each step and the yield that the mode can deliver.
4. Use equation (18.4) and (18.8) to compute the degree of over-expression of each enzyme along the elementary mode.

In theory this strategy should work, however there are a number of pitfalls. This includes the inability to up regulate all the necessary enzymes and secondly the possibility be not begin able to be precise enough when a particular enzyme needs to be up regulated by a specific amount.

## Further Reading

---

1. Wiechert W. (2001)  $^{13}\text{C}$  metabolic flux analysis. *Metabolic Engineering* Jul;3(3):195-206.

2. Stephanopoulos, Gregory (1998). "Chapter 9: Methods for the Experimental Determination of Metabolic Fluxes by Isotope Labeling". *Metabolic engineering: principles and methodologies*. San Diego: Academic Press. pp. 356–404. ISBN 0-12-666260-6.



# **Part III**

# **Appendices**



# A

## *List of Symbols and Abbreviations*

### Symbols

---

$A$	Name and amount of species
$a$	Concentration of species $A$
$s_a, s_b, \dots$	Stoichiometric amounts for species $A, B, \dots$
$c_a, c_i$	Stoichiometric coefficient for species $A$ or species $i$
$\Delta$	Change
$\delta_s$	Small change to species $S$
$\varepsilon_s^v$	Elasticity coefficient
$\xi$	Extent of reaction
$h$	Hill coefficient
$k, k_i$	Rate constant
$\Gamma$	Mass-action ratio
$\gamma$	Normalized activator concentration, $A / K_A$
$\Delta_r G$	Reaction free energy change
$\Delta_r G^\circ$	Reaction free energy change
$J$	Flux
$K_a$	Association constant
$K_d$	Dissociation constant
$K_{eq}$	Equilibrium constant
$M_i$	Fractional molar amount of species $i$ .
$T$	Total number of moles in a given conserved cycle

---

$K_H$	Half-maximal activity
$K_i$	Inhibition constant
$K_m$	Michaelis constant
$K_s$	Michaelis constant with respect to substrate
$K_p$	Michaelis constant with respect to product
$L$	Allosteric constant
$L$	Liter (US spelling)
$M$	Modifier
mol	Mole
$n_i$	Amount of substance $i$
$n$	Number of subunits
$P$	Product concentration
$P_A$	Permeability coefficient
$\alpha, \beta$	Normalized reactant concentration, $A/K_A$ , etc
$\pi, \pi_i, q, q_i$	Normalized product concentration, $P_i/K_m$ , etc
$\rho$	Disequilibrium ratio
$q$	Heat transferred
$R$	Gas constant
$\sigma$	Modifier factor
$S$	Substrate concentration or entropy depending on context
$T$	Temperature
$t$	Time
$t_{1/2}$	Half-Life
$U$	Internal energy
$v_E$	Extensive reaction rate
$v_f$	Forward reaction rate
$v_I$	Intensive reaction rate
$v_i$	$i^{\text{th}}$ reaction rate
$v_r$	Reverse reaction rate
$V$	Volume
$V_m$	Maximal velocity
$Y$	Fractional saturation
$C_e^J$	Flux Control Coefficient of flux $J$ with respect to enzyme $e$
$C_e^s$	Concentration Control Coefficient of species $s$ with respect to enzyme $e$
$\varepsilon_s^v$	Elasticity Coefficient of reaction rate $v$ with respect to species $s$

---

## Non-Mathematical Abbreviations

---

AMP, ADP, ATP	Adenine nucleotides
ATCase	Aspartate transcarbamylase
cAMP	Cyclic AMP
CI	Lambda phage repressor protein
Cro	Lambda phage repressor protein
CRP	Catabolite activator protein
CTP	Cytidine triphosphate
DNA	Deoxyribonucleic acid
F6P	Fructose-6-Phosphate
KNF	Koshland, Nemethy and Filmer model
IPTG	Isopropyl $\beta$ -D-1-thiogalactopyranoside
LacI	Lactose Operon Repressor
mRNA	Messenger RNA
MWC	Monod, Wyman and Changeux model
PEP	Phosphoenolpyruvate
PFK	Phosphofructokinase
RBS	Ribosome binding site
RNA	Ribonucleic acid
TF	Transcription factor

---



# B

## ***Math Fundamentals***

---

### **B.1 Notation**

**Sum and Product:**

$$a_1 + a_2 + a_3 + \dots + a_n = \sum_{i=1}^n a_i$$
$$a_1 \times a_2 \times a_3 \times \dots \times a_n = \prod_{i=1}^n a_i$$

**Vectors and Matrices:**

Bold lower case letters indicate vectors, for example:  $\mathbf{v}, \mathbf{s}$

Bold upper case letters indicate matrices, for example:  $\mathbf{N}, \mathbf{X}$ .

**Derivatives:**

On the left is Leibniz's notation and on the right Lagrange's notation:

$$\frac{df}{dx} \equiv f'(x)$$

$$\frac{d^2f}{dx^2} \equiv f''(x)$$

$$\frac{d^n f}{dx^n} \equiv f^{(n)}(x)$$

## B.2 Short Table of Derivatives

---

$$\frac{d}{dx}[c] = 0$$

$$\frac{d}{dx}[x] = 1$$

$$\frac{d}{dx}[cu] = c \frac{du}{dx}$$

$$\frac{d}{dx}[u + v] = \frac{du}{dx} + \frac{dv}{dx}$$

$$\frac{d}{dx}[uv] = u \frac{dv}{dx} + v \frac{du}{dx}$$

$$\frac{d}{dx}[u/v] = \frac{v \frac{du}{dx} - u \frac{dv}{dx}}{v^2}$$

$$\frac{d}{dx}[u^n] = nu^{n-1} \frac{du}{dx}$$

$$\frac{d}{dx}[f(u)] = \frac{df}{du} f(u) \frac{du}{dx}$$

$$\frac{d}{dx}[\ln u] = \frac{1}{u} \frac{du}{dx}$$

$$\frac{de^u}{dx} = e^u \frac{du}{dx}$$

$$\frac{d}{dx}[\sin(u)] = \cos(u) \frac{du}{dx}$$

$$\frac{d}{dx}[\cos(u)] = -\sin(u) \frac{du}{dx}$$

## B.3 Logarithms

---

$$\log(AB) = \log(A) + \log(B)$$

$$\log(A/B) = \log(A) - \log(B)$$

$$\log(A^n) = n \log(A)$$

$$x^n \times x^m = x^{n+m}$$

$$\frac{x^n}{n^m} = x^{n-m}$$

$$(x^n)^m = x^{n \times m}$$

## B.4 Partial Derivatives

---

If the value of a given function depends on two variables then we write this function in the form:

$$u = f(x, y)$$

If it is possible to change  $x$  without affecting  $y$  then  $x$  and  $y$  are called independent variables. The rate of change of  $u$  with respect to  $x$  when  $x$  varies but  $y$  remains constant is called the

**partial derivative** of  $u$  with respect to  $x$ . Partial derivatives are denoted using the partial symbol,  $\partial$ . That is the partial derivative of  $u$  with respect to  $x$  is:

$$\frac{\partial u}{\partial x}$$

Likewise the partial derivative of  $u$  with respect to  $y$  is given by:

$$\frac{\partial u}{\partial y}$$

To find a partial derivative, we simply differentiate with respect to the variable of interest while treating the remaining variables as constants. For example, the reaction rate for a given reaction is  $v = k_1 s - k_2 p$ , where  $S$  is the reactant,  $P$  the product and  $k_1$  and  $k_2$  the rate constants. In a controlled environment we should in principle be able to change  $S$  and  $P$  independently. Therefore we can write down the partial derivatives of the reaction rate with respect to  $S$  and  $P$  as follows:

$$\frac{\partial v}{\partial s} = k_1$$

$$\frac{\partial v}{\partial p} = -k_2$$

In order to indicate what variables are kept constant in the partial derivative the following notation is sometimes used, particularly in thermodynamics:

$$\left( \frac{\partial v}{\partial s} \right)_p = k_1$$

$$\left( \frac{\partial v}{\partial p} \right)_s = -k_2$$

Or for function with many variables,  $x, y, z, \dots$ , the notation would extend to:

$$\left( \frac{\partial u}{\partial x} \right)_{y,z,\dots}$$

Like derivatives, partial derivatives are defined in terms of a limit. For example the partial derivatives for the function,  $f(x, y)$  are defined as:

$$\frac{\partial f(x, y)}{\partial x} = \lim_{h \rightarrow 0} \frac{f(x + h, y) - f(x, y)}{h}$$

$$\frac{\partial f(x, y)}{\partial y} = \lim_{h \rightarrow 0} \frac{f(x, y + h) - f(x, y)}{h}$$

The graphical interpretation of a partial derivative,  $\partial f(x, y)/\partial x$  is that it represents the slope of the function,  $f(x, y)$  in the  $x$  direction.

## B.5 Differential Equations

---

Differential equations are equations that contain derivatives. For example, the following is a differential equation:

$$\frac{dy}{dx} + y^2 = 0$$

An **ordinary differential equation** is where the derivatives are functions of the same variable. For example, the following equations are ordinary differential equations:

$$\frac{dy}{dx} = ay$$

$$\frac{dy}{dx} = 2x + 3y - 8$$

$$\frac{d^2y}{dx^2} - x \frac{du}{dx} = 0$$

A differential equation expressed in terms of the first derivative ( $dy/dx$ ) is called a first-order differential equation. A differential equation that is expression in terms of second order derivatives ( $d^2y/dx^2$ ) is called a second-order differential equation. When solving differential equations the objective is to find the function  $y(x)$  such that when differentiated gives the original differential equation. For example the solution to:

$$\frac{dy}{dx} = ay$$

is

$$y = y_0 e^{ax} \quad (\text{B.1})$$

If we differentiate the solution (B.1) we get back the original differential equation.

Differential equations are used very often to model physical systems where they describe the rate of change of some variable with respect to time,  $t$ . The reason why they are used is because we may not explicitly know the solution  $y(t)$  but we will often know the rate of change the variable has at any given moment in time,  $dy/dt$ . This means we can at least obtain a numerical solution to  $y(t)$  even if the analytical solution is unobtainable.

Differential equations can be further classified as autonomous or non-autonomous. Autonomous differential equations are the most common in biochemical models. These equations do not depend on time, that is the right-hand side of the differential equation has no terms relating explicitly to time. For example equation B.2 is autonomous equation B.3 is non-autonomous:

$$\frac{dx}{dt} = x^2 + 10 \quad (\text{B.2})$$

$$\frac{dx}{dt} = x^2 + t - 5 \quad (\text{B.3})$$

A **partial differential equation** is one where the derivatives are functions of more than one derivative. For example the equation is a partial differential equation:

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} = \frac{\partial p}{\partial x}$$

Note the use of the partial  $d$  ( $\partial$ ) in the partial differential equation to indicate that the function  $u$  is differentiated with respect to more than one variable. The partial derivative also indicates that when a derivative is made, other constants are assumed to be held constant.

## B.6 Polynomials

---

An equation of the form:

$$a_n s^n + a_{n-1} s^{n-1} + \dots + a_1 s + a_o$$

is called a polynomial. Such equations involve only addition, subtraction, multiplication and variables raised to positive integer powers. Given this definition, the equations,  $3s^{-2}$  and  $2/(s + 2)$  are not polynomials because the exponents are negative. The degree of a polynomial is the highest integer power, for example the degree of the polynomial  $7s^4 + 3s^2 + 2s$  is four. Of particular interest in many applications are the roots of the polynomial function:

$$F(s) = 0 = a_n s^n + a_{n-1} s^{n-1} + \dots + a_1 s + a_o$$

An important theorem related to polynomials is the factor theorem. This states that given a polynomial  $F(s)$ , then  $(s - p)$  is a factor of  $F(s)$  if and only if  $F(p) = 0$ . That is  $p$  is a root of  $F(s)$ . For example, if  $F(s) = s^3 - 8$  then  $s - 2$  is a factor because 2 is a root of  $F(s)$ . In general we can factor a polynomial  $F(s)$  of  $n^{\text{th}}$  degree into  $n$  linear factors:

$$F(s) = A(s - p_n)(s - p_{n-1}) \dots (s - p_o) = 0$$

Consider the second-order polynomial where we factor and collect terms:

$$Q_2(s) = s^2 + a_1 s + a_o = (s - p_1)(s - p_2) = s^2 - (p_1 + p_2)s + p_1 p_2 \quad (\text{B.4})$$

Likewise we can do the same with a third-order polynomial:

$$\begin{aligned} Q_3(s) &= s^3 + a_2 s^2 + a_1 s + a_o = (s - p_1)(s - p_2)(s - p_3) \\ &= s^3 - (p_1 + p_2 + p_3)s^2 + (p_1 p_2 + p_1 p_3 + p_2 p_3)s - p_1 p_2 p_3 \end{aligned} \quad (\text{B.5})$$

By induction we can extend this to an  $n^{\text{th}}$ -order polynomial:

$$Q_n(s) = s^n + a_{n-1} s^{n-1} + \dots + a_1 s + a_o \quad (\text{B.6})$$

so that the coefficients are given by:

$a_{n-1}$  = the negative of the sum of all roots

$a_{n-2}$  = the sum of products of all possible combinations taken two at a time

$a_{n-3}$  = the negative sum of products of all possible combinations taken three at a time

⋮

$a_0 = (-1)^n$  multiplied by the product of all roots

From this analysis we can discern a useful pattern. First we make sure that the polynomial must be in the form B.6, that is all coefficients are non-zero. The one exception is if  $a_o$  is zero, we can divide throughout by  $s$  and this we yield the form B.6.

Looking at the second-order polynomial, we can see that if the roots  $p_i$  are negative then the terms  $-(p_1 + p_2)$  and  $p_1 p_2$  must be positive. For the general case,  $Q_n(s)$  we can make the same assertion, however note that the reverse statement, if all the coefficients are positive then all the roots will be negative is **not** necessarily true. This rule applies to the general case  $Q_n(s)$  and can be stated as.

A **necessary** but not sufficient condition for negative roots is that all coefficients must be positive. This includes the case when the roots are complex where the real parts will be negative.

Note that for all roots to be negative no coefficient can be zero.

1. If any coefficient  $a_i$  is equal to zero then one or more roots will be positive
2. If any coefficient  $a_i$  is negative, then at least one root is positive
3. If all coefficients are positive then roots may be negative or positive

Note that for a polynomial higher than second order, the rules provide only a necessary condition that all roots will be negative. For higher order polynomials, additional rules must be invoked such as constructing the Routh-Hurwitz table. Using the Routh method the following rules can be derived for 2nd 3rd and 4th order polynomials.

Quadratic:  $a_2 s^2 + a_1 s + a_o = 0$

All roots are negative if

$$a_2 > 0, a_1 > 0, a_o > 0$$

Cubic:  $a_3 s^3 + a_2 s^2 + a_1 s + a_o = 0$

All roots are negative if:

$$a_3 > 0, a_2 > 0, a_1 > 0, a_o > 0$$

and

$$a_1 a_2 > a_0 a_3$$

Quartic:  $a_4 s^4 + a_3 s^3 + a_2 s^2 + a_1 s + a_0 = 0$

All roots are negative if:

$$a_4 > 0, a_3 > 0, a_2 > 0, a_1 > 0, a_0 > 0$$

and

$$a_1 a_2 - a_0 a_3 > 0$$

and

$$a_1 a_2 a_3 - a_1^2 a_4 - a_0 a_3^2 > 0$$

## B.7 Taylor Series

---

Expressions like  $1 + 2x + 6x^2$  and  $2 + 4x + x^2 - 3x^3$  that consist of the sum of a number of terms raised to a positive power are called polynomials. The only operations allowed in a polynomial are addition, subtraction, multiplication and non-negative integer powers. One of the simplest polynomials is the straight line,  $y = a + bx$ , termed a polynomial of first degree. The coefficients,  $a$  and  $b$  can be chosen so that the line will pass through any two points. That is we can express any straight line using  $y = a + bx$ . Similarly for a polynomial of second degree,  $y = a + bx + cx^2$ , a parabola, we can choose the constants,  $a, b$  and  $c$  so that the curve can pass through any three points.

It follows that we can find a polynomial equation of  $n^{\text{th}}$  degree that will pass through any  $n + 1$  points. If the polynomial has an infinite number of terms, then we can imagine that the polynomial can be made to follow any function,  $f(x)$  by suitable adjustment of the polynomial coefficients. Although this statement may not always be true, in many cases it is which makes the polynomial series very useful.

A polynomial of infinite degree is called a polynomial series:

$$f(x) = c_0 + c_1 x + c_2 x^2 + c_3 x^3 + \dots$$

The question is, how can we find the polynomial series that will represent a given function, for example  $\sin(x)$ ? To answer this we have to determine the constants,  $c_0, c_1$  etc. in the polynomial equation. Let us assume that we wish to know the value of  $\sin(x)$  at  $x = 0$  using a polynomial series. At  $x = 0$ , all terms vanish except for  $c_0$ , therefore at  $x = 0$ :

$$f(0) = c_0$$

We can therefore interpret the first constant,  $c_0$  as the value of the function at  $x = 0$ . What about  $c_1$ ? Let us take the derivative of the series, that is:

$$f'(x) = c_1 + 2c_2 x + 3c_3 x^2 + \dots$$

Function	Second order approximation
$\frac{1}{1+x}$	$1 + x + x^2$
$\sqrt{1+x}$	$1 + \frac{x}{2} + \frac{x^2}{8}$
$\sin(x)$	$x$

**Table B.1** Examples of common approximations

If we set  $x = 0$ , we find that:

$$f'(0) = c_1$$

That is the second constant,  $c_1$ , in the polynomial series is the first derivative of the function. If we take the second derivative we can also show that at  $x = 0$ ,  $f''(x_0) = 2c_2$ , that is  $c_2 = f''(0)/2$ . For the third derivative we can show that  $f'''(0) = 3(2)c_3$ , that is  $c_3 = f'''(0)/(3!)$ . This pattern continues for the remaining terms in the polynomial so that we can now write:

$$f(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \frac{f'''(0)}{3!}x^3 + \dots$$

This series is called the **Maclaurin series** for the function,  $f(x)$ . It approximates the function around the specific value of  $x = 0$ . To illustrate the use of the Maclaurin series consider expanding  $\sin(x)$  around  $x = 0$ .  $f(0)$  will equal  $\sin(0) = 0$ .  $f'(0) = \cos(0) = 1$  and so on. We can therefore write the series as:

$$\sin(x) = 0 + 1x + 0 - \frac{1}{3!}x^3 + 0 + \frac{1}{5!}x^5 - \dots$$

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots$$

What if we wanted to approximate a function about an arbitrary value of  $x$ ? To do this we would use the Taylor series which is a generalization of the Maclaurin series. The **Taylor series** is defined by:

$$f(x) = f(x_o) + \frac{\partial f}{\partial x_o}(x - x_o) + \frac{1}{2!} \frac{\partial^2 f}{\partial x_o^2}(x - x_o)^2 + \dots + \frac{1}{n!} \frac{\partial^n f}{\partial x_o^n}(x - x_o)^n + \dots \quad (\text{B.7})$$

where the approximation is now centered on  $x_o$ . If we set  $x_o$  equal to zero we will obtain the Maclaurin series.

## B.8 Total Derivative

---

Consider the function:

$$f(t) = f(x(t), y(t))$$

The derivative of  $f(t)$  with respect to  $t$ , is given by the chain rule:

$$\frac{df}{dt} = \frac{\partial f}{\partial x} \frac{dx}{dt} + \frac{\partial f}{\partial y} \frac{dy}{dt}$$

Note the use of partial derivatives. This equation is often abbreviated to:

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy$$

where it is called the **total derivative**. Often the variable,  $t$  is not specified in the total derivative. Operationally the total derivative computes the change in  $f$ , given small changes in  $x$  and  $y$ .

## B.9 Eigenvalues and Eigenvectors

---

A square matrix such as  $\mathbf{A}$  can be used to transform a given vector,  $\mathbf{v}$  in specific ways. For example, if the matrix  $\mathbf{A}$  is:

$$\begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}$$

then the result of multiplying  $\mathbf{A}$  into  $\mathbf{v}$  will yield a vector that is similar to  $\mathbf{v}$  but where the first element is scaled by 2 and the second element by 4.

For an arbitrary square matrix, if it is possible to find a vector  $\mathbf{v}$  such that when we multiply the vector by  $\mathbf{A}$  we get a scaled version of  $\mathbf{v}$ , then we call the vector  $\mathbf{v}$  the **eigenvector** of  $\mathbf{A}$  and the scaling value, the **eigenvalue** of  $\mathbf{A}$ . For a matrix of dimension  $n$ , there will be at most  $n$  eigenvalues and  $n$  eigenvectors. In the case of the simple example above the eigenvalues must be 2 and 4 respectively while the two eigenvector will be:

$$\begin{bmatrix} \alpha \\ 0 \end{bmatrix} \quad \begin{bmatrix} 0 \\ \alpha \end{bmatrix}$$

The definition of an eigenvector and eigenvalue is often given in the form:

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$$

We can rearrange this equation as follows:

$$\begin{aligned} \mathbf{A}\mathbf{v} &= \lambda\mathbf{I}\mathbf{v} \\ \mathbf{A}\mathbf{v} - \lambda\mathbf{I}\mathbf{v} &= 0 \\ (\mathbf{A} - \lambda\mathbf{I})\mathbf{v} &= 0 \end{aligned}$$

From linear algebra we know that there will be non-zero solutions to  $(A - \lambda I)v = 0$  if  $\det(A - \lambda I) = 0$ . We can use this observation to compute the eigenvalues and eigenvectors of a matrix. For example consider the matrix:

$$\begin{bmatrix} 3 & 6 \\ 1 & 4 \end{bmatrix}$$

Computing  $A - \lambda I$  yields:

$$\begin{aligned} A - \lambda I &= \begin{bmatrix} 3 - \lambda & 6 \\ 1 & 4 - \lambda \end{bmatrix} \\ \det(A - \lambda I) &= (3 - \lambda)(4 - \lambda) - 6 \\ &= \lambda^2 - 7\lambda + 6 \\ &= (\lambda - 6)(\lambda - 1) \end{aligned}$$

The eigenvalues are therefore 6 and 1. With two eigenvalues there will be two eigenvectors. First we consider  $\lambda = 6$ .

$$\begin{aligned} (A - \lambda I)v &= 0 \\ \left( \begin{bmatrix} 3 & 6 \\ 1 & 4 \end{bmatrix} - \begin{bmatrix} 6 & 0 \\ 0 & 6 \end{bmatrix} \right) v &= 0 \\ \begin{bmatrix} -3 & 6 \\ 1 & -2 \end{bmatrix} v &= 0 \end{aligned}$$

By inspection we can see that the eigenvector is:

$$\begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

satisfied this equation. Likewise we can do the same for the other eigenvalue,  $\lambda = 1$  where the corresponding eigenvector is found to be:

$$\begin{bmatrix} -3 \\ 1 \end{bmatrix}$$

### de Moivre's Theorem

A complex number can be expressed in terms of sine and cosines (polar form):

$$a + bj = z = r(\cos \theta + j \sin \theta)$$

The square of  $z$  is given by:

$$\begin{aligned} z^2 &= r^2(\cos \theta + j \sin \theta)^2 \\ &= r^2(\cos^2 \theta - \sin^2 \theta + 2j \cos \theta \sin \theta) \\ &= r^2(\cos 2\theta + \sin 2\theta) \end{aligned}$$

This result can be generalized to any integer  $n$  such that:

$$z^n = r^n(\cos n\theta + j \sin n\theta)$$

The above equation is called de Moivre's Theorem. The theorem has a number of applications, for example we can use de Moivre's theorem to evaluate  $(1+j)^{16}$ . First we determine the magnitude and angle for the complex number  $(1+j)$ . The magnitude is  $\sqrt{1+1} = \sqrt{2}$ . The angle is  $\tan^{-1} 1/1 = \pi/4$ . We can therefore write:

$$\begin{aligned} (1+j)^{16} &= r^2(\cos(n\theta) + j \sin(n\theta)) \\ &= (\sqrt{2})^{16}(\cos(16\pi/4) + j \sin(16\pi/4)) \\ &= 256(1+0) = 256 \end{aligned}$$

## Finding Roots

Of more interest to the work in this book is using de Moivre's theorem to find the roots of complex numbers. For example, let us find the roots to  $\sqrt[3]{27j}$ . First we write  $27j$  in polar form. The magnitude  $r = \sqrt{0^2 + (27)^2} = 27$ . The angle is given by  $\tan^{-1} 27/1 = \pi/2$ . We can therefore write  $27j$  as:

$$27j = 27 \left( \cos \frac{\pi}{2} + j \sin \frac{\pi}{2} \right)$$

Since we are looking for  $\sqrt[3]{27j}$  we can also look at this term in the following equivalent form:  $z^3 = 27j$ . By de Moivre's theorem:

$$r^3(\cos 3\theta + j \sin 3\theta) = 27j = 27 \left( \cos \frac{\pi}{2} + j \sin \frac{\pi}{2} \right)$$

What possible values are there for  $\theta$ ? It is true that:

$$\cos 3\theta = \cos \frac{\pi}{2} \quad \text{and} \quad \sin 3\theta = \sin \frac{\pi}{2}$$

$\cos 3\theta$  will repeat itself every  $2\pi/3$ . Therefore it must be true that:

$$3\theta = \frac{\pi}{2} + 2\pi k$$

where  $k$  is an integer. Since there are three possible roots we can set  $k$  to 0, 1, and 2 and evaluate the roots in turn.

$k = 0$ ,  $3\theta = \pi/2$ , therefore  $\theta = \pi/6$

$$\begin{aligned} z_1 &= 3(\cos \frac{\pi}{6} + j \sin \frac{\pi}{6}) \\ &= 3 \left( \frac{\sqrt{3}}{2} + j \frac{1}{2} \right) \\ &= \frac{3\sqrt{3}}{2} + j \frac{3}{2} \end{aligned}$$

$k = 1$ ,  $3\theta = \pi/2 + 2\pi(1)$ , therefore  $\theta = 5\pi/6$

$$\begin{aligned} z_1 &= 3 \left( \cos \frac{5\pi}{6} + j \sin \frac{5\pi}{6} \right) \\ &= 3 \left( \frac{-\sqrt{3}}{2} + j \frac{1}{2} \right) \\ &= -\frac{3\sqrt{3}}{2} + j \frac{3}{2} \end{aligned}$$

$k = 2$ ,  $3\theta = \pi/2 + 2\pi(2)$ , therefore  $\theta = 9\pi/6 = 3\pi/2$

$$\begin{aligned} z_1 &= 3 \left( \cos \frac{3\pi}{2} + j \sin \frac{3\pi}{2} \right) \\ &= 3(0 + j(-1)) \\ &= -3j \end{aligned}$$

Setting  $k = 3$  gives the same result as  $k = 0$ , setting  $k = 4$  gives the same result as  $k = 1$  and the therefore the three unique cube roots of  $-27j$  are:

$$\begin{aligned} z_1 &= \frac{3\sqrt{3}}{2} + j \frac{3}{2} \\ z_2 &= -\frac{3\sqrt{3}}{2} + j \frac{3}{2} \\ z_3 &= -3j \end{aligned}$$

In general the  $n^{th}$  root of  $z$  is given by:

$$\begin{aligned} \sqrt[n]{z} &= \sqrt[n]{r} \left( \cos \left( \frac{\theta + 2\pi k}{n} \right) + j \sin \left( \frac{\theta + 2\pi k}{n} \right) \right) \\ \text{for } k &= 0, 1, 2, \dots, n-1 \end{aligned}$$

---

**Further Reading**

---

1. Smail LL (1953) Analytical Geometry and Calculus. Appleton-Century-Crofts ISBN: 978-0982477311



# C

## *Control Equations*

### C.1 Linear Pathways

---

#### Two Step pathway



**Figure C.1** Two Step Pathway.

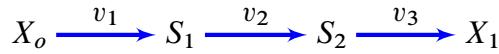
$$C_{e_1}^J = \frac{\varepsilon_1^2}{\varepsilon_1^2 - \varepsilon_1^1}$$

$$C_{e_2}^J = -\frac{\varepsilon_1^1}{\varepsilon_1^2 - \varepsilon_1^1}$$

$$C_{e_1}^{s_1} = \frac{1}{\varepsilon_1^2 - \varepsilon_1^1}$$

$$C_{e_2}^{s_1} = -\frac{1}{\varepsilon_1^2 - \varepsilon_1^1}$$

### Three Step pathway



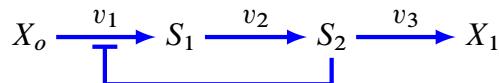
**Figure C.2** Three Step Pathway.

The denominator,  $D$  is given by:

$$D = \varepsilon_1^2 \varepsilon_2^3 - \varepsilon_1^1 \varepsilon_2^3 + \varepsilon_1^1 \varepsilon_2^2$$

$$\begin{aligned} C_{e_1}^J &= \frac{\varepsilon_1^2 \varepsilon_2^3}{D} & C_{e_2}^J &= -\frac{\varepsilon_1^1 \varepsilon_2^3}{D} & C_{e_3}^J &= \frac{\varepsilon_1^1 \varepsilon_2^2}{D} \\ C_{e_1}^{s_1} &= \frac{\varepsilon_2^3 - \varepsilon_2^2}{D} & C_{e_1}^{s_2} &= \frac{\varepsilon_1^2}{D} \\ C_{e_2}^{s_1} &= -\frac{\varepsilon_2^3}{D} & C_{e_2}^{s_2} &= -\frac{\varepsilon_1^1}{D} \\ C_{e_3}^{s_1} &= \frac{\varepsilon_2^2}{D} & C_{e_3}^{s_2} &= \frac{\varepsilon_1^1 - \varepsilon_1^2}{D} \end{aligned}$$

### Three Step pathway with Negative Feedback

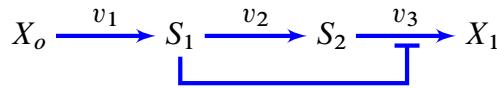


**Figure C.3** Three Step Pathway.

$$D = \varepsilon_1^1 \varepsilon_2^2 - \varepsilon_1^1 \varepsilon_2^3 + \varepsilon_1^2 \varepsilon_2^3 - \varepsilon_2^1 \varepsilon_1^2$$

$$\begin{aligned} C_{e_1}^J &= \frac{\varepsilon_1^2 \varepsilon_2^3}{D} & C_{e_2}^J &= -\frac{\varepsilon_1^1 \varepsilon_2^3}{D} & C_{e_3}^J &= \frac{\varepsilon_1^1 \varepsilon_2^2 - \varepsilon_2^1 \varepsilon_1^2}{D} \\ C_{e_1}^{s_1} &= \frac{\varepsilon_2^3 - \varepsilon_2^2}{D} & C_{e_2}^{s_1} &= -\frac{\varepsilon_2^3 - \varepsilon_2^1}{D} & C_{e_3}^{s_1} &= \frac{\varepsilon_2^2 - \varepsilon_2^1}{D} \\ C_{e_1}^{s_2} &= \frac{\varepsilon_1^2}{D} & C_{e_2}^{s_2} &= -\frac{\varepsilon_1^1}{D} & C_{e_3}^{s_2} &= \frac{\varepsilon_1^1 - \varepsilon_1^2}{D} \end{aligned}$$

### Three Step Pathway with Feedforward Loop



**Figure C.4** Three Step Pathway.

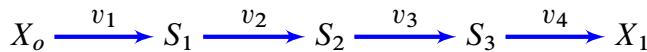
$$D = \varepsilon_1^1 \varepsilon_2^2 + \varepsilon_1^2 \varepsilon_2^2 - \varepsilon_1^1 \varepsilon_2^3 + \varepsilon_1^2 \varepsilon_2^3$$

$$C_{e1}^J = \frac{\varepsilon_2^2 \varepsilon_1^3 + \varepsilon_1^2 \varepsilon_2^3}{D} \quad C_{e2}^J = -\frac{\varepsilon_1^1 \varepsilon_2^3}{D} \quad C_{e3}^J = \frac{\varepsilon_1^1 \varepsilon_2^2}{D}$$

$$C_{e1}^{s1} = \frac{\varepsilon_2^3 - \varepsilon_2^2}{D} \quad C_{e2}^{s1} = -\frac{\varepsilon_2^3}{D} \quad C_{e3}^{s1} = \frac{\varepsilon_2^2}{D}$$

$$C_{e1}^{s2} = \frac{\varepsilon_1^2 + \varepsilon_1^3}{D} \quad C_{e2}^{s2} = \frac{\varepsilon_1^1 + \varepsilon_1^3}{D} \quad C_{e3}^{s2} = \frac{\varepsilon_1^1 - \varepsilon_1^2}{D}$$

### Four Step pathway



**Figure C.5** Four Step Pathway.

$$D = \varepsilon_1^1 \varepsilon_2^2 \varepsilon_3^3 - \varepsilon_1^1 \varepsilon_2^2 \varepsilon_3^4 + \varepsilon_1^1 \varepsilon_2^3 \varepsilon_3^4 - \varepsilon_1^2 \varepsilon_2^3 \varepsilon_3^4$$

$$C_{e1}^J = -\frac{\varepsilon_1^2 \varepsilon_2^3 \varepsilon_3^4}{D} \quad C_{e2}^J = \frac{\varepsilon_1^1 \varepsilon_2^3 \varepsilon_3^4}{D} \quad C_{e3}^J = -\frac{\varepsilon_1^1 \varepsilon_2^2 \varepsilon_3^4}{D} \quad C_{e4}^J = \frac{\varepsilon_1^1 \varepsilon_2^2 \varepsilon_3^3}{D}$$

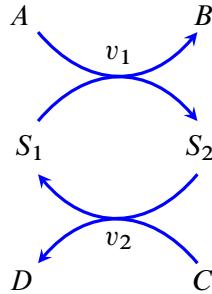
$$C_{e1}^{s1} = \frac{-\varepsilon_2^2 \varepsilon_3^3 + \varepsilon_2^2 \varepsilon_3^4 - \varepsilon_2^3 \varepsilon_3^4}{D} \quad C_{e2}^{s1} = -\frac{\varepsilon_2^3 \varepsilon_3^4}{D} \quad C_{e3}^{s1} = -\frac{\varepsilon_2^2 \varepsilon_3^4}{D} \quad C_{e4}^{s1} = \frac{\varepsilon_2^2 \varepsilon_3^3}{D}$$

$$C_{e1}^{s2} = \frac{\varepsilon_1^2 \varepsilon_3^3 - \varepsilon_1^2 \varepsilon_3^4}{D} \quad C_{e2}^{s2} = \frac{-\varepsilon_1^1 \varepsilon_3^3 + \varepsilon_1^1 \varepsilon_3^4}{D} \quad C_{e3}^{s2} = \frac{-\varepsilon_1^1 \varepsilon_3^4 + \varepsilon_1^2 \varepsilon_3^4}{D} \quad C_{e4}^{s2} = \frac{\varepsilon_1^1 \varepsilon_3^3 - \varepsilon_1^2 \varepsilon_3^3}{D}$$

$$C_{e1}^{s3} = -\frac{\varepsilon_1^2 \varepsilon_2^3}{D} \quad C_{e2}^{s3} = \frac{\varepsilon_1^1 \varepsilon_2^3}{D} \quad C_{e3}^{s2} = -\frac{\varepsilon_1^1 \varepsilon_2^2}{D} \quad C_{e4}^{s2} = \frac{\varepsilon_1^1 \varepsilon_2^2 - \varepsilon_1^1 \varepsilon_2^3 + \varepsilon_1^2 \varepsilon_2^3}{D}$$

## C.2 Cycles

---



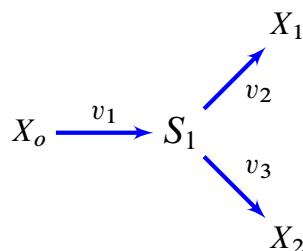
**Figure C.6** Simple Conserved cycle where  $s_1 + s_2 = \text{constant} = M_t$ .

$$C_{e_1}^{s_2} = \frac{s_1}{\varepsilon_1^1 s_2 + \varepsilon_2^2 s_1}$$

$$C_{e_2}^{s_2} = -\frac{s_1}{\varepsilon_1^1 s_2 + \varepsilon_2^2 s_1}$$

## C.3 Branches

---



**Figure C.7** Three Step Pathway.

Let  $\alpha = \frac{v_2}{v_1}$  and  $\varepsilon_1 \equiv \varepsilon_1^1 \quad \varepsilon_2 \equiv \varepsilon_1^2 \quad \varepsilon_3 \equiv \varepsilon_1^3$

$$d = \varepsilon_2 \alpha + \varepsilon_3 (1 - \alpha) - \varepsilon_1$$

$$\begin{aligned}
 C_{e_1}^{J_1} &= \frac{\varepsilon_3(1-\alpha) + \varepsilon_2\alpha}{d} & C_{e_2}^{J_2} &= \frac{\varepsilon_3(1-\alpha) - \varepsilon_1}{d} > 0 \\
 C_{e_1}^{J_1} &= \frac{-\varepsilon_1\alpha}{d} & C_{e_1}^{J_2} &= \frac{\varepsilon_2}{d} > 0 \\
 C_{e_1}^{J_1} &= \frac{-\varepsilon_1(1-\alpha) + \varepsilon_2\alpha}{d} & C_{e_3}^{J_2} &= \frac{-\varepsilon_2(1-\alpha)}{d} < 0
 \end{aligned}$$

For the concentration control coefficients:

$$\begin{aligned}
 C_{e_1}^s &= \frac{1}{\varepsilon_2\alpha + \varepsilon_3(1-\alpha) - \varepsilon_1} \\
 C_{e_2}^s &= \frac{-\alpha}{\varepsilon_2\alpha + \varepsilon_3(1-\alpha) - \varepsilon_1} \\
 C_{e_3}^s &= \frac{-(1-\alpha)}{\varepsilon_2\alpha + \varepsilon_3(1-\alpha) - \varepsilon_1}
 \end{aligned}$$



# D

## *Modeling with Python*

---

### D.1 Introduction

---

In this appendix a brief description of the Python programming language will be given plus a brief introduction to the Antimony reaction network format and libRoadRunner.

**Python** Python is an easy to learn general purpose interactive programming language. It has similar usability characteristics to Matlab or Basic. As such it is a good language to use for doing pathway simulations and is easily learned by new users. In recent years Python has also become more widely used as a general purpose scientific programming language and now supports many useful libraries and tools for modelers. All the scripts we provide in this book are written in Python.

**Antimony** SBML has become a de facto standard for exchanging models of biological pathways. Any tool we use should therefore be able to support SBML. However SBML is a computer readable language and it is not easy for humans to read or write SBML. Instead more human readable formats have been developed. In this text book we will be using the Antimony pathway description language [125]. Models can be described in Antimony then converted to SBML or vice versa.

**libRoadRunner** To support SBML from within Python we developed a C/C++ simulation library called libRoadRunner [127] that can read and run models based on SBML. In order to use libRoadRunner within Python, we also provide a Python interface that makes it easy to carry out simulations with Python.

---

**Spyder** Integration of the various tools including Python is achieved by using spyder2 (<https://code.google.com/p/spyderlib/>). Spyder2 offers a Matlab like experience in a friendly, cross-platform environment.

## D.2 Introduction to Python

---

One great advantage of the Python language is that it runs on many computer platforms, most notably Windows, Mac and Linux and is freely downloadable from the Python web site. To execute Python code we will need access to what is often referred to as a Python IDE (Integrated Development Environment). In the Python world there are many IDEs to choose from, ranging from very simple consoles to sophisticated development systems that includes documentation, debuggers and other visual aids. In this book we use the cross-platform IDE called spyder2 (<https://code.google.com/p/spyderlib/>).

The best way to learn Python is to download a copy and start using it. We have prepared installers that install all the relevant components you need, these can be found at [tellurium.analogmachine.org](http://tellurium.analogmachine.org). The Tellurium distribution includes some additional helper routines which can make life easier for new users. The Tellurium version can be downloaded for Mac and Windows computers. We will use the Windows version here. To download the installer go to the web site [tellurium.analogmachine.org](http://tellurium.analogmachine.org), and click on the first link you see called *Download Windows version here*. Run the installer and follow the instructions.

Once Tellurium is installed go to the start menu, find Tellurium and select the application call Tellurium spyder. If successful you should see something like the screen shot in Figure D.1 but without the plotting window. The screen-shot shows three important elements, on the left we see an editor, this is where models can be edited. On the lower right is the Python console where Python commands can be entered. At the top right we show plotting window that illustrates some output from a simulation. For those familiar with IPython, the latest version of spyder2 supports the IPython console directly.

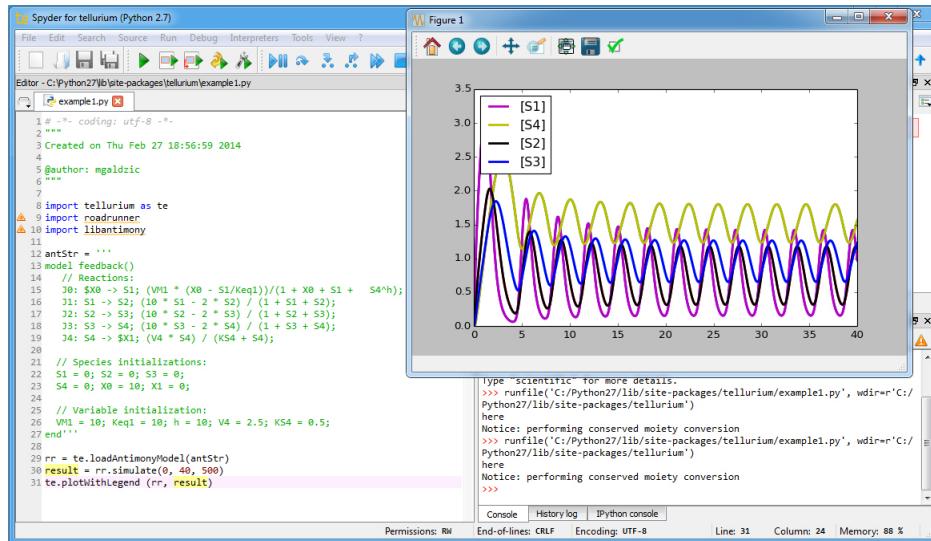
Once you have started the Tellurium IDE, let us focus on the Python console at the bottom right of the application. A screen-shot of the console is shown in Figure D.2.

The >>> symbol marks the place where you can type commands. The following examples are based on Python 2.7. To add two numbers, say 2 + 5, we would type the following:

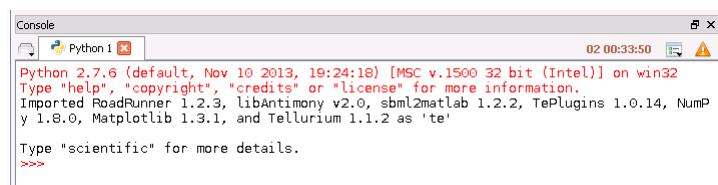
```
>>> print 2 + 5
7
>>>
```

**Listing D.1** Simple Arithmetic

Just like Matlab or Basic we can assign values to variables and use those variables in other calculations:



**Figure D.1** Screen-shot of Tellurium, showing editor on the left, Python console bottom right and plotting window top-right.



**Figure D.2** Screen-shot of Tellurium, focusing on the Python console.

```
>>> a = 2
>>> b = 5
>>> c = a + b
>>> print c
7
>>>
```

### Listing D.2 Assigning values to variables

The types of values we can assign to variables include values such as integers, floating point numbers, Booleans (True or False), strings and complex numbers.

```
>>> a = 2
>>> b = 3.1415
>>> c = False
>>> d = "Hello Python"
```

```
>>> e = 3 + 6j  
>>>
```

### **Listing D.3** Different kinds of values

Many functions in Python are accessible via modules. For example to compute the sin of a number we can't simply type `sin (30)`. Instead we must first load the math module. We can then call the sin function:

```
>>> import math  
>>> print sin (3.1415)  
9.265358966049026e-05  
>>>
```

### **Listing D.4** Importing modules (libraries) into Python

In Tellurium we preload some libraries including the math library.

## **Repeating Calculations**

One of the commonest operations we do in computer programming is iteration. We can illustrate this with a simple example that loops ten times, each time printing out the loop index. This example will allow us to introduce the IDE editor. The editor is the panel on the left side of the IDE. In the editor we can type Python code, for example we could type:

```
a = 4.0  
b = 8.0  
c = a/b  
print "The answer is:", c
```

### **Listing D.5** Writing a simple program in the IDE editor

When we've finished typing this in the editor window, we can save our little program to a file (Select Menu: File/Save As...) and run the program by clicking on the green arrow in the tool bar of the IDE (Figure D.3). If we run this program we will see:

```
The answer is: 0.5  
>>>
```

### **Listing D.6** Writing a simple program in the IDE editor

The IDE allows a user to have as many program files open at once, each program file is given its own tab so that it is easy to move from one to the other. This is useful if one is working on multiple models at the same time.



**Figure D.3** Screen-shot of Tellurium, focusing on the Toolbar with the run button circled.

We will now use the editor to write the simple program that loops ten times, this is shown below:

```
for i in range (10):
    print i,
```

**Listing D.7** A simple loop in python

This will generate the sequence:

```
0 1 2 3 4 5 6 7 8 9
```

**Listing D.8** Result from simple loop program

There are a number of new concepts introduced in this small looping program. The first line contains the `for` keyword can be translated into literal English as “for all elements in a list, do this”. The list is generated from the `range()` function and in this case generates a list of 10 numbers starting at 0. `i` is the loop index and within the loop, `i` can be used in other calculations. In this case we will just print the value of `i` to the console. Each time the program loops it extracts the next value from the list and assigns it to `i`.

Two things are important to note in the print line. The first and most important is that the line has been indented four spaces. This isn’t just for aesthetic reasons but is actually functional. It tells Python what code should be executed *within* the loop. To elaborate we could add more lines to the loop, such as:

```
for i in range (10):
    a = i
    b = a*2
    print b,
print "Finished Loop"
```

**Listing D.9** A simple loop illustrating multiple statements

In this example there are three indented lines, this means that these three lines will be executed within the loop. The last line which prints a message, is not indented and therefore will not be executed within the loop. This means we only see the message appear once right at the end. The output for this little program is shown below.

```
0 2 4 6 8 10 12 14 16 18 Finished Loop
```

Another important point worth noting is the use of the , after the loop print statement. The comma is used to suppress a newline. This is why the output appears on one line only. If we had left out the comma each print statement would be on its own line.

A final word about `range()`. Range takes up to three arguments. In the example we only gave one argument, 10. A single argument means create a list starting at zero, incrementing one for each item until the incremented value reaches 10. A second argument such as `range(5, 10)` means start the list at 5 rather than zero. Finally, a third argument can be used to specify the increment size. For example the command `range(1, 10, 2)` will yield the list:

```
[1, 3, 5, 7, 9]
```

The easiest way to try out the various options in range is to type them at the console to get immediate feedback.

The use of variables, printing results, importing libraries and looping are probably the minimum concepts one needs to start using Python. However there are a huge range of resources online to help learn Python. Of particular interest is the codcademy web site (<http://www.codcademy.com/>). This site offers an interactive means to learn Python (including other programming languages).

### D.3 Describing Reaction Networks using Antimony

---

The code shown in the panel below illustrates the description of a very simple model using the Antimony syntax [125] followed by two lines of Python that uses libRoadRunner to run a simulation of the model. In this section we will briefly describe the Antimony syntax. A more detailed description of Antimony can be found at <http://antimony.sourceforge.net/index.html>.

```
import tellurium as te

rr = te.loada """
S1 -> S2; k1*S1;
S1 = 10; k1 = 0.1
"""

rr.simulate (0, 50, 100)
rr.plot()
```

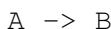
**Listing D.10** Simple model Antimony and simulated using libRoadRunner

The main purpose of Antimony is to make it straight forward to specify complex reaction networks using a familiar chemical reaction notation.

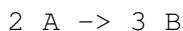
A chemical reaction can be an enzyme catalyzed reaction, a binding reaction, a phosphorylation, a gene expressing a protein or any chemical process that results in the conversion of one of more species (reactants) to a set of one or more other species (products). In Antimony, reactions are described using the notation:



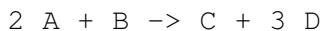
where the reactants are on the left side and products on the right side. The left and right are separated by the  $\rightarrow$  symbol. For example:



describes the conversion of reactant A into product B. In this case one molecule of A is converted to one molecule of B. The following example shows non-unity stoichiometry:



which means that two molecules of A react to form three molecules of B. Bimolecular and other combinations can be specified using the + symbol, that is:



tells us that two molecules of A combine with one molecule of B to form one molecule of C and three molecules of D.

To specify species that do not change in time (boundary species), add a dollar character in front of the name, for example:

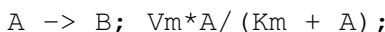


means that during a simulation A is fixed.

Reactions can be named using the syntax J1 :, for example:



means the reaction has a name, J1. Named reaction are useful if you want to refer to the flux of the reaction; kinetic rate laws come immediately after the reaction specification. If only the stoichiometry matrix is required, it is not necessary to enter a full kinetic law, a simple  $\dots \rightarrow S1; v;$  is sufficient. Here is an example of a reaction that is governed by a Michaelis-Menten rate law:



Note the semicolons. Here is a more complex example involving multiple reactions:

```
MainFeed: $X0 -> S1; Vm*X0/(Km + X0);
TopBranch: S1 -> $X1; Vm1*S1/(Km1 + S1);
BottomBranch: S1 -> $X2; Vm2*S1/(Km2 + S1);
```

There is no need to pre-declare the species names shown in the reactions or the parameters in the kinetic rate laws. Strictly speaking, declaring the names of the floating species is optional, however this feature is for more advanced users who wish to define the order

of rows that will appear in the stoichiometry matrix. For normal use there is no need to pre-declare the species names. To pre-declare parameters and variables see the example below:

```
const Xo, X1, X2; // Boundary species
var S1;           // Floating species

MainFeed:      $X0 -> S1; Vm*X0/(Km + X0);
TopBranch:     S1 -> $X1; Vm1*S1/(Km1 + S1);
BottomBranch: S1 -> $X2; Vm2*S1/(Km2 + S1);
```

We can load an Antimony model into libRoadRunner using the short-cut command `loada`. For example:

```
rr = te.loada '''
    const Xo, X1, X2; // Boundary species
    var S1;           // Floating species

    MainFeed:      $X0 -> S1; Vm*X0/(Km + X0);
    TopBranch:     S1 -> $X1; Vm1*S1/(Km1 + S1);
    BottomBranch: S1 -> $X2; Vm2*S1/(Km2 + S1);
'''
```

To reference model properties and methods, the property or method must be proceeded with the roadrunner variable. e.g. `rr.S1 = 2.3;`

When loaded into libRoadRunner the model will be converted into a set of differential equations. For example, consider the following model:

```
$Xo -> S1; v1;
S1 -> S2; v2;
S2 -> $X1; v3;
```

will be converted into:

$$\begin{aligned}\frac{ds_1}{dt} &= v_1 - v_2 \\ \frac{ds_2}{dt} &= v_2 - v_1\end{aligned}$$

Note that there are no differential equations for  $X_o$  and  $X_1$ . This is because they are fixed and do not change in time. If the reactions have non-unity stoichiometry, this is taken into account when the differential equations are derived.

### D.3.1 Initialization of Model Values

To initialize the concentrations and parameters in a model we can add assignments after the network is declared, for example:

```
MainFeed:      $X0 -> S1;  Vm*X0 / (Km + X0);
TopBranch:    S1 -> $X1; Vm1*S1 / (Km1 + S1);
BottomBranch: S1 -> $X2; Vm2*S1 / (Km2 + S1);

X0 = 3.4;   X1 = 0.0;
S1 = 0.1;
Vm = 12;   p.Km = 0.1;
Vm1 = 14;  p.Km1 = 0.4;
Vm2 = 16;  p.Km2 = 3.4;
```

## D.4 Using libRoadRunner in Python

libRoadRunner is a high performance simulator [127] that can simulate models described using SBML. In order to use Antimony with libRoadRunner it is necessary to first convert an Antimony description into SBML and then load the SBML into libRoadRunner. Tellurium provides a handy routine called `loadAntimonyModel` to help with this task (The short-cut name is `loada`). To load an Antimony model we first assign an Antimony description to a string variable, for example:

```
model = '''
S1 -> S2; k1*S1;

S1 = 10; k1 = 0.1;
'''
```

We now use the `loadAntimonyModel` (`model`) or `loada` to load the model into libRoadRunner.

```
>>> rr = te.loadAntimonyModel (model)
```

### Listing D.11 Loading an Antimony model

In this book we generally use the short-cut command as follows:

```
rr = te.loada ('''
S1 -> S2; k1*S1;

S1 = 10; k1 = 0.1;
''')
>>>
```

---

**Listing D.12** Loading an Antimony model using the short-cut command

Note that `loadAntimonyModel` and `loada` are part of the Tellurium Python package supplied with the Tellurium installer. If the Tellurium packages hasn't been loaded, use the following command to load the Tellurium package:

```
>>> import tellurium as te
```

**Listing D.13** Importing the Tellurium Package

#### D.4.1 Time Course Simulation

Once a model has been loaded into `libRoadRunner`, performing a simulation is very straight forward. To simulate a model we use the `libRoadRunner` `simulate` method. This method has many options but for everyday use four options will suffice. The following panel illustrates a number examples of how to use `simulate`.

```
>>> result = rr.simulate ()  
>>> result = rr.simulate (0, 10)  
>>> result = rr.simulate (0, 10, 100)  
>>> result = rr.simulate (0, 10, 100, ['time', 'S1'])
```

**Listing D.14** Calling the `simulate` method

---

Argument	Description
1st	Start Time
2nd	End Time
3rd	Number of Points
4th	Selection List

---

Let us focus on the forth version of the `simulate` method that takes four arguments. This call will run a time course simulation starting at time zero, ending at time 10 units, and generating 100 points. The results of the run are deposited in the matrix variable, `result`. At the end of the run, the `result` matrix will contain columns corresponding to the time column and all the species concentrations as specified by the forth argument. The forth argument can be used to change the columns that are returned from the `simulate` method. For example:

```
>>> result = rr.simulate (0, 10, 1000, ['S1'])}
```

will return a matrix 1,000 rows deep and one column wide that corresponds to the level of species S1.

Note that the special variable `Time` is available and represents the independent time variable in the model.

To visualize the output in the form of a graph, one can pass the matrix of results to the `plot` command. In the following example we return one species level, S1 and three fluxes. Finally we plot the results.

```
result = rr.simulate (0, 10, 1000, ['Time', 'S1', 'J1', 'J2', 'J3']);
te.plotWithLegend (rr, result)
```

or if we are not interested in the result data itself we can use the `libRoadRunner` plot:

```
rr.simulate (0, 10, 1000, ['Time', 'S1', 'J1', 'J2', 'J3']);
rr.plot()
```

It is possible to set the output column selections separately using the command:

```
rr.selections = ['time', 'S1']
```

This can save some typing each time a simulation needs to be carried out. By default the selection is set to time as the first column followed by all molecular species concentrations. As such it is more common to simply enter the command:

```
>>> result = rr.simulate (0, 10, 50)
```

In fact even the start time and end time and number of points are optional and if missing, `simulate` will revert to its defaults.

```
>>> result = rr.simulate()
```

## D.4.2 Plotting Simulation Results

Tellurium comes with Matplotlib which is a common plotting package used by many Python users. To simplify its use we provide two simple plotting calls:

```
te.plot (array)
te.plotWithLegend (rr, array)
```

The first takes the resulting array generated by a call to `simulate` and uses the first column as the *x* axis and all subsequent columns as *y* axis data. The second call takes the `roadrunner` variable as well as the array and does the same kind of plot but this time adds a legend to the

plot. We will use the first plotting command in the next section where we merge together multiple simulations.

### D.4.3 Applying Perturbations to a Simulation

Often in a simulation we may wish to perturb a species or parameter at some point during the simulation and observe what happens. One way to do this in Tellurium is to carry out two separate simulations where a perturbation is made in between the two simulations. For example, let's say we wish to perturb the species concentration for a simple two step pathway and watch the perturbation decay. First, we simulate the model for 10 time units; this gives us a transient and then a steady state.

```
import numpy # Required for vstack
import tellurium as te

rr = te.loada '''
$X0 -> S1; k1*X0;
S1 -> $X1; k2*S1;

X0 = 10; k1 = 0.3; k2 = 0.15;
''')

m1 = rr.simulate (0, 40, 50)
```

We then make a perturbation in  $S1$  as follows:

```
rr.S1 = rr.S1 * 1.6
```

which increases  $S1$  by 60%. We next carry out a second simulation:

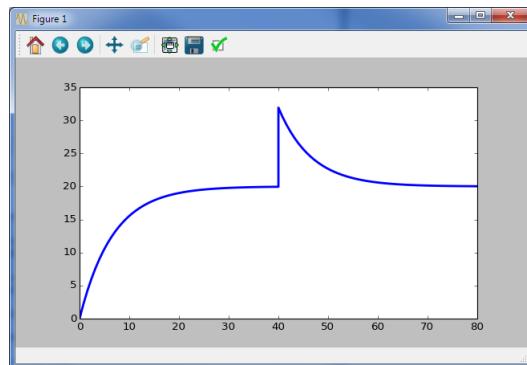
```
m2 = rr.simulate (40, 80, 50)
```

Note that we set the time start of the second simulation to the end time of the first simulation. Once we have the two simulations we can combine the matrices from both simulations using the Python command `vstack`

```
% Merge the two result array together
m = numpy.vstack ((m1, m2))
```

Finally, we plot the results, screen-shot shown in Figure D.4.

```
te.plotArray (m)
```



**Figure D.4** Screen-shot from Matplotlib showing effect of perturbation in S1.

#### D.4.4 Steady State and Metabolic Control

To evaluate the steady-state first make sure the model values have been previously initialized, then enter the following statement at the console.

```
>>> rr.getSteadyState()
```

This statement will attempt to compute the steady state and return a value indicating how effective the computation was. It returns the norm of the rate of change vector (i.e.  $\sqrt{\text{Sum of } dydt}$ ). The closer this is to zero, the better the approximation to the steady state. Anything less than  $10^{-4}$  usually indicates that a steady state has been found.

Once a steady state has been evaluated, the values of the metabolites will be at their steady state values, thus S1 will equal the steady state concentration of S1.

The fluxes through the individual reactions can be obtained by either referencing the name of the reaction (e.g. J1), or via the short-cut command rr.v. The advantage to looking at the reaction rate vector is that the individual reaction fluxes can be accessed by indexing the vector (see example below). **Note that indexing is from zero.**

```
>>> print rr.J1, rr.J2, rr.J3
3.4, ...
>>> for i in range (0, 2):
...     print rr.rv()[i]
3.4
etc
->
```

To compute control coefficients use the statement:

getCC (Dependent Measure, Independent parameter)

The dependent measure is an expression usually containing flux and metabolite references,

for example, `S1`, `J1`. The independent parameter must be a simple parameter such as a Vmax, Km, ki, boundary metabolite (`X0`), or a conservation total such as `cm_xxxx`. Examples include:

```
rr.getCC ('J1', 'Vmax1')
rr.getCC ('J1', 'Vm1') + rr.getCC ('J1', 'Vm2')
rr.getCC ('J1', 'X0')
rr.getCC ('J1', 'cm_xxxx')
```

To compute elasticity coefficients use the statement:

`getEE (Reaction Name, Parameter Name)`

For example:

```
rr.getEE ('J1', 'X0')
rr.getEE ('J1', 'S1')
```

Since `getCC` and `getEE` are built-in functions, they can be used alone or as part of larger expressions. Thus, it is easy to show that the response coefficient is the product of a control coefficient and the adjacent elasticity by using:

```
R = rr.getCC ('J1', 'X0')
print R - rr.getCC ('J1', 'Vm') * rr.getEE ('J1', 'X0')
```

To obtain the conservation matrix for a model use the model method, `getConservationMatrix`. Note that in the Antimony text we use the `var` word to predeclare the species so that we can set up the rows of the stoichiometry matrix in a certain order if we wish. This allows us to obtain conservation matrices with only positive terms.

```
import tellurium as te

r = te.loada '''
var ES, S1, S2, E;

J1: E + S1 -> ES; v;
J2: ES -> E + S2; v;
J3: S2 -> S1; v;
v = 0
'''

r.conervedMoietyAnalysis = True
print r.getConservationMatrix()

# Output
          ES, S1, S2, E
_CSUM0 [[ 1,  1,  1,  0],
```

```
_CSUM1 [ 1, -0, 0, 1]
```

The result given above indicates that the conservation relations,  $ES + S1 + E$  and  $E + ES$  exist in the model. As a result, Tellurium would generate two internal parameters of the form  $cm$  corresponding to the two relations.

#### D.4.5 Other Model Properties of Interest

There are a number of predefined objects associated with a reaction network model which might also be of interest. For example, the stoichiometry matrix,  $sm$ , the rate vector  $rv$ , the species levels vector and  $dv$  which returns the rates of change.

```
print rr.sm()
print rr.rv()
print rr.sv()
print rr.dv()
```

The names for the parameters and variables in a model can be obtained the short-cuts:

```
print rr.fs() # List of floating species names
print rr.bv() # List of boundary species names
print rr.ps() # List of parameter names
print rr.rs() # List of reaction names
print rr.vs() $ List of compartment names
```

The jacobian matrix can be returned using the command: `rr.getFullJacobian()`.

## D.5 Generating SBML and Matlab Files

Tellurium can import and export standard SBML [59] as well as export Matlab scripts for the current model. To load a model in SBML, load it directly into libRoadRunner. For example:

```
>>> rr = roadrunner.RoadRunner ('mymodel.xml')
>>> result = rr.simulate (0, 10, 100)
```

There are two ways to retrieve the SBML, one can either retrieve the original SBML loaded using `rr.getSBML()` or retrieve the *current* SBML using `rr.getCurrentSBML()`. Retrieving the current SBML can be useful if the model has been changed. To save the SBML to a file we can use the Tellurium helper function `saveToFile ()`, for example:

```
>>> te.saveToFile ('mySBMLModel.xml', rr.getCurrentSBML())
```

To convert an SBML file into Matlab, use the `getMatlab` method:

```
import tellurium as te

rr = te.loada '''
S1 -> S2; k1*S1;
S2 -> S3; k2*S2;
S1 = 10; k1 = 0.1; k2 = 0.2;
'''

# Save the SBML
te.saveToFile ('model.xml', rr.getSBML())

# Save the Matlab
te.saveToFile ('model.mat', rr.getMatlab())
```

## D.6 Exercise

Figure D.5 shows a two gene circuit with a feedforward loop. Assume the following rate laws for the four reactions:

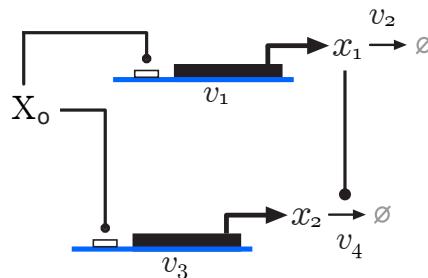
$$v_1 = k_1 X_o$$

$$v_2 = k_2 x_1$$

$$v_3 = k_3 X_o$$

$$v_4 = k_4 x_1 x_2$$

Assume that all rate constants are equal to one and that  $X_o = 1$ . Assume  $X_o$  is a fixed species.



**Figure D.5** Two gene circuit with feedforward loop.

1. Use Tellurium to model this system.
2. Run a simulation of the system from 0 to 10 time units.
3. Next, change the value of  $X_o$  to 2 (double it) and rerun the simulation for another 10 time units from where you left off in the last simulation. Combine both simulations and plot the result, that is time on the x-axis, and  $X_o$  and  $x_2$  on the y-axis.
4. What do you see?
5. Write out the differential equations for  $x_1$  and  $x_2$ .
6. Show algebraically that the steady state level of  $x_2$  is independent of  $X_o$ .



## References

- [1] Acerenza, L., and A. Cornish-Bowden. 1997. Biochem J **327** ( Pt 1):217–224.
- [2] Albe, K. R., M H Butler, and B E Wright. 1990. J Theor Biol **143** (2):163–195.
- [3] Aris, R. 1965. Arch. Rational Mech. Anal **19**:81–99.
- [4] Aström, K. J., and Richard M Murray. 2010. *Feedback systems: an introduction for scientists and engineers*. Princeton university press.
- [5] Bakker, B. M., P. A. Michels, F. R. Opperdoes, and H. V. Westerhoff. 1997. J Biol Chem **272** (6):3207–3215.
- [6] Berg, H. 2000. Physics Today **53** (1):24–30.
- [7] Black, H. 1977. IEEE spectrum **14** (12):55–61.
- [8] Blackman, F. F. 1905. Ann. Botany **19**:281–295.
- [9] Blinov, M. L., J R Faeder, B Goldstein, and W S Hlavacek. 2004. Bioinformatics **20** (17):3289–3291.
- [10] Blüthgen, N., F J Bruggeman, S Legewie, H Herzel, H V Westerhoff, and B N Kholodenko. 2006. FEBS J **273** (5):895–906.
- [11] Brown, G. C. 1991. Journal of theoretical biology **153** (2):195–203.
- [12] Bruggeman, F., JL Snoep, and HV Westerhoff. 2008. IET systems biology **2** (6):397–410.
- [13] Bruggeman, F. J., H. V. Westerhoff, J. B. Hoek, and B. N. Kholodenko. 2002. J. Theor Biol. **218**:507–20.
- [14] Buchler, N. E., and Frederick R Cross. 2009. Molecular systems biology **5** (1):272.
- [15] Buchler, N. E., and Matthieu Louis. 2008. J Mol Biol **384** (5):1106–1119.
- [16] Bujara, M., Michael Schümperli, René Pellaux, Matthias Heinemann, and Sven Panke. 2011. Nature chemical biology **7** (5):271–277.

- [17] Burrell, M. M., P. J. Mooney, M. Blundy, D. Carter, F. Wilson, J. Green, K. S. Blundy, and T. ap Rees. 1994. *Planta* **194**:95—101.
- [18] Burton, A. C. 1936. *Journal Cellular and Comparative Physiology* **9**(1):1–14.
- [19] Chan, T., and P.C. Hansen. 1992. *SIAM Journal on Scientific and Statistical Computing* **13**:727.
- [20] Chance, B., and G R Williams. 1955. *J Biol Chem* **217** (1):429–438.
- [21] Chassagnole, C., Rais B, E. Quentin, D. A. Fell, and J.-P. Mazat. 2001. *Biochem. J* **356**:415–423.
- [22] Clarke, B. L. 1980. *Stability of complex reaction networks.*, vol. 42 of *Adv. Chem. Phys.* Wiley, New York.
- [23] Cornish-Bowden, A. 2013. *FEBS letters* **587** (17):2725–2730.
- [24] Cornish-Bowden, A., and J. H. Hofmeyr. 2002. *J. theor. Biol* **216**:179–191.
- [25] Cornish-Bowden, A., J. H. Hofmeyr, and M. Cardenas. 2002. *Biochemical Society Transactions* **30**:43–47.
- [26] Curien, G., Stephane Ravanel, and Renaud Dumas. 2003. *The FEBS Journal* **270** (23):4615–4627.
- [27] Davies, S. E., and Kevin M Brindle. 1992. *Biochemistry* **31** (19):4729–4735.
- [28] De Camp, L. S. 1990. *The ancient engineers*. Barnes & Noble Publishing.
- [29] Del Vecchio, D., Aaron J Dy, and Yili Qian. 2016. *Journal of The Royal Society Interface* **13** (120):20160380.
- [30] Ehlde, M., and G. Zacchi. 1997. *Chemical Engineering Science* **52**:2599–2606(8).
- [31] Eisenthal, R., and A. Cornish-Bowden. 1998. *Journal of Biological Chemistry* **273** (10):5500–5505.
- [32] Emmerling, M., J E Bailey, and U Sauer. 2000. *Biotechnol Bioeng* **67** (5):623–627.
- [33] Eschrich, K., W. Scellenberger, and E. Hofmann. 1980. *Arch. Biochem. Biophys.* **205**:114–121.
- [34] Eschrich, K., W. Schellenberger, and E. Hofmann. 1983. *Arch. Biochem. Biophys.* **222**:657–660.
- [35] Eschrich, K., W. Schellenberger, and E. Hofmann. 1990. *European Journal Of Biochemistry* **188**:697–703.

- [36] Faeder, J. R., Michael L Blinov, and William S Hlavacek. 2009. Systems biology :113–167.
- [37] Fell, D. 1997. *Understanding the Control of Metabolism*. Portland Press., London.
- [38] Fell, D., and K. Snell. 1988. Biochem. J. **256**:97–101.
- [39] Fell, D. A., and H. M. Sauro. 1985. Eur. J. Biochem. **148**:555–561.
- [40] Fell, D. A., and H. M. Sauro. 1990. Eur. J. Biochem. **192**:183–187.
- [41] Flint, H. J., R. W. Tateson, I. B. Bartelmess, D. J. Porteous, W. D. Donochie, and H. Kacser. 1981. Biochem. J. **200**:231–246.
- [42] Gekas, V., and M. Lopez-Leiva. 1985. Process biochemistry **20** (1):2–12.
- [43] Giersch, C. 1995. The FEBS Journal **227** (1-2):194–201.
- [44] Gillespie, D. T. 2007. Annu Rev Phys Chem **58**:35–55.
- [45] Goldbeter, A., and D. E. Koshland. 1981. Proc. Natl. Acad. Sci **78**:6840–6844.
- [46] Goldbeter, A., and D. E. Koshland. 1984. J. Biol. Chem. **259**:14441–7.
- [47] Goodwin, B. 1965. Advances in Enzyme Regulation **3**:425–438.
- [48] Groen, A. K., R. J. A. Wanders, H. V. Westerhoff, R. van der Meer, and J. M. Tager. 1982. J. Biol. Chem. **257**:2754–2757.
- [49] Hanekom, A. J. 2006. Master's thesis University of Stellenbosch.
- [50] Hearon, J. Z. 1952. Physiol. Rev. **32**:499–523.
- [51] Heinisch, J. 1986. Mol. Gen Genet. **202**:75–82.
- [52] Heinrich, R., and T. A. Rapoport. 1974. Eur. J. Biochem. **42**:89–95.
- [53] Heinrich, R., and S Schuster. 1996. *The Regulation of Cellular Systems*. Chapman and Hall.
- [54] Higgins, J. 1963. Ann. N. Y. Acad. Sci. **108**:305–321.
- [55] Hofmeyr, J.-H. 2001. p. 291–300. In: Proceedings of the Second International Conference on Systems Biology . Caltech,
- [56] Hofmeyr, J.-H., and A. Cornish-Bowden. 1997. Comput Appl Biosci **13** (4):377–385.
- [57] Hofmeyr, J.-H., H. Kacser, and K. J. van der Merwe. 1986. Eur. J. Biochem. **155**:631–641.

- [58] Huang, C. F., and J. E. Ferrell. 1996. Proc. Natl. Acad. Sci **93**:10078–10083.
- [59] Hucka, M., A. Finney, H. M. Sauro, H. Bolouri, J. C. Doyle, H. Kitano, A. P. Arkin, B. J. Bornstein, D. Bray, A. Cornish-Bowden, A. A. Cuellar, S. Dronov, E. D. Gilles, M. Ginkel, V. Gor, I. I. Goryanin, W. J. Hedley, T. C. Hodgman, J. H. Hofmeyr, P. J. Hunter, N. S. Juty, J. L. Kasberger, A. Kremling, U. Kummer, N. Le NovÃÆ’Â“re, L. M. Loew, D. Lucio, P. Mendes, E. D. Mjolsness, Y. Nakayama, M. R. Nelson, P. F. Nielsen, T. Sakurada, J. C. Schaff, B. E. Shapiro, T. S. Shimizu, H. D. Spence, J. Stelling, K. Takahashi, M. Tomita, J. Wagner, and J. Wang. 2003. Bioinformatics **19**:524–531.
- [60] Ingalls, B. 2013. *Mathematical Modeling in Systems Biology: An Introduction*. MIT Press.
- [61] Ingalls, B. P. 2004. Journal of Physical Chemistry B **108**:1143–1152.
- [62] Kacser, H. 1983. Biochem. Soc. Trans. **11**:35–40.
- [63] Kacser, H., and J. A. Burns. 1973. In: Davies, D. D. , (ed.), Rate Control of Biological Processes, vol. 27 of *Symp. Soc. Exp. Biol.* p. 65–104. Cambridge University Press.
- [64] Kalisky, T., E Dekel, and U Alon. 2007. Physical Biology **4** (4):229–245.
- [65] Khammash, M. 2016. BMC biology **14** (1):1.
- [66] Kholodenko, B. N. 2006. Nat Rev Mol Cell Biol **7** (3):165–176.
- [67] Kholodenko, B. N., J. B. Hoek, H. V. Westerhoff, and G. C. Brown. 1997. FEBS Letters **414**:430–434.
- [68] Kim, K. H., and Herbert M Sauro. 2015. In: Computational Methods in Synthetic Biology p. 287–297. Springer.
- [69] King, E., and C. Altman. 1956. The Journal of Physical Chemistry **60** (10):1375–1378.
- [70] Klipp, E., and Reinhart Heinrich. 1999. Biosystems **54** (1-2):1–14.
- [71] Kolch, W., Melinda Halasz, Marina Granovskaya, and Boris N Kholodenko. 2015. Nature reviews. Cancer **15** (9):515.
- [72] Kornberg, A., et al.. 1960. Science **131** (3412):1503–1508.
- [73] Koshland, D. E., G. Némethy, and D. Filmer. 1966. Biochemistry **5** (1):365–385.
- [74] Kouril, T., Dominik Esser, Julia Kort, Hans V Westerhoff, Bettina Siebers, and Jacky L Snoep. 2013. FEBS Journal **280** (18):4666–4680.
- [75] LaPorte, D. C., K Walsh, and D E Koshland. 1984. J Biol Chem **259** (22):14068–14075.
- [76] Liebermeister, W., and E. Klipp. 2006. Theor Biol Med Model **3**:42.

- [77] Liebermeister, W., J. Uhendorf, and E. Klipp. 2010. Bioinformatics **26** (12):1528.
- [78] Liu, A. P., and Daniel A Fletcher. 2009. Nature reviews. Molecular cell biology **10** (9):644.
- [79] Longabaugh, W., E.H. Davidson, and H. Bolouri. 2005. Developmental Biology **283** (1):1–16.
- [80] Marín-Hernández, A., Sara Rodríguez-Enríquez, Paola A Vital-González, Fanny L Flores-Rodríguez, Marina Macías-Silva, Marcela Sosa-Garrocho, and Rafael Moreno-Sánchez. 2006. FEBS Journal **273** (9):1975–1988.
- [81] Markevich, N. I., J B Hoek, and B. N. Kholodenko. 2004. J. Cell Biol. **164**:353–9.
- [82] Maus, C., Stefan Rybacki, and Adelinde M Uhrmacher. 2011. BMC Systems Biology **5** (1):166.
- [83] Mayr, O. 1975. *The Origins of Feedback Control*. The MIT Press.
- [84] Mindell, D. 2000. Technology and Culture **41** (3):405–434.
- [85] Monod, J., J Wyman, and J. P. Changeux. 1965. J Mol Biol **12**:88–118.
- [86] Mor, I., EC Cheung, and KH Vousden. 2011. p. 211–216. In: Cold Spring Harbor symposia on quantitative biology vol. 76. Cold Spring Harbor Laboratory Press.,
- [87] Morales, M. F. 1921. Journal Cellular and Comparative Physiology **30**:303–313.
- [88] Moreno-Sánchez, R., Alvaro Marín-Hernández, Emma Saavedra, Juan P Pardo, Stephen J Ralph, and Sara Rodríguez-Enríquez. 2014. The international journal of biochemistry & cell biology **50**:10–23.
- [89] Müller, S., Friedrich K Zimmermann, and Eckhard Boles. 1997. Microbiology **143** (9):3055–3061.
- [90] Müller-Hill, B. 1996. *The lac Operon: A Short History of a Genetic Paradigm*. Walter de Gruyter, Berlin.
- [91] Niekerk, D. D., Gerald P Penkler, Francois Toit, and Jacky L Snoep. 2016. The FEBS journal **283** (4):634–646.
- [92] Olivier, B. G., J. M. Rohwer, and J. H. Hofmeyr. 2005. Bioinformatics **21**:560–1.
- [93] Ortega, F., J L Garcés, F Mas, B N Kholodenko, and M Cascante. 2006. FEBS J **273** (17):3915–3926.
- [94] Palsson, B. O. 2007. *Systems Biology: Properties of Reconstructed Networks*. Cambridge University Press.

- [95] Papadopoulos, E. 2007. Distinguished figures in mechanism and machine science :217–245.
- [96] Park, D. 1988. Computers & chemistry **12** (2):175–188.
- [97] Penkler, G., Francois du Toit, Waldo Adams, Marina Rautenbach, Daniel C Palm, David D van Niekerk, and Jacky L Snoep. 2015. FEBS Journal **282** (8):1481–1511.
- [98] Petersen, S., A A de Graaf, L Eggeling, M Möllney, W Wiechert, and H Sahm. 2000. J Biol Chem **275** (46):35932–35941.
- [99] Qiao, L., R.B. Nachbar, I.G. Kevrekidis, S.Y. Shvartsman, and A. Asthagiri. 2007. PLoS Comput Biol **3** (9):e184.
- [100] Rao, C. V., H. M. Sauro, and A. P. Arkin. 2004. 7th International Symposium on Dynamics and Control of Process Systems, DYCOPS, 7 7:0–0.
- [101] Rapoport, T. A., and R. Heinrich. 1975. Biosystems **7**:120–129.
- [102] Reder, C. 1988. J. Theor. Biol. **135**:175–201.
- [103] Reich, J. G., and E. E. Selkov. 1981. *Energy metabolism of the cell*. Academic Press, London.
- [104] Ro, D. K., E M Paradise, M Ouellet, K J Fisher, K L Newman, J M Ndungu, K A Ho, R A Eachus, T S Ham, J Kirby, M C Chang, S T Withers, Y Shiba, R Sarpong, and J D Keasling. 2006. Nature **440** (7086):940–943.
- [105] Rohwer, J. M., A. J. Hanekom, and J. H. Hofmeyr. 2007. p. 175–187. In: Proc. 2nd Int. Symp. on Experimental Standard Conditions of Enzyme Characterizations (ESEC 2006) . Beilstein Institute, Frankfurt am Main, Germany,
- [106] Ruijter, G., H Panneman, and J Visser. 1997. Biochimica et Biophysica Acta (BBA)- General Subjects **1334** (2):317–326.
- [107] Sauro, H. M. 1994. BioSystems **33**:15–28.
- [108] Sauro, H. M. 2000. In: Hofmeyr, J.-H. S., J. M. Rohwer, and J. L. Snoep. , (ed.), *Animating the Cellular Map: Proceedings of the 9th International Meeting on BioThermoKinetics*, . Stellenbosch University Press.
- [109] Sauro, H. M. 2012. *Enzyme Kinetics for Systems Biology*. Ambrosius Publishing. 2nd Edition.
- [110] Sauro, H. M. 2015. *Systems Biology: Linear Algebra for Pathway Modeling*. Ambrosius Publishing, Seattle.
- [111] Sauro, H. M., and D. A. Fell. 1991. Mathl. Comput. Modelling **15**:15–28.

- [112] Sauro, H. M., and Brian Ingalls. 2007. arXiv preprint arXiv:0710.5195 .
- [113] Savageau, M. A. 1972. Curr. Topics Cell. Reg. **6**:63–130.
- [114] Savageau, M. A. 1975. J Mol Evol **5** (3):199–222.
- [115] Savageau, M. A. 1976. *Biochemical systems analysis: a study of function and design in molecular biology*. Addison-Wesley, Reading, Mass.
- [116] Schaaff, I., J. Heinisch, and F. K. Zimmermann. 1989. Yeast **5** (4):285–290.
- [117] Schellenberger, W., K. Eschrich, and E. Hofmann. 1978. Acta biol. med. germ. **37**:1425–1441.
- [118] Schellenberger, W., Klaus Eschrich, and Eberhard Hofmann. 1981. Advances in enzyme regulation **19**:257–284.
- [119] Schuster, S., T Dandekar, and D A Fell. 1999. Trends Biotechnol **17** (2):53–60.
- [120] Schuster, S., D. A. Fell, and T. Dandekar. 2000. Nature Biotechnlogy **18**:326–332.
- [121] Schuster, S., and T. Hofer. 1991. J. Chem. Soc. Faraday Trans. **87**:2561–2566.
- [122] Segel, I. H. 1993. *Enzyme kinetics: behavior and analysis of rapid equilibrium and steady-state enzyme systems (Paper)*. Wiley-Interscience.
- [123] Small, J. R., and D. A. Fell. 1990. Eur. J. Biochem. **191**:405–411.
- [124] Smallbone, K., Hanan L Messiaha, Kathleen M Carroll, Catherine L Winder, Naglis Malys, Warwick B Dunn, Ettore Murabito, Neil Swainston, Joseph O Dada, Farid Khan, et al.. 2013. FEBS letters **587** (17):2832–2841.
- [125] Smith, L. P., Frank T Bergmann, Deepak Chandran, and Herbert M Sauro. 2009. Bioinformatics **25** (18):2452–2454.
- [126] Snell, K., and D. A. Fell. 1990. Adv. Enzyme Reg. **30**:13–32.
- [127] Somogyi, E. T., Jean-Marie Bouteiller, James A. Glazier, Matthias König, J. Kyle Medley, Maciej H. Swat, and Herbert M. Sauro. 2015. Bioinformatics (Oxford, England) .
- [128] Stephanopoulos, G. N., and A. A. Aristidou. 1998. *Metabolic Engineering: Principles and Methodologies*. Academic Press.
- [129] Stitt, M., W. P. Quick, U. Schurr, E.-D. Schulze, S. R. Rodermel, and L. Bogorad. 1991. Planta **183**:555–566. 10.1007/BF00194277.
- [130] Storey, K. B. 2005. *Functional metabolism: regulation and adaptation*. John Wiley & Sons.

- [131] Sturm, O. E., Richard J Orton, Joan Grindlay, *et al.*. 2010. *Sci Signal* **3**(153).
- [132] Szent-Györgyi, A. 1963. *Annual review of biochemistry* **32** (1):1–15.
- [133] Szurmant, H., and G.W. Ordal. 2004. *Microbiology and Molecular Biology Reviews* **68** (2):301.
- [134] Thomas, S., Peter JF Mooney, Michael M Burrell, *et al.*. 1997. *Biochemical Journal* **322** (1):119–127.
- [135] Thron, C. 1991. *Bulletin of mathematical biology* **53** (3):403–424.
- [136] Torres, N., F Mateo, E Melendez-Hevia, and H Kacser. 1986. *Biochem. J* **234**:169–174.
- [137] Trinh, C., Aaron Wlaschin, and Friedrich Srienc. 2009. *Applied Microbiology and Biotechnology* **81**:813–826. 10.1007/s00253-008-1770-1.
- [138] Tyson, J., and H. G. Othmer. 1978. *Progress in Theoretical Biology* (R. Rosen & F.M. Snell, Eds.) **5**:1–62.
- [139] Tyson, J. J., K. C. Chen, and B. Novak. 2003. *Current Opinion in Cell Biology* **15**:221–231.
- [140] Urbano, A. M., Helen Gillham, Yoram Groner, and Kevin M Brindle. 2000. *Biochemical Journal* **352** (3):921–927.
- [141] Vallabhajosyula, R. R., V Chickarmane, and H M Sauro. 2006. *Bioinformatics* **22** (3):346–353.
- [142] van Eunen, K., and Barbara M Bakker. 2014. *Perspectives in Science* **1** (1):126–130.
- [143] Varma, A., and B. O. Palsson. 1993a. *Journal of Theoretical Biology* **165**(4):477–502.
- [144] Varma, A., and B. O. Palsson. 1993b. *Journal of Theoretical Biology* **165**(4):503–522.
- [145] Varusai, T. M., Walter Kolch, Boris N Kholodenko, and Lan K Nguyen. 2015. *Molecular BioSystems* **11** (10):2750–2762.
- [146] Waley, S. G. 1964. *Biochem. J.* **91** (3):514–0.
- [147] Walsh, K., and Daniel E Koshland. 1985. *Proceedings of the National Academy of Sciences* **82** (11):3577–3581.
- [148] Wang, H., Marie La Russa, and Lei S Qi. 2016. *Annual review of biochemistry* **85**:227–264.
- [149] Webb, J. 1963. *Enzyme and Metabolic Inhibitors: General principles of inhibition. Volume 1*. Number v. 1. Academic Press.

- [150] Westermark, P. O., J. Hellgren-Kotaleski, and A. Lansner. 2004. WSEAS Transactions on Biology and Medicine **1**:91–98.
- [151] Wilkinson, D. J. 2006. *Stochastic Modelling for Systems Biology*. Chapman and Hall.
- [152] Woods, J. H., and H. M. Sauro. 1997. Comput Appl Biosci **13** (2):123–130.
- [153] Wright, S. 1934. The American Naturalist **68**:24–53.



## ***History***

### 1. VERSION: 1.0

**Date:** 2018-04-31

**Author(s):** Herbert M. Sauro

**Title:** Introduction to Metabolic Control Analysis

**Modification(s):** First Edition Release



# Index

## Symbols

<b>N</b> .....	53
$s_d$ .....	293
$s_i$ .....	293
$\mathcal{E}$ .....	100
$\partial$ .....	100
$\varepsilon$ .....	100

## A

activator.....	42
allosteric constant.....	34
AMP.....	62
amplification.....	215
Antimony .....	44
antimony.....	411, 416
Antimony, Tellurium .....	53
apparent kinetic order.....	101
arginine biosynthesis.....	174
asymmetry: thermodynamic .....	200
ATP.....	62

## B

Bakker .....	325
Bell Labs .....	209
bifurcation plot.....	258, 265, 343
biochemical systems theory .....	101
biomodels .....	343
Biotapestry.....	46
bistability .....	341
bistable.....	259, 260
Blackman.....	143
branch point theorem.....	231
branch point ultrasensitivity .....	233
branch point: algebraic .....	235
branched model: cyclic .....	374
branched pathways .....	229
Briggs-Haldane .....	111

## C

canonical control coefficients .....	160
cas9 .....	185
cascades .....	345
cascades: negative feedback.....	349
chock point.....	143
citrate synthase.....	175
close to equilibrium .....	192
closed loop gain .....	212
closed solution .....	188
coenzyme A .....	63
committed step.....	190
competitive .....	30
competitive inhibition .....	30
complex eigenvalues .....	255
complex number .....	257
computational approaches .....	296
computer exchange format.....	10
computer simulation.....	183
concentration connectivity theorem .....	154
concentration control coefficient.....	127
concentration control coefficients.....	196
concentration control: distribution .....	201
conjugate pair .....	257
connectivity theorems .....	151
connectivity theorems: interpretation ...	156
conservation matrix .....	309
conserved cycles .....	60, 288
conserved laws: python.....	304
constraining species levels .....	324
control .....	21, 123
control coefficients .....	124, 127
control equations.....	161, 405
control equations: algebraic method ....	163
Cornish-Bowden .....	38, 325
covalent modification cycles .....	345

- CRISPR ..... 185  
cubic equation ..... 279  
cycle amplification ..... 330  
cycle connectivity theorems ..... 334  
cyclic branched model ..... 374
- D**
- de Movire's theorem ..... 400  
definition of pathway ..... 376  
Denton and Pogson ..... 143  
dependent species ..... 293  
derivative ..... 100  
derive control equations ..... 161  
design of simulation software ..... 318  
determined systems ..... 358  
determined systems: algebra ..... 361  
difference quotient ..... 104  
differential equations ..... 394  
dimerization ..... 339  
discrete events ..... 45  
disequilibrium ratio ..... 19, 109, 110  
distribution of control ..... 128  
distribution: concentration control ..... 201  
disturbance ..... 244  
dominant genes ..... 139  
double modulation ..... 179  
*Drosophila melanogaster* ..... 174  
dual cycles ..... 336
- E**
- echelon form ..... 299  
eigenvalues ..... 246, 252, 277, 399  
eigenvectors ..... 399  
Eisenthal ..... 325  
elasticity ..... 265  
elasticity coefficient ..... 100  
elasticity matrix ..... 249  
elasticity rules ..... 116  
elasticity: cooperativity ..... 113  
elasticity: enzyme ..... 112  
elasticity: enzyme catalyzed ..... 111  
elasticity: Hill equation ..... 113  
elasticity: log/log plot ..... 108
- elasticity: mass-action kinetics ..... 109  
elasticity: operational definition ..... 102  
elementary matrices ..... 297  
elementary modes ..... 371  
elementary modes: conditions ..... 372  
elementary operations ..... 296  
elementary reaction ..... 39  
elementary reactions ..... 39  
endergonic ..... 62  
enzyme action ..... 27  
enzyme elasticity ..... 112  
enzyme inhibitor complex ..... 30  
enzyme-reactant complex ..... 27  
error signal ..... 210  
experimental methods ..... 173  
explicit regulation ..... 58  
exponential terms ..... 254  
exponentials ..... 246  
export Matlab ..... 425  
export SBML ..... 425  
extensive property ..... 13  
extent of reaction ..... 13  
external stability ..... 251
- F**
- fast and slow scales ..... 289  
feedback amplifier ..... 209  
feedback control: Greek ..... 207  
feedback: stability ..... 275  
first-order ..... 101  
fixed species ..... 45  
flux balance laws ..... 355  
flux connectivity theorem ..... 152  
flux constraint ..... 137  
flux constraints ..... 355  
flux control coefficient ..... 127  
flux expression ..... 188  
flux summation theorem ..... 135  
forward elimination ..... 296  
fractional amounts ..... 11  
free energy ..... 20  
front loading ..... 190, 198  
futile cycle amplification ..... 237

- futile cycles ..... 236
- G**
- gain ..... 212
- Gaussian elimination ..... 296
- gene networks ..... 58
- general solution ..... 246
- generalized power law ..... 168
- generic feedback ..... 210
- genetic engineering ..... 174
- genetics ..... 174
- getCC ..... 141, 423
- getEE ..... 424
- glucose-6-phosphate ..... 324
- glyceraldehyde-3-phosphate ..... 324
- glycerol 3-phosphate ..... 324
- glycolysis ..... 324
- glyosome ..... 324
- governors ..... 208
- Greek science ..... 207
- H**
- Harold Black ..... 209
- heterozygotes ..... 139
- Hill equation ..... 113, 328
- I**
- impedance ..... 215
- implicit regulation ..... 58
- independent species ..... 293
- inhibitor ..... 42
- inhibitor titrations ..... 176
- initial rate ..... 29
- intensive property ..... 14
- internal stability ..... 244
- interpretation: connectivity theorems ... 156
- interpreting summation theorems ..... 139
- IPTG ..... 175
- irreversible bistability ..... 267
- irreversible inhibitors ..... 179
- irreversible step ..... 191
- J**
- Jacobian ..... 276
- Jacobian matrix ..... 246, 250, 320
- Jacobian: biochemical system ..... 249
- Jarnac ..... 44
- K**
- kernel: Scilab ..... 307
- key step ..... 143
- kinetic mechanism ..... 27
- kinetic order ..... 100, 168
- Ktesibios ..... 207
- L**
- large number of molecules ..... 49
- left null space ..... 310
- Lehninger ..... 143
- libRoadRunner ..... 411
- linear dependence ..... 293
- linear differential equation ..... 253
- linear pathway ..... 50, 125
- linear pathways ..... 187
- linear programming ..... 366
- linearization ..... 214
- linearly dependent rows ..... 293
- link matrix ..... 308, 310
- link-zero matrix ..... 308
- linking parts to the whole ..... 149
- local changes ..... 135
- local equation ..... 114, 116
- log form ..... 106
- logarithmic gains ..... 168
- logarithmic scales ..... 107
- logarithms ..... 392
- loop gain ..... 212
- M**
- MAPK ..... 325
- Markevich switch ..... 341
- mass-action kinetics ..... 109
- mass-action ratio ..... 19
- mass-balance equation ..... 46
- Mathematica ..... 119
- Mathematica: null space ..... 358
- Matlab ..... 52, 425

- Matlab: null space ..... 358  
matrix method ..... 161  
maximum yield predictions ..... 377  
MCA of conserved cycles ..... 325  
MCA, stability ..... 276  
MCA, ultrasensitivity ..... 330  
MCA: Tellurium ..... 423  
measuring control coefficients ..... 173  
mechanistic details ..... 43  
metabolic flux analysis ..... 366  
Metatool ..... 372  
Michaelis-Menten ..... 111  
mitogen-activated protein kinase ..... 325  
model composition ..... 45  
modulating cycle totals ..... 348  
modulating input signal ..... 346  
moiety ..... 61  
moiety conservation ..... 288  
Moiety Conserved Cycles ..... 323  
moiety conserved cycles ..... 60  
moiety constraints ..... 287  
Monod, Wyman and Changeux ..... 34  
multicompartment Systems ..... 319  
multiple steady states ..... 259
- N**
- NAD/NADH ..... 63  
negative cooperativity ..... 33  
negative elasticity ..... 102  
negative feedback amplifier ..... 351  
negative feedback: biochemical ..... 215  
negative feedback: control analysis ..... 217  
negative feedback: graphical ..... 216  
negative feedback: historical ..... 207  
negative feedback: simple analysis ..... 211  
network topology ..... 250  
*Neurospora crassa* ..... 174  
Newton's difference quotient ..... 103  
non-elementary reactions ..... 42  
non-unity stoichiometry ..... 41  
notation ..... 124  
null space ..... 358, 367  
null space of  $\mathbf{N}^T$  ..... 305
- numerical methods ..... 311  
numerical stability ..... 319
- O**
- one-step pathway ..... 134  
op amp ..... 209  
open loop gain ..... 212  
operating point ..... 104  
operational definition ..... 102  
optimal allocation of protein ..... 203
- P**
- pacemaker step ..... 143  
Palmer ..... 38  
Palsson ..... 65  
parameterless control coefficients ..... 160  
partial derivative ..... 100  
partial derivatives ..... 392  
pathway engineering ..... 382  
pathway length: stability ..... 282  
pathwayDesigner ..... 53  
percentage changes ..... 102, 128  
periodic behavior ..... 254  
periodic solutions ..... 256  
perturbations ..... 126  
perturbations transmission ..... 190  
perturbations: Tellurium ..... 422  
perturbations: transmission ..... 200  
PFK paradox ..... 220  
phase plot ..... 252  
phase portrait ..... 252  
phase shift ..... 276  
Philon ..... 207  
phosphate ..... 63  
phosphorylation cycle ..... 287  
plotting ..... 421  
polynomials ..... 395  
positive cooperativity ..... 33  
positive elasticity ..... 102  
positive feedback loop ..... 259  
positive feedback: stability ..... 264  
predicting concentration changes ..... 131  
predicting flux changes ..... 131

- product inhibition ..... 31, 191  
 product insensitive ..... 195  
 propagation ..... 111  
 proportional change ..... 107  
 proportional changes ..... 107  
 pure competitive inhibitor ..... 30  
 PySCeS ..... 44  
 python ..... 411
- Q**
- QR Factorization ..... 313  
 quasi-equilibrium ..... 71
- R**
- R state ..... 34  
 rank ..... 293  
 rapid-equilibrium ..... 43  
 rate of change ..... 8  
 rate of reaction ..... 13  
 rate-limiting steps ..... 143  
 rational fractions ..... 11  
 reaction order ..... 100  
 reaction rates ..... 54  
 recessive genes ..... 139  
 reconstituted systems ..... 181  
 reduced echelon ..... 359  
 reduced echelon form ..... 299  
 regulation ..... 224  
 regulator ..... 210  
 regulatory link ..... 43  
 relative changes ..... 102  
 relaxation times ..... 192  
 relaxed ..... 34  
 response coefficient ..... 158  
 response coefficient: non-MCA ..... 328  
 reversibility ..... 40, 55  
 robustness ..... 215, 226  
 roots, polynomial ..... 395  
 Routh-Hurwitz ..... 396  
 rref: Matlab;/Scilab ..... 304  
 rref: python ..... 303  
 RRQR ..... 316  
 rubisco ..... 175
- S**
- S-Systems: relationship ..... 166  
 saddle node ..... 255  
 saturable kinetics ..... 193  
 Sauro ..... 38  
 SBGN ..... 46  
 SBML ..... 11, 52, 411, 425  
 scaled  $L$  ..... 310  
 second-order ..... 101  
 Segel ..... 38  
 sensitivity ..... 213  
 sensitivity to  $T$  ..... 333  
 sequestration ..... 339  
 serine pathway ..... 184  
 servomechanism ..... 210  
 set point ..... 210  
 short table of derivatives ..... 392  
 sigmoid response ..... 33  
 signal transmission ..... 209  
 signaling networks ..... 57  
 simple branch ..... 137, 357  
 singular value decomposition ..... 311  
 software: bifurcation ..... 266  
 spiral trajectories ..... 256  
 Spyder ..... 411  
 stability ..... 227, 243  
 stability threshold ..... 280  
 stability, MCA ..... 276  
 stability: four-step pathway ..... 278  
 stability: pathway length ..... 282  
 stability: three-step pathway ..... 276  
 stable ..... 244  
 stable node ..... 254  
 stable point ..... 244  
 steady state ..... 74  
 steady state: Tellurium ..... 423  
 steam engine ..... 208  
 stoichiometric amount ..... 8  
 stoichiometric coefficient ..... 9, 50  
 stoichiometric network ..... 39  
 stoichiometry matrix ..... 53  
 stoichiometry matrix partitioned ..... 307

- subelasticity terms ..... 118  
 substrate cycles ..... 236  
 sum and product ..... 391  
 sum of the elasticities ..... 110  
 summation theorems ..... 134, 139  
 supply/demand ..... 226  
 SVD ..... 311  
 svd: Scilab ..... 312  
 sympy ..... 303  
 system equation ..... 63
- T**
- T state ..... 34  
 Taylor series ..... 397  
 Tellurium ..... 64, 310  
 tense ..... 34  
 text representation ..... 44  
 thermodynamic equilibrium ..... 74  
 three-point estimation ..... 103, 104  
 three-point formula ..... 103  
 three-step pathway ..... 194  
 time scale ..... 60  
 toggle switch ..... 344  
 topology ..... 250  
 total derivative ..... 115, 132, 399  
 tracking ..... 210  
 tracking device ..... 351  
 traffic congestion ..... 144  
 transcription factor ..... 58  
*Trypanosoma brucei* ..... 324  
 two-step pathway ..... 128, 135  
 type I elementary matrices ..... 297  
 type II elementary matrices ..... 297  
 type III elementary matrices ..... 297  
 Tyson ..... 267
- U**
- ultrasensitivity ..... 328  
 ultrasensitivity via sequestration ..... 339  
 ultrasensitivity: implicit differentiation ..... 332  
 underdetermined systems ..... 366  
 unscaled elasticity ..... 100, 249, 277  
 unstable node ..... 254
- unstable point ..... 244
- V**
- vectors and matrices ..... 391  
 Voet and Voet ..... 143  
 volume ..... 14
- W**
- water clocks ..... 207  
 Watt ..... 208  
 well-stirred reactor ..... 47
- Z**
- zero-order ..... 101