Kinetics, Control and Regulation of Metabolic Systems

Jan-Hendrik S. Hofmeyr¹ Jacky L. Snoep¹,²
Hans V. Westerhoff²

November 29, 2002

¹ Triple-J Group for Molecular Cell Physiology, Dept. of Biochemistry, University of Stellenbosch, Private Bag X1, Matieland 7602, South Africa
² Molecular Cell Physiology, Vrije Universiteit, de Boelelaan 1087, 1081 HV Amsterdam, The Netherlands

© 2002, J.-H.S. Hofmeyr (Chapters 1–4)

Programme

08:30 The kinetic model
- The ordinary differential equations
- System states: transient, steady and equilibrium
- Steady-state flux relationships: simple examples
- Conservation relationships: simple examples

09:30 Exercises
10:30 Coffee
11:00 Structural analysis
- The stoichiometric matrix, N
- Relationships between ODEs (the L-matrix) and conservation relationships
- Steady-state flux relationships (the K-matrix)

11:30 Computational structural analysis
12:00 Lunch
13:00 Metabolic Control Analysis
- Definitions of the coefficients (R, C, ε): operational, graphical, and mathematical
- Understanding the relationships between the coefficients using thought experiments
  - The summation property
  - The partitioned response property
  - The connectivity property
- Combining structure and control: the general control matrix equation (C·E=I)
- Expressing control in terms of elasticity

15:00 Break
15:30 Illustration of metabolic engineering strategy using MCA, and kinetic modelling on the JWS server
16:30 End of tutorial
Introduction

Systems biology is one of the most exciting fields of research today. With the sequencing of complete genomes and the rapid development of whole systems approaches in the proteomics and metabolomics fields, systems biologists can now attempt to build realistic models of whole cells, including signal transduction and gene regulation. The possibilities for research with such models are immense, ranging from fundamental questions of cell function and regulation, to more philosophical ones about studying complex systems and aspects of life itself, to applied projects such as drug development. The field of systems biology falls within the context of Bioinformatics where its role lies in the integration of the data sets obtained within the Genomics, Transcriptomics, Proteomics and Metabolomics fields. Of course kinetic modelling and systems biology have a long history starting in the 1960's, well before the recent bioinformatics development. The traditional term for these approaches is computational biology. In the 1960's Chance, Garfinkel and Hess started with kinetic modelling of metabolic pathways. Since then both the immense increase in computer power and the extended knowledge on enzyme kinetics have made it possible to make more realistic models than run much faster. However, most of the models are limited to the metabolic level, due to the dearth of precise kinetic data on other cellular levels. The recent boost in Bioinformatics has led to a huge increase in availability of data that can be used to validate models. This is very important in the building of larger realistic models, which can be used for several purposes: as tools to control processes or to predict effects of certain perturbations, or, and this will be our main interest, for the quantitative understanding of systems. Detailed kinetic models tend to become as complex as the system that is being modelled and with such models it is important to use a higher level language to understand the model. The framework we use for this is metabolic control analysis. Originally developed in the 1970s by Kacser and Burns and Heinrich and Rapoport, this framework makes it possible to quantify to which degree each of the enzymes determines the system variables and how this relates to enzyme properties.

Aim and motivation

Metabolic control analysis has become the quantitative framework of choice for many systems biologists interested in control and regulation in reaction networks. It is applicable to many levels of description: from elementary steps that form a reaction mechanism, through enzyme-catalysed steps that form metabolic pathways, to organs that form an organism, to organisms that form a population, to ecosystems. As such, it is an indispensable tool of systems biology.

The aim of this tutorial is to teach the basic concepts of metabolic control analysis starting from basic principles. In the first half of the tutorial a general kinetic model for any system of coupled chemical reactions is used as a point of departure for understanding the different dynamic states of such systems. The network topology can be captured in the form of a stoichiometric matrix which can be analysed to yield both the relationships between the differential equations that describe the system (and, therefore, the mass conservation constraints) and the steady-state flux relationships. The second half of the tutorial is devoted to control analysis, where we use descriptive, graphical, and mathematical thought experiments to explain both the definitions of and the relationships between the coefficients of control analysis. Structural analysis of the stoichiometric matrix and control analysis is then combined into a simple and elegant description of the relationship between the systemic control properties of the steps in the network and the local elasticity properties of these steps. Lectures are augmented with exercise sessions.

Recommended background

Basic knowledge of the following is assumed:

Mathematics
- Calculus (differentiation)
- Linear algebra (elementary matrix manipulation)

Biology/Chemistry
- Reaction kinetics (enzyme rate equations)
- Equilibrium thermodynamics (the equilibrium constant and mass-action ratio)

Expected outcomes

After completing the tutorial participants should be able to
- write a kinetic model for any reaction network;
- obtain the steady state flux relationships (K-matrix) and the conservation relationships (from the L-matrix), through analysis of the stoichiometric matrix (N) either by hand (for simple systems) or by using a computer package such as PySCeS, Gepasi or Jarnac;
- understand the definitions of the main coefficients of control analysis (C, R, ϵ);
- use thought experiments to explain the relationships between the coefficients in control analysis (partitioned response, summation and connectivity properties).
combine the control properties in the general control matrix equation
solve the control matrix equation to express control (systemic) properties in
terms of elasticity (local) properties
apply the concepts of control analysis to real systems

Contributors

Jacky Snoep
received his PhD in 1992 on the control of pyruvate catabolism in
microorganisms from the University of Amsterdam. He subsequently worked as a
post doctoral fellow, first specializing in molecular techniques at the University ofFlorida and later together with ... of kinetic models that can be run viathe internet (http://jjj.biochem.sun.ac.za or http://www.jjj.bio.vu.nl). These mod-
els will be part of a bigger silicon cell effort (http://www.siliconcell.net/). Snoep
has given numerous international workshops on kinetic modelling and metaboliccontrol analysis.

Hans V. Westerhoff
did a Ph.D. on Thermodynamics and Control of Biological
Free Energy Transduction in 1983 with Karel van Dam in Amsterdam. As visit-
ing scientist at the National Institutes of ... reactivity and fast adaptation and for regulated gene expression. For some
7 years now he is leader of the Maxima Initiative for Theoretical Experimental
with Wayne Getz from Berkeley and Snoep ... of the European Systems Biology Initiative for Genomics and Health (ES-BIGH; www.systembiology.net).

Jannie Hofmeyr
started working in the fields of metabolic modelling and control
and developed a number of new ways of understanding the control and regulation of
metabolic pathways. Together with Henk Snoep and Johann Rohwer.

5. Hofmeyr, J.L. Snoep & H.V. Westerhoff—Metabolic Kinetics, Control and Regulation

6. Hofmeyr, J.L. Snoep & H.V. Westerhoff—Metabolic Kinetics, Control and Regulation
Contents

1 The Kinetic Model for a Network of Coupled Reactions 10
  1.1 The basic linkages and structures in metabolism ........... 10
  1.2 The kinetic model ..................................... 12
  1.3 A hierarchy of kinetic models ............................. 17
    1.3.1 The mass-action form of the kinetic model ............ 18
    1.3.2 The rate law form of the kinetic model ................ 20
    1.3.3 The power-law form of the kinetic model ............... 20
  1.4 Metabolic states ........................................ 21
    1.4.1 Equilibrium ........................................ 21
    1.4.2 Steady state ........................................ 21
    1.4.3 Transient state ........................................ 22
  1.5 The kinetic model in steady state ......................... 22
  1.6 Time hierarchy in metabolism ............................. 24
  1.7 Parameters and variables .................................. 24

2 The Analysis of Stoichiometric Matrices 26
  2.1 Matrix formulation of the kinetic model .................... 27
  2.2 Stoichiometric matrix analysis ........................... 31
    2.2.1 Relationships between steady-state fluxes .............. 31
    2.2.2 Relationships between differential equations — conservation relationships .......... 34
  2.3 A general procedure for stoichiometric matrix analysis ....... 40
    2.3.1 Steady-state flux relationships and the K-matrix ....... 42
    2.3.2 Conservation relationships and the L-matrix .......... 44
  2.4 Some complex examples .................................... 47
    2.4.1 The photosystem cent of Rhodobacter sphaeroides—a network of interwoven moiety-conserved cycles ........ 47
    2.4.2 The bacterial glutamine synthase system—a complexity branched network .................. 50
  2.5 Analysis of the basic metabolic structures ................. 52
    2.5.1 A 3-enzyme linear system (Fig. 2.5A) .......... 52
    2.5.2 A 3-enzyme branched system (Fig. 2.5B) ............. 55
    2.5.3 A 3-membered moiety-conserved cycle (Fig. 2.5C) ...... 57
    2.5.4 A 2-membered moiety-conserved cycle (Fig. 2.5D) ...... 60

3 Metabolic Control Analysis 66
  3.1 Quantifying metabolic control .............................. 66
    3.1.1 Local coefficients: Elasticity coefficients ............ 67
    3.1.2 Global coefficients: Response and control coefficients 70
  3.2 Control Properties ...................................... 74
    3.2.1 Summation properties of control coefficients .......... 74
    3.2.2 Connectivity properties ............................ 75

4 The Algebra of Metabolic Control Analysis 78
  4.1 Introduction ........................................... 78
  4.2 The kinetic model ....................................... 79
    4.2.1 Functional relationships in the steady state ........ 80
    4.2.2 Flux-relationships in the steady state ................ 82
  4.3 Differentiation of the steady-state equation ................ 82
  4.4 Metabolic control analysis ............................... 85
    4.4.1 Concentration response with respect to p ............ 85
    4.4.2 Concentration-control coefficients .................. 86
    4.4.3 Flux-response and control coefficients ............... 87
    4.4.4 Concentration-response with respect to T ............. 87
    4.4.5 Flux-response with respect to T ...................... 88
    4.4.6 Normalising the central equations ..................... 89
    4.4.7 Summation theorems .................................. 91
    4.4.8 Connectivity theorems ................................ 91
    4.4.9 The control-matrix equation ......................... 91
    4.4.10 The inverse problem .................................. 92
  4.5 Discussion ............................................... 93
  4.6 Appendix: The Jacobian matrix ............................ 93
  4.7 Appendix: An explicit example ............................ 95

5 Metabolic Regulation: Supply-Demand Analysis 99
  5.1 Introduction ........................................... 99
  5.2 Metabolic regulation, organisation and function ............ 99
  5.3 Quantitative analysis of supply-demand systems ............. 101
  5.4 Discussion ............................................... 106

6 Metabolic engineering of branched systems: redirecting the main pathway flux 107
  6.1 Summary ................................................. 107
  6.2 Introduction ............................................. 107
  6.3 Theory .................................................. 108
  6.4 Core Models ............................................. 111
1 The Kinetic Model for a Network of Coupled Reactions

When faced with the metabolic map we are struck by its seemingly bewildering complexity. Just as we search for structural motifs to try to make sense of protein structure, so should we look for underlying structure in metabolic networks. When wanting to study the dynamic behaviour of metabolic networks, we are not so much interested in reaction chemistry and mechanism (the main concern of metabolic studies thus far), but rather in the different ways in which reactions are coupled. In the first section of this chapter we show that the metabolic network is built up from a few basic linkage types and structures. This provides a good background against which a kinetic model of metabolism can be developed. We are of course primarily interested in metabolic structures in which reactions are catalysed by enzymes, but it should be emphasised that the kinetic model we develop in this chapter is applicable to any system of coupled chemical reactions. Although coupled reaction networks show many of the properties of electrical and hydrodynamic networks, there is one property of reaction networks that makes them unique, richer in behaviour, and therefore more interesting—this property is stoichiometry, the fixed ratios in which molecules react with each other. The existence of stoichiometry complicates the study of network behaviour, but this is more than compensated for by the new possibilities it opens up for novel behaviour. Although stoichiometry will of course be incorporated into the kinetic model right from the start, a systematic study of stoichiometry is left for Chapter 2.

1.1 The basic linkages and structures in metabolism

A fundamental feature of metabolic pathways is that any two functional steps, usually enzyme-catalysed reactions or transport steps, are linked by metabolites common to both (the product of one step becomes the substrate of the next). The two different types of linkages are shown in Fig. 1.1:

An important difference between these two types of linkage is that a type 1 linking metabolite is part of the material being metabolised and is passed on, whereas a type 2 linking metabolite alternates between two or more forms of a chemical group (called a moiety); it forms a permanent part of the system. The concentration of type 1 linker is in theory free to assume any value (however, there may be physical constraints on the concentrations of metabolic intermediates due to limits in the solvent capacity of water). On the other hand, the total concentration of
Figure 1.1: The two ways in which enzymic reactions are linked. A. The second enzyme takes the linking metabolite, Y, on to form a new substance, Z. B. The second enzyme reverses the action of the first enzyme as far as the linking metabolite is concerned. For example, if the first enzyme oxidises X, the second reduces it, or if the first enzyme phosphorylates X to form Y, the second dephosphorylates Y back to X. Of course, the second reaction is not a simple reversal of the first, but a different reaction catalysed by a different enzyme (e.g., the first enzyme may be a kinase that uses ATP to phosphorylate X, in which case the second enzyme usually is a phosphatase that hydrolysates the phosphoester bond). This is the well-known metabolic feature of reactions that are coupled by the interconversion of different forms of cofactors, such as NAD\(^+\)/NADH or ATP/ADP. These reactions are usually bimolecular reactions in which some or other chemical group (such as a hydride or phosphoryl group in the above examples) is transferred from a donor to an acceptor compound.

the different forms of a type 2 linking metabolite is often constant. The condition is that the formation and degradation of the moiety be slow in comparison to the interconversion of the different forms. Such type 2 structures are called moiety-conserved cycles. For example, ATP, ADP and AMP form an adenylate-conserving system, whereas pyridine nucleotide is conserved in the NAD\(^+\)/NADH couple. The concept of moiety conservation is so important that Reich and Sel’kov described energy metabolism as 'open flow through networks of moiety-conserved cycles'.

In metabolism there is a tendency to form chains consisting of only one type of linkage (Fig. 1.2). Many of the familiar metabolic pathways have as backbone a chain of type 1 linkages. On the other hand, in the electron transfer chain all linkages are of the second type (each carrier existing in two forms, the total amount being conserved). Often a type 2 chain might start from one of the steps of a type 1 chain and perhaps terminate in another type 1 chain or even in a later step of the same type 1 chain.

The two fundamental types of linkage are combined to form four basic metabolic structures (Fig. 1.3). In branched pathways depicted in Fig. 1.3B, a branchpoint occurs where a metabolite is committed by different enzymes to more than one end-product. A central branchpoint metabolite, such as acetyl-CoA or glutamine, can be the junction of a number of branches. Fig. 1.3B shows a branch with diverging fluxes. Reversing the direction of fluxes changes the system to one with converging branches. When two branches in a pathway reconverge a loop is formed (Fig. 1.3C). Flux through the two limbs of the loop can be either parallel or opposed. The simplest forms of loops are, first, reactions catalysed by two isozymes and, second, the substrate loop (also called the futile loop), which consists of two enzymes catalysing opposite directions of the same transformation (e.g. the kinase/phosphatase loops of glycolysis). Note that we do not, as is common in the literature, call such a structure a substrate or futile cycle; the term cycle is reserved for the distinct structural type depicted in Fig. 1.3D. Whereas linear, branched and looped structures can be formed by monomolecular reactions, the numerous bimolecular metabolic reactions can give rise to cycles (note that cycles consisting of monomolecular reactions are identical to loops without input and output; they form closed systems with no metabolic significance).

The main difference between substrate loops and cycles is that the input to and output from loops are at branchpoints formed by metabolites common to more than two reactions, while in cycles they are through bimolecular reactions (reactions 1 and 3 in Fig. 1.3D). The simplest type of cycle is the 2-member cycle mentioned above under type 2 linkages.

1.2 The kinetic model

In general, therefore, a metabolic system consists of a network of coupled enzymic reactions and membrane-associated transport steps. The structure of such networks is commonly displayed on metabolic maps, where each reaction is described in terms of the participating enzyme, metabolites and cofactors and the reaction stoichiometry. These chemical reactions and transport steps can be thought of as the primary connections between metabolite pools through which the pools affect each other by mass action. Metabolites and enzymes can also interact through regulatory loops, i.e., feedback and feedforward interactions; these can be thought of as secondary connections for transmitting information through the network. An-
Figure 1.3: The four basic metabolic structures: A. Linear chain B. Branched chain C. Loop D. Cycle. 

other important feature of many metabolic systems is that they are divided into different compartments with the same metabolite sometimes occurring in two or more compartments. For the purpose of analysing the dynamic behaviour of such systems the compartmentalised pools of the same metabolite must be regarded as separate metabolites even if they have the same chemical structure.

In large compartments and when diffusion processes are slow there may be a concentration gradient of a specific metabolite, which therefore cannot be represented by a uniform pool with a measurable size. However, as this complicates the analysis considerably and is, more often than not, inapplicable, we make the simplifying assumption that metabolites are at all times uniformly distributed through the space occupied by the compartment in which they occur, i.e., no concentration gradients exist within the system, ... Such metabolites can be regarded as thermodynamically defined in the sense that they can be assigned a chemical potential.

Consider a specific metabolite pool, S with concentration \( s \) (Fig. 1.4)\(^1\). The rate at which \( s \) changes depends on the magnitudes of individual rates of the reactions that produce or consume S. The rate at which \( s \) changes at any specific instance can be written as

\[
\frac{ds}{dt} = v_1 + v_2 + v_3 - v_4 - v_5 - v_6 = (v_1 + v_3 + v_5) - (v_4 + v_2 + v_6) \tag{1.1}
\]

where \( v_1 \) is the net rate of reaction 1, \( v_2 \) that of reaction 2, etc. These equations symbolise the statement that the rate at which \( s \) changes is equal to the sum of the rates of the reactions for which \( S \) is a product minus the sum of the rates of the reactions for which \( S \) is a substrate.

The net rate of each individual reaction can be expressed as the rate of change of the concentration of any reactant or product of that reaction. Thus, for a simple reaction, \( A + B \rightarrow C \), the rates of consumption of A and B are \( \frac{da}{dt} \) and \( \frac{db}{dt} \), while the rate of production of C is \( \frac{dc}{dt} \). For this reaction one can write

\[
v = -\frac{da}{dt} = -\frac{db}{dt} = \frac{dc}{dt}, \tag{1.2}
\]

Because one molecule of A must react with one molecule of B to form one molecule of C, it is immaterial which of these derivatives is used as a measure of the \( v \), as long as one keeps the signs correct.

However, for a reaction such as \( 2A + B \rightarrow 3C \) the situation is more complicated. Here, for example, \( a \) changes twice as fast as \( b \), while \( c \) changes three times as fast as \( b \). The three time derivatives are related by

\[
-\frac{1}{2} \frac{da}{dt} = -\frac{db}{dt} = \frac{1}{3} \frac{dc}{dt}. \tag{1.3}
\]

It should be clear that, in order to be unambiguous about what we mean by the ‘rate’ of a chemical reaction, we should state which time derivative we have chosen to represent the rate (e.g. \( -\frac{da}{dt} \) or \( \frac{dc}{dt} \)), and we should supply the balanced chemical equation for the reaction. Note that if the stoichiometries are not one, the value of the rate depends on which time derivative is chosen.

This procedure is sufficient if only one isolated reaction is studied, but as soon as reactions are coupled by common intermediates that can participate in these reactions with different stoichiometries, we need a more consistent definition of reaction rate. This can be given in terms of the degree of advancement, \( \xi \) (the Greek letter \( \xi \)), of the reaction, also called the extent of reaction. Consider a simple

---

\(^1\)Names of chemical species are denoted by a capital letters \( E, X, S, \ldots \) and their concentrations by lowercase italic letters \( e, x, s, \ldots \) rather than by the usual square brackets \([E], [S] \) or \([X]\).
reaction \( A \rightleftharpoons B \). If an infinitesimal amount \( d\xi \) moles of \( A \) is converted into \( B \), we can write
\[
\frac{\text{change in the amount of } A}{\text{change in the amount of } B} = -d\xi.
\]

The more complex reaction \( 2A + B \rightleftharpoons 3C \) serves to illustrate that, in general, any reaction can be treated in the same way. We could write the reaction in an alternative way:
\[
0 = -2A - B + 3C.
\]
The numbers \( -2 \), \( -1 \) and \( +3 \) are called the stoichiometric coefficients, the negative signs going with the reactants and the positive ones with the products. We shall use the symbol \( c \) for a stoichiometric coefficient, so that in general we can write for any equation:
\[
0 = c_A A + c_B B + c_C C + c_D D + \ldots
\]

When the reaction proceeds by an amount \( d\xi \), the amounts of the reactants and products change as follows:
\[
\begin{align*}
\text{change in the amount of } A &= -2d\xi, \quad \text{or } c_A d\xi \text{ in general} \\
\text{change in the amount of } B &= -d\xi, \quad \text{or } c_B d\xi \text{ in general} \\
\text{change in the amount of } C &= +3d\xi, \quad \text{or } c_C d\xi \text{ in general}.
\end{align*}
\]

Now that we understand what \( \xi \) means, we can define the rate, \( v \), of the reaction as
\[
v = \frac{d\xi}{dt},
\]
the rate of change of advancement of the reaction. In the above we saw that, say, a reaction \( A \rightleftharpoons B \) had rate of change \( da/dt = c_A d\xi/dt \) or, keeping in mind the definition of reaction rate in eq. 1.4, \( da/dt = c_A v \).

Therefore, reaction rate can be expressed as the rate of change of a reagent scaled by the inverse of its stoichiometric coefficient (i.e. \( v = (1/c_A) da/dt \)). For our example \( v \) can be seen to be any of the three scaled time derivatives in eq. 1.3.

In Fig. 1.4 we assumed that the stoichiometries with which \( S \) participates in the different reactions are all equal to one. Of course this need not be so and with the above general definition of reaction rate we have a way of accounting for differing stoichiometries in such equations. Let us explore this by considering a simple system that consists of two reactions, one of which produces compound \( S_1 \) while the other consumes it.

\[
\begin{align*}
X_0 &\xrightarrow{\mathbb{v}_1} 2S_1 \\
S_1 &\xrightarrow{\mathbb{v}_2} X_2
\end{align*}
\]

The individual reaction rates are
\[
\begin{align*}
v_1 &= -dx_0/\frac{dt}{2} = \frac{dS_1}{dt} \\
v_2 &= -dS_1/\frac{dt}{2} = dx_2/\frac{dt}{2}
\end{align*}
\]

It follows that for reactions 1.5 and 1.6 individually
\[
\frac{dS_1}{dt} = 2v_1 \quad \text{and} \quad \frac{dS_1}{dt} = -v_2
\]

If the two enzymes occur together, the rate at which \( S_1 \) changes at any time will be the sum of the individual rates at which the enzymes produce or consume \( S_1 \).

In general, therefore, the rates which affect the concentration of a metabolite will, in differential equations such as eq. 1.1, be weighted by the stoichiometric coefficients; \( c > 0 \) if the metabolite is a product of the reaction, \( c < 0 \) if it is a substrate, and, of course, \( c = 0 \) if the metabolite does not participate in the reaction.

With the above background we can now develop a general way of describing the rates at which the metabolite concentrations in a system change. Consider a system that consists of \( n \) steps that interconvert \( m \) metabolites. We use \( j \) as a counter for the \( 1 \) to \( m \) metabolites. Therefore \( S \) participates in step \( j \).

For any metabolite \( S_1 \) we can now write
\[
\frac{dS_i}{dt} = c_{i1}v_1 + c_{i2}v_2 + c_{i3}v_3 + \cdots + c_{i(n-1)}v_{n-1} + c_mv_n
\]

In shorthand summation notation this differential equation becomes
\[
\frac{dS_i}{dt} = \sum_{j=1}^{n} c_{ij}v_j
\]

Such an equation can be written for each metabolite so that the dynamic behaviour of the whole system is described by the following system of \( m \) differential equations which may be regarded as is the general kinetic model for a system of coupled reactions:
\[
\frac{dS_j}{dt} = \sum_{j=1}^{n} c_{ij}v_j \quad \text{for } i, j = 1, \ldots, m
\]

In general, \( S_i \) can denote enzymic intermediates, metabolite pools or groups of metabolite pools; \( v_j \) denotes the rate of functional steps, which can be elementary, enzyme, translocator, non-catalysed, or groups of reactions. The stoichiometric coefficients, \( c_{ij} \), describe the mass-action linkages in the reaction network, i.e. its topology. In general, the rates \( v_j \) are non-linear functions of the \( S_i \) and, except for very simple networks, the system has no analytical solution. However, starting from specified initial conditions, the time-course behaviour of \( S_i \) can be simulated by numerical integration of the differential equations, usually by digital computer.
Figure 1.5: A schematic representation of the enzymatic reaction $c_1S_1 + c_2S_2 = c_3S_3 + c_4S_4$ which is imbedded in a metabolic network. Step numbers are boxed and metabolites represented by either a subscripted X (for external metabolites) or a subscripted S (for internal variable metabolites). Stoichiometric coefficients are associated with the line connecting a metabolite to an enzyme box. Arrowheads are in accordance with the direction defined in the rate equation, and do not imply irreversibility. Effectors such as $X_e$ are linked to the enzyme box by arrows and the type of interaction indicated by $+$ (activation) or $-$ (inhibition).

Figure 1.6: A schematic representation of the reactions in eq. 1.5 and 1.6

Let us again consider the simple system described by reactions 1.5 and 1.6. Fig. 1.5 explains the schematic method that will be used throughout to depict the way in which reactions are coupled. One of its useful features is that it allows stoichiometric coefficients to be shown unambiguously.

Using this schematic method our simple model is depicted in Fig. 1.6. The full kinetic model for this system is (according to eq. 1.12)

$$
\begin{align*}
\frac{dx_0}{dt} &= -v_1 \\
\frac{ds_1}{dt} &= (2)v_1 - (1)v_2 \\
\frac{dx_2}{dt} &= (1)v_1 + (1)v_2 - v_2
\end{align*}
$$

Given explicit rate equations for reactions 1 and 2, this is a complete description of how $x_0$, $s_1$ and $x_2$ change with time.

1.3 A hierarchy of kinetic models

Up to now we have been rather vague about the identity of the metabolites, about the reactions that produce or consume them and about the equations that describe the rates of reaction $v$. The kinetic model is extremely versatile in that it can be used to describe many different levels of complexity; these levels differ in terms of how reactions are aggregated; in turn this aggregation is inextricably bound up with the time hierarchy of chemical interactions. At one extreme, chemical interactions can be described on the quantum level where the players are elementary particles. Through the atomic level one progresses to the individual molecular level to the level of populations of molecules. Here the kinetic model becomes a valid way of describing dynamics. The lowest rung on this part of the hierarchy is where reaction are described in terms of elementary reactions where two molecules associate to form a complex (e.g., protein-ligand association), a complex dissociates into two molecules, or a molecule or complex isomerises. The rates of these reactions are described in terms of first or second-order rate constants in combination with reactant concentrations. This can be called the mass-action level where molecular events happen in the micro to millisecond time-scale. Biochemists usually aggregate the elementary steps associated with an enzyme catalytic process into a single reaction, the rate of which is described by a rate-law, the most familiar being the Michaelis-Menten rate equation. At this level of the time scale events happen within seconds to minutes—this we regard as the intermediary metabolic time scale. The processes of transcription, translation, protein synthesis and degradation are slower (minutes to hours) and can be thought of as the genetic time-scale.

We now describe three important approaches that have been used to describe the dynamic behaviour of metabolic systems; they all depend on the way that enzymic reactions and their rates are represented mathematically.

1.3.1 The mass-action form of the kinetic model

In the mass-action approach each enzyme reaction is written in terms of elementary mechanistic steps. Reaction rates of each step are then described in terms of mass-action kinetics. The $s_i$ in the kinetic model then signify all metabolites and intermediary enzyme complexes (with a differential equation for each of them); the rates $v_j$ of the elementary steps in each enzyme mechanism are then usually expressed by simple first or second order rate equations.

For the system in Fig. 1.6 the mass-action approach will be based on the scheme in Fig. 1.7, which assumes a simple reversible mechanism for each enzyme reaction. The kinetic model is expressed in the following set of differential equations (here we use square brackets to depict concentrations):

$$
\begin{align*}
\frac{d[X_0]}{dt} &= -v_{1a} \\
\frac{d[E_1X_0]}{dt} &= v_{1a} - v_{1b} \\
\frac{d[E_2S_1]}{dt} &= v_{1b} - v_{1c} \\
\frac{d[S_1]}{dt} &= 2v_{1c} - v_{2a} \\
\frac{d[E_2S_1]}{dt} &= v_{2a} - v_{2b} \\
\frac{d[E_2X_2]}{dt} &= v_{2b} - v_{2c} \\
\frac{d[X_2]}{dt} &= v_{2c}
\end{align*}
$$
The rate law form of the kinetic model

In the rate-law approach, the rate of each enzymic reaction is described by a aggregated rate law (a net-flux equation) obtained from steady state kinetic studies. This most useful approximation follows from the assumption that enzyme concentrations are in general much lower than the concentrations of intermediary metabolites. However, this assumption only has to be made when the concentrations of substrates and products that enter the rate equations are total concentrations (i.e. the sum of free plus bound metabolite). If only free metabolite concentrations enter the rate equation then it is valid even at high enzyme concentration. Well-known examples of such rate laws are the Michaelis-Menten, Hill or Monod-Wyman-Changeux equations (or, more generally, the type of equations described by Cleland). Each rate law is a non-linear function of substrate, product and effector concentration (but linear in enzyme concentration), in conjunction with either fundamental rate constants (King-Altman form) or phenomenological constants such as $k_{cat}$, $K_{eq}$, $K_m$ and $K_i$ (Cleland form). An important point to note is that any allosteric feedback or feedforward effects are accounted for in these rate equations. In practice, the rate-law approach gives results that agree satisfactorily with those obtained by solving the mass-action form of the model and simplifies analysis considerably. If we assume that rate laws are known for reactions 1 and 2 in the system in Fig. 1.6, then the differential equations that form the kinetic model are those in eq. 1.13, while the rate laws themselves could be something like:

$$v_1 = \frac{V_1}{K_{1(X_0)}} \left( x_0 - \frac{s_1}{K_{eq1}} \right) \left( 1 + \frac{x_0}{K_{1(X_0)}} + \frac{s_1}{K_{1(S_1)}} \right)$$  

$$v_2 = \frac{V_2}{K_{2(S_1)}} \left( \frac{s_1}{K_{eq2}} - \frac{x_2}{K_{2(X_2)}} \right) \left( 1 + \frac{s_1}{K_{2(S_1)}} + \frac{x_2}{K_{2(X_2)}} \right)$$

(1.29)  

(1.30)

We shall mostly use the rate-law approach in our modelling exercises.

1.3.3 The power-law form of the kinetic model

In this form of the kinetic model, pioneered and used extensively by Savageau, all the reactions that produce a metabolite are aggregated as are those that consume the metabolite. The rates of each of these aggregated reactions are then described by a power-law equation. Any differential equation in the kinetic model therefore contains only two rate terms which simplifies the mathematical and numerical analysis considerably. For the system in Fig. 1.6, the detailed kinetic model in power-law notation will be (as this is just illustrative, irreversibility of both reactions is assumed):

$$dx_0/dt = -\alpha_1 x_0^{\beta_1(x_0)}$$

(1.31)
\[
\frac{ds_1}{dt} = \alpha_1 x_{g_1}^0 - \alpha_2 s_1 \quad \quad (1.32)
\]
\[
\frac{dx_2}{dt} = \alpha_2 s_1 \quad \quad (1.33)
\]
The $\alpha$-terms are rate constants and the $g$-terms apparent kinetic orders. Although use of this approach has its advantages, these aggregated reactions are often difficult to interpret in physical terms. We shall not use this approach at all.

In summary it can be seen that the essential ingredients of the description of a metabolic network are the stoichiometric reactions and an algebraic expression for the rate of each reaction. Any constraint on the concentrations of a group of metabolites is described by a conservation equation (this will be discussed fully in a later section).

### 1.4 Metabolic states

With the background of the previous sections we are now well equipped to understand the three possible states a metabolic system can be in:

1.4.1 Equilibrium

All metabolic pools have time-invariant values and the individual net reaction rates are zero (each reaction is itself in equilibrium). This situation, where there is detailed balance in each reaction, is, biochemically speaking, of little interest, as only isolated systems can achieve equilibrium. Systems in equilibrium are dead.

1.4.2 Steady state

Functioning metabolic systems are open, with several or all reactions being either ‘pushed’ or ‘pulled’ from the outside. As these pushing and pulling forces are not in equilibrium with each other, they ensure that a metabolic system rather approaches a steady state (also called a stationary state). Here the metabolite pools are also time-invariant (their net rates of change are zero), but the individual reaction rates are not zero, so that there are constant fluxes of matter through the system. In a non-growing system an additional condition for steady state is that of conservation of mass; the net import of mass into the system per unit time is equal to the net export per unit time. In a growing system in steady state the metabolite pools will not be constant because of a constant ‘flux to expansion’; after correcting for this the normal time-invariant steady state will be seen to obtain.

Theoretically, a steady state is never actually reached, only approached asymptotically. However, in practice one can regard a system as being ‘in’ the steady state. When for a certain set of conditions a steady state is ‘unique’, it will be approached no matter what the initial concentrations of the metabolite pools are. This should be distinguished from situations where more than one steady state is possible for a given set of parameters (multistationarity or multiple steady states); here, which steady state is approached depends on the initial conditions, i.e., on the history of the system. Metabolite pools may also show constant periodic (oscillatory) behaviour around a steady-state point; the oscillating motion of metabolite concentrations is then said to be in a ‘limit cycle’.

The concept of stability is of great importance in connection with the steady state. A metabolic system is continually subjected to fluctuations in its variables (even the parameters, which we regard as constant, fluctuate). A steady state cannot exist or be maintained if it is not stable with respect to these fluctuations and perturbations. In this discussion we assume the steady state to be asymptotically stable. This means that if we regard the set of steady-state concentrations as describing a point in a multidimensional space, then there exists a neighbourhood around that point from which the steady-state point is always approached in time, i.e. if the initial concentrations correspond to any point in that neighbourhood, the system eventually reaches the same steady state. Besides this dynamic stability we also assume structural stability, which means that after a perturbation in one of the parameters the system relaxes to a closely-neighbouring steady state.

1.4.3 Transient state

A metabolic system could be moving towards a steady-state from some initial condition, from one steady state to another after a perturbation in some parameter of the system, or back to the old steady state after a fluctuation in one of the internal metabolites; it is then said to be in the transient state.

Let us now review our kinetic model in terms of the steady state.

### 1.5 The kinetic model in steady state

Biochemical systems have two important properties: the functional units of the system (such as enzymes and translocators) display non-linear kinetics and the systems are open, that is, they interact with their environment through the exchange of matter and energy. We should know how to define the system so that it is open.

Open systems are not always easy to set up experimentally. One possibility is to inject pathway substrate(s) at a constant rate and remove pathway product(s) by some method (precipitation, adsorption). This means we supply the system with constant fluxes into and out of the system (source fluxes and sink fluxes respectively). Another possibility is to ensure that the concentrations of source substrates are so high or the extra-system volume so large that, for all practical purposes, they remain constant within the time scale of the experiment; this is the situation most frequently encountered. This means providing the system with a practically inexhaustible supply of pathway substrate and a practically unsaturable pool of waste
products. Sometimes a substrate or product of a pathway is a solid such as glyco-
meter, its numerical value could be measured as either that of \( J_1 \) or that of \( J_2 \); although \( J_0 \) and \( J_2 \) are clamped, one can measure flux in terms of the rate of disappearance of a radio-active tracer \( J_0 \) or of appearance of a radio-active tracer \( J_2 \); this flux-value is used in practice depends on which rate of change is measured.

Fortunately, for the purposes of modelling such experimental detail can be ig-
mored. We just define the concentrations of all pathway substrates, products andexternal effectors as constant (we thus regard these 'external' metabolites as 'fixed'). All those metabolites in the system that are either

\[
\sum_{i=1}^{n} c_i J_i \quad \text{for} \quad i = 1, \ldots, m
\]

Note that we use the symbol \( J \) to remind us that the reaction rates are now steady-
state fluxes. The term 'flux' is often used rather loosely as a general synonym for

reaction rate; in terms of this definition, flux has no meaning in the transient

equilibrium, so that the net rate is zero and the metabolite concentrations are time-invariant (i.e. \( dx_i/dt = 0 \)). If our general kinetic model is in steady state, we call the resulting system of

\[
0 = \sum_{i=1}^{n} J_i = \sum_{i=1}^{n} J_i
\]


dx_i/dt = J_i = J_0 \quad \text{(1.34)}


dx_i/dt = J_i = J_2 \quad \text{(1.35)}

Therefore, numerical solution by computer is the only way to obtain the steady state metabolite concentrations and rates. What is important isthat the steady state is fully determined by the parameters.

For our simple model in Fig. 1.6 there is only one balance equation, i.e. for S

\[
S
\]

1.7 Parameters and variables

Within a chosen time-scale, certain quantities can thus be regarded as constant, unless they are manipulated by the investigator. These are called the system parameters, which we shall also call the controllers. Obvious examples are environmental factors such as temperature and, if the system is well-buffered, pH. From the viewpoint of the metabolic time-scale one sees a system of slowly drifting equilibria and slowly drifting steady

states.

• The concentrations of those chemical species that are synthesised and de-

graded very slowly in relation to the metabolic reactions, e.g. enzymes, translo-

cators and certain metabolites in cofactors.
The concentrations of the external metabolites of the system: initial substrates, final products and other external effectors such as inhibitors, activators and hormones. These external metabolites must be buffered by the environment. They must be constant to keep the system open and allow a steady state to be approached.

Equilibrium constants, \( K_{eq} \), and enzymic constants such as \( k_{cat} \), \( K_M \) and \( K_i \) are expanded and generalised into the scaled matrix formulation of the kinetic model and the accompanying analysis that forms the basis of metabolic model and control analysis. For this purpose, a model is described that forms the basis of demonstration and leaves out many important features of the system, such as, for example, the variable nature of ATP and ADP and the branch from pyruvate to acetyl-CoA. The model is given in Fig. 2.1 and is a simplified representation of the system allowing a branch to the pentose phosphate cycle via glucose-6-phosphate dehydrogenase to be introduced.

The kinetic model for this system can be written down by inspection (we assume that the external metabolites glucose, lactate, ADP, ATP, \( P_i \), NADP⁺ and NADPH are).

The previous chapter gave a general introduction to the important concepts and ingredients of metabolic modelling, using a very simple model. Here these ideas are expanded and generalised into the scaled matrix formulation of the kinetic model and the accompanying analysis that forms the basis of demonstration and leaves out many important features of the system, such as, for example, the variable nature of ATP and ADP and the branch from pyruvate to acetyl-CoA. The model is given in Fig. 2.1 and is a simplified representation of the system allowing a branch to the pentose phosphate cycle via glucose-6-phosphate dehydrogenase to be introduced.
fixed parameters and therefore do not include them in the kinetic model:

\[
\begin{align*}
\frac{ds_1}{dt} &= v_1 - v_2 - v_6 \quad (2.1) \\
\frac{ds_2}{dt} &= 2v_2 - v_3 \quad (2.2) \\
\frac{ds_3}{dt} &= v_3 - v_4 \quad (2.3) \\
\frac{ds_4}{dt} &= v_4 - v_5 \quad (2.4) \\
\frac{ds_5}{dt} &= v_5 - v_7 \quad (2.5) \\
\frac{ds_6}{dt} &= v_7 - v_5 \quad (2.6)
\end{align*}
\]

### 2.1 Matrix formulation of the kinetic model

Although for reasonably simple systems such as the glycolytic skeleton model the representation eqs. 2.1–2.6 is adequate, it does not make the network structure of the system clear, something that is important when large and intricately networked systems are considered. We shall now develop an alternative representation, which, besides being more elegant, has the advantage of being amenable to the formidable analytical tools of linear algebra. This is the so-called matrix formulation. We first write the above equation system in matrix notation and then explain what everything means:

\[
\begin{bmatrix}
\frac{ds_1}{dt} \\
\frac{ds_2}{dt} \\
\frac{ds_3}{dt} \\
\frac{ds_4}{dt} \\
\frac{ds_5}{dt} \\
\frac{ds_6}{dt}
\end{bmatrix} =
\begin{bmatrix}
1 & -1 & 0 & 0 & 0 & -1 \\
0 & 2 & -1 & 0 & 0 & 0 \\
0 & 0 & 1 & -1 & 0 & 0 \\
0 & 0 & 0 & 1 & -1 & 0 \\
0 & 0 & -1 & 0 & 1 & 0 \\
0 & 0 & 1 & 0 & -1 & 0
\end{bmatrix}
\begin{bmatrix}
v_1 \\
v_2 \\
v_3 \\
v_4 \\
v_5 \\
v_6
\end{bmatrix}
\]

(2.7)

Let us start in the middle with the table of integers called the stoichiometric matrix. Such a table is an extremely useful and compact way of representing the network structure of any metabolic system of coupled reactions. There is nothing mysterious about this notation. A matrix is the name used in mathematics for a table (or array) of numbers\(^1\). In general, each row (of \(m\) rows) of the stoichiometric matrix refers to a metabolite and each column (of \(n\) columns) to a reaction. In the above 6 by 6 stoichiometric matrix the row and column identities are:

\[
\begin{array}{cccccc}
R_1 & R_2 & R_3 & R_4 & R_5 & R_6 \\
S_1 & 1 & -1 & 0 & 0 & -1 \\
S_2 & 0 & 2 & -1 & 0 & 0 \\
S_3 & 0 & 0 & 1 & -1 & 0 \\
S_4 & 0 & 0 & 0 & 1 & -1 \\
S_5 & 0 & 0 & -1 & 0 & 1 \\
S_6 & 0 & 0 & 1 & 0 & -1
\end{array}
\]

The individual elements of the matrix are the stoichiometric coefficients \(c_{ij}\), each of which specifies the number of molecules of metabolite \(S_i\) that participates in the stoichiometrically balanced equation of reaction \(j\). As noted before, a stoichiometric coefficient is positive if the metabolite is a product, negative if a substrate, and zero if a metabolite does not participate in a reaction. Whether a metabolite is regarded as a substrate or product depends on the convention used in the reaction equation.

Similarly, the columns

\[
\begin{bmatrix}
\frac{ds_1}{dt} \\
\frac{ds_2}{dt} \\
\frac{ds_3}{dt} \\
\frac{ds_4}{dt} \\
\frac{ds_5}{dt} \\
\frac{ds_6}{dt}
\end{bmatrix}
\]

and

\[
\begin{bmatrix}
v_1 \\
v_2 \\
v_3 \\
v_4 \\
v_5 \\
v_6
\end{bmatrix}
\]

are matrices, both having dimensions of 6 rows by 1 column. Such a matrix that consists of one column of entities is called a column vector (similarly, a matrix of one row by \(n\) columns is called a row vector). You may recall that, although a vector in three-dimensional space is often represented by an arrow from the origin to a point in the space, it may just as well be given by the coordinates of the point \((x,y,z)\). This coordinate notation is more useful in that it can be extended into \(n\)-dimensional space by just adding more elements to the vector, whereas it is impossible to represent such an \(n\)-dimensional vector geometrically. To illustrate the equivalence of the geometrical and matrix representations, consider a three-dimensional rate vector\(^2\):

\[
\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2.5 \\ 2 \end{bmatrix}
\]

(2.8)

Geometrically, it can be represented (Fig. 2.2) as the vector resulting from the addition of its three components along the rates axes.

\[
\begin{bmatrix}
1 \\
2.5 \\
2
\end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 2.5 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}
\]

(2.9)

\(^1\)This simple description does injustice to the power of matrix notation; a matrix can, for instance, be multi-dimensional, or an array of other matrices or an array of mathematical functions.

\(^2\)Vectors are usually represented by boldface lowercase symbols, matrices by boldface uppercase symbols.
Remember that eq. 2.7 is a representation of eqs. 2.1-2.6; therefore matrix multiplication must work in such a way that it reproduces the original system of differential equations. For example, the first component of the $\frac{ds}{dt}$ vector must result from multiplication of the rate vector into the first row of the stoichiometric matrix:

$$\begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & -1 \\ 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \end{bmatrix} = v_1 - v_2 - v_6 \quad (2.13)$$

Similarly, the second component of the $\frac{ds}{dt}$ vector must result from multiplication of the second row of the stoichiometric matrix into the rate vector:

$$\begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \end{bmatrix} = \begin{bmatrix} 0 & 2 & -1 & 0 & 0 & 0 \\ 1 & 2 & 1 & 0 & 0 & 0 \\ 3 & 0.5 & 1 & 4 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \\ -1 \\ -1 \end{bmatrix} = 2v_2 - v_3 \quad (2.14)$$

and so on. From these examples it should be clear that the multiplication operation is performed by multiplying the 1st element of the row with the 1st element of the column, the 2nd row element with the 2nd column element, the 3rd with the 3rd, and so on, after which all of these products are summed.

These operations are fundamental to all matrix multiplications; it produces a single number (called an inner product) from a row vector and column vector of equal length, e.g.

$$\begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \\ -1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 2 & -1 & 0 & 0 \\ 0 & 0 & 3 & 0.5 & 1 \\ 0 & 0 & 1 & 2 & 4 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} = 1.5 \quad (2.15)$$

Another useful way of viewing multiplication of a matrix by a vector is that the product $Nv$ is actually a combination of the six columns of $N$ where each column is weighted by a component of $v$:

$$\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ -1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ -1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ -1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \end{bmatrix} \quad (2.16)$$
It is now a simple matter to generalise this rule and extend it. We already have a way of referring to each entry of the stoichiometric matrix: the entry in the \(i\)th row (metabolite) and the \(j\)th column (reaction) is denoted by \(c_{ij}\). If \(N\) is an \(m\) by \(n\) matrix, then index \(i\) ranges from 1 to \(m\) and index \(j\) from 1 to \(n\); altogether there are \(mn\) entries that form a rectangular matrix. One subscript is enough for the \(J\)th component of the stoichiometric matrix: the entry in the \(J\)th column (reaction) is denoted by \(c_{ij}\). To describe the product \(Nv\), we could use the summation notation

\[
\sum_{j=1}^{n} c_{ij} v_j = \text{the } i\text{th component of the product } Nv \tag{2.17}
\]

This sum takes us along the \(i\)th row of \(N\), each entry multiplied by the corresponding component of \(v\). Note again that the length of the rows (number of columns of \(N\)) of must match the length of \(v\); an \(m\) by \(n\) matrix multiplies an \(n\)-dimensional vector.

A last point must be made clear: matrix multiplication, in contrast to normal algebraical multiplication, is \textit{not} associative. This means that \(A \cdot B \neq B \cdot A\). Even if the matrices are square and of the same dimension the two multiplications do not necessarily result in the same matrix. Another point is that if the product \(A \cdot B\) is at all to make sense, \(A\) must have the same number of columns as \(B\) has rows. The resulting product matrix then has the number of rows of \(A\) and the number of columns of \(B\). For example, the multiplication of a \(3 \times 4\) by a \(4 \times 2\) matrix results in a \(3 \times 2\) matrix.

### 2.2 Stoichiometric matrix analysis

So, at this stage we understand the matrix formulation of the kinetic model. Now comes a very important part of the modelling process: that of analysing the kinetic model in order to understand (i) the relationships between steady-state fluxes (which allows the partitioning of the fluxes into a set of independent and dependent fluxes), and (ii) the relationships, if any, between the differential equations (which allows us to find the constraints on the concentrations of metabolites—the conservation equations—and the partitioning of the concentrations into a set of independent and dependent concentrations).

#### 2.2.1 Relationships between steady-state fluxes

Before developing a general procedure based on the matrix formulation (in section 2.3), let us first do the analysis by hand in order to understand what the procedure involves. First, the steady-state flux relationships. In steady state, the differential equations 2.1–2.6 lead to the balance equations:

\[
J_1 - J_2 - J_6 = 0 \tag{2.18}
\]

\[
2J_2 - J_3 = 0 \tag{2.19}
\]

\[
J_3 - J_4 = 0 \tag{2.20}
\]

\[
J_4 - J_5 = 0 \tag{2.21}
\]

\[
J_5 - J_3 = 0 \tag{2.22}
\]

\[
J_3 - J_5 = 0 \tag{2.23}
\]

Note that we use \(J\) instead of \(v\) to indicate that we consider steady-state fluxes. Purely by inspection of these balance equations it is clear that in the steady state

\[
J_1 = J_2 + J_6 \tag{2.24}
\]

\[
2J_2 = J_3 = J_4 = J_5 \quad \text{or} \quad J_2 = 0.5J_3 = 0.5J_4 = 0.5J_5 \tag{2.25}
\]

This means that in steady state there are three general fluxes in the pathway: one through reaction 1, one through reaction 6, and one through reactions 2, 3, 4, 5 (note that \(J_2\) is related to \(J_3, J_4 \text{ and } J_5\) by a fixed stoichiometric factor). The fluxes are related by the above two equations. Usually, even for quite complicated models, finding the steady-state fluxes relationships is quite easy, but even so one can make mistakes and the more reliable, automatic procedure to be discussed further on is preferable.

There is a compact and elegant way to describe these flux relationships. It depends on a partitioning of the fluxes into a set of independent fluxes and a set of dependent fluxes. An independent set contains a minimum number of fluxes in terms of which all the dependent fluxes can be expressed. We shall develop a systematic procedure for partitioning into independent and dependent fluxes further on. Here we see, purely by inspection, that a valid set of independent fluxes would be, for example, \([J_2, J_6]\). All the other fluxes can be expressed in terms of these two fluxes as follows:

\[
J_1 = J_2 + J_6 \tag{2.26}
\]

\[
J_2 = J_2 \tag{2.27}
\]

\[
J_3 = 2J_2 \tag{2.28}
\]

\[
J_4 = 2J_2 \tag{2.29}
\]

\[
J_5 = 2J_2 \tag{2.30}
\]

\[
J_6 = J_6 \tag{2.31}
\]

In vector form this can be written as:

\[
\begin{bmatrix}
J_1 \\
J_2 \\
J_3 \\
J_4 \\
J_5 \\
J_6
\end{bmatrix} = \begin{bmatrix}
1 \\
1 \\
2 \\
2 \\
2 \\
0
\end{bmatrix} \begin{bmatrix}
J_2 \\
J_6
\end{bmatrix} \tag{2.32}
\]
Even more compactly, this equation can be expressed in terms of a matrix formed by combining the column vectors:

\[
\begin{bmatrix}
J_1 \\
J_2 \\
J_3 \\
J_4 \\
J_5 \\
J_6
\end{bmatrix}
= 
\begin{bmatrix}
1 & 1 \\
1 & 0 \\
2 & 0 \\
2 & 0 \\
2 & 0 \\
0 & 1
\end{bmatrix}
\begin{bmatrix}
J_2 \\
J_6
\end{bmatrix}
\tag{2.33}
\]

We give this special matrix that expresses all the flux relationships the symbol K. The symbol derives from the fact that the column vectors of this matrix actually generate the kernel (or nullspace) of the stoichiometric matrix. This so-called vector space is an important entity in linear algebra, and we shall explain its meaning in section 2.5 using simpler models where the kernel can actually be visualised geometrically.

By re-ordering the rows so as to bring those referring to independent fluxes to the top, it is always possible to form a submatrix that contains 1’s on the diagonal from top left to bottom right and 0’s everywhere else:

\[
\begin{bmatrix}
J_2 \\
J_6 \\
\dots \\
J_4 \\
J_5 \\
J_6
\end{bmatrix}
= 
\begin{bmatrix}
1 & 0 \\
0 & 1 \\
\dots & \dots \\
1 & 1 \\
2 & 0 \\
2 & 0
\end{bmatrix}
\begin{bmatrix}
J_2 \\
J_6
\end{bmatrix}
\tag{2.34}
\]

This special square matrix is called an identity matrix; in matrix algebra it fulfills the role that the number one fulfills in ordinary algebra, i.e., any matrix or vector multiplied by an identity matrix of the correct dimension equals the original matrix or vector. Try it for yourself: multiply the right-hand vector of independent fluxes into the identity submatrix and convince yourself that it gives the first two elements of the left-hand vector. The symbol used for the identity matrix is I, or, if the dimension is important, I_{n,n} (here I_{2,2}).

This re-ordering of the flux vector into independent fluxes J_i and dependent fluxes J_d ensures that the K-matrix can always be partitioned into two submatrices as follows:

\[
\begin{bmatrix}
J_i \\
J_d
\end{bmatrix}
= 
K_{0,i}
\begin{bmatrix}
I_{n-r} \\
K_0
\end{bmatrix}
\begin{bmatrix}
J_i
\end{bmatrix}
\tag{2.35}
\]

where the bottom r rows of the submatrix \(K_0\) correspond to the dependent fluxes. It should also be clear that there will always be \(n-r\) independent fluxes (remember that there are in general \(n\) reactions in the system). As we shall see further on, the number \(r\) (called the rank of a matrix) turns out to be an important property of stoichiometric matrices.

### 2.2.2 Relationships between differential equations — conservation relationships

**By inspection**

The relationships between the differential equations, if there are any, are often not obvious, especially in complicated models. These relationships usually arise from cycles in the model; writing down a scheme of the pathway in which none of the variable metabolites occur more than once is of great help in finding such cycles by inspection (however, for models with interweaved cycles this can lead to quite messy schemes).

Consider the glycolysis scheme in Fig. 2.1B. An obvious cyclic structure is formed by metabolites \(S_5\) and \(S_6\), which are interconverted by reactions 3 and 5. So, let us look at the differential equations that describe the rates of change of \(S_5\) and \(S_6\):

\[
\frac{ds_5}{dt} = v_3 - v_5
\tag{2.36}
\]

\[
\frac{ds_6}{dt} = v_5 - v_3
\tag{2.37}
\]

It is immediately obvious that

\[
\frac{ds_3}{dt} + \frac{ds_6}{dt} = 0
\tag{2.38}
\]

which can also be written as

\[
\frac{d}{dt}(s_5 + s_6) = 0
\tag{2.39}
\]

If the derivative of something is zero, then that something must be a constant, so that

\[
s_5 + s_6 = T_1
\tag{2.40}
\]

where \(T_1\) is a constant. Note that this constraint on the concentrations of \(S_5\) and \(S_6\) (and any other such constraints that follow from linear relationships between differential equations) is true for all metabolic states: the transient, steady and equilibrium states. We say the sum of concentrations \(T_1\) is conserved and we call an equation such as eq. 2.40 a conservation equation.

What does this constraint mean in physical terms? In Chapter 1 we started off our discussion of the kinetic model by describing the basic structures of metabolism. With this background we see that eq. 2.40 is a (moiety)-conservation equation for the cycle containing the oxidised and reduced forms of NAD. The equation states that, although the individual concentrations of NAD+ and NADH in this model definitions vary with the state of the system, the sum of their concentrations must remain constant.

Not so obvious is a second constraint on the concentrations of \(S_3\), \(S_4\) and \(S_5\). Their differential equations also sum to zero:

\[
\frac{ds_3}{dt} + \frac{ds_4}{dt} + \frac{ds_5}{dt} = v_3 - v_4 + v_4 - v_5 + v_5 - v_3 = 0
\tag{2.41}
\]
so that

\[ s_1 + s_4 + s_5 = T_2 \]  

(2.42)

This is a conservation equation for oxidised intermediates of the glycolytic pathway. The problem is that, once pointed out, the cycle is obvious, but to the untrained eye it is not easy to spot. We therefore need a foolproof way of finding all conservation equations. Note that mass conservations occur when some subset of the differential equations are linearly dependent\(^3\), by which we mean some some differential equations can be written as a linear combination of others, e.g.,

\[
\frac{ds_6}{dt} = -\frac{ds_5}{dt} \tag{2.43}
\]

\[
\frac{ds_4}{dt} = -\frac{ds_1}{dt} - \frac{ds_5}{dt} \tag{2.44}
\]

By Gaussian elimination

For simple models conservation equations can often be found just by inspection, but for complex systems a more systematic procedure is needed. The matrix generalisation (section 2.3) of the procedure described in the following paragraphs has the advantage of generating all the flux relationships and conservation equations, and helps with the partitioning of fluxes and concentrations into independent and dependent sets. It can be done by hand or by computer. First we show how the conservation equations can be obtained.

The procedure is formally called Gaussian elimination, which is just a fancy name for the set of rules for solving linear equations in more than one variable that we all learnt at school. The idea is to eliminate systematically one variable from all of the equations but one; this remaining equation we can then think of being stored in the bottom of a ‘box’. Then the next variable is eliminated from all of the remaining equations but one, which is then put on top of the equation already in the box. Then the third variable, and so on until one is left with the last equation which contains only one variable, which is therefore solved. This is the elimination part of the procedure. To solve all the remaining variables, the process then backtracks as follows: starting with the equation that lies on the top of the stack in the box and contains two variables, we substitute into it the solved variable and solve for the second variable; then these two variables are substituted into the next equation on top of the stack, solving for the third variable. Iteration of this process finally solves all the variables. This is called back substitution.

For the purpose of generating the required relationships, we actually need only the elimination part of the procedure, so let us try it on the set of equations 2.1–2.6. Regard the equations \( \mathbf{A} \mathbf{x} = \mathbf{b} \) as a case of the general set of linear equations

\[
\mathbf{Ax} = \mathbf{b} \tag{2.45}
\]

and subtract the 3rd from the 6th to obtain a new 6th equation:

\[
(v_3 - v_1) - (v_3 - v_4) = v_4 - v_5 = \frac{ds_6}{dt} - \frac{ds_5}{dt} \tag{2.46}
\]

Now \( v_1 \) occurs only in the 3rd equation, which goes into the box. Next we have to eliminate \( v_4 \) from the new 5th equation by adding the 4th to the new 5th:

\[
(v_5 - v_4) + (v_4 - v_3) = 0 = \frac{ds_5}{dt} + \frac{ds_4}{dt} + \frac{ds_5}{dt} \tag{2.47}
\]

and from the new 6th by subtracting the 4th from the new 6th:

\[
(v_4 - v_3) - (v_4 - v_5) = 0 = \frac{ds_6}{dt} - \frac{ds_5}{dt} - \frac{ds_4}{dt} \tag{2.48}
\]

Now \( v_4 \) occurs only in the 4th equation, which goes into the box. Adding eqs. 2.47 and 2.48 gives the more familiar constraint that we found before

\[
0 = \frac{ds_5}{dt} - \frac{ds_6}{dt} \tag{2.49}
\]

This demonstrates a very important point about conservation relationships: Unless there is only one conservation equation, the set of conservation equations is not unique, because multiples of any two equations can be combined to give a new (and valid) relationship. Nevertheless, one should always try to write the conservation equations in a form that makes physical sense. This means juggling them to form a set that contains only additions and no subtractions. We return to this matter again in the next section.

So, after the elimination procedure we are left with a set of equations that is still equivalent to the original set of differential equations but in which the constraints on the differential equations are made explicit

\[
v_1 - v_2 - v_6 = \frac{ds_1}{dt} \tag{2.50}
\]

\[
2v_2 - v_3 = \frac{ds_2}{dt} \tag{2.51}
\]

\[
v_3 - v_4 = \frac{ds_3}{dt} \tag{2.52}
\]

\[
v_4 - v_5 = \frac{ds_4}{dt} \tag{2.53}
\]

\[
0 = \frac{ds_5}{dt} + \frac{ds_4}{dt} + \frac{ds_5}{dt} \tag{2.54}
\]

\[
0 = \frac{ds_5}{dt} + \frac{ds_6}{dt} \tag{2.55}
\]
From these relationships it is evident that the time derivatives can be partitioned into a set of independent time derivatives, say \( \frac{ds_1}{dt}, \frac{ds_2}{dt}, \frac{ds_3}{dt}, \frac{ds_4}{dt}, \frac{ds_5}{dt} \), in terms of which all the other time derivatives could be expressed; these others \( \frac{ds_6}{dt}, \frac{ds_7}{dt} \) form the set of dependent time derivatives. There is always one dependent time derivative for each conservation equation. But, whereas \( \frac{ds_1}{dt} \) and \( \frac{ds_2}{dt} \) must be part of the independent set of time derivatives (they cannot be expressed in terms of the others), the choice of \( \frac{ds_3}{dt} \) and \( \frac{ds_5}{dt} \) as independent is not unique; there are other possibilities, such as \( \frac{ds_4}{dt} \) instead of \( \frac{ds_5}{dt} \). A consistent procedure for forming the independent and dependent sets of time derivatives will be given in the next section.

In vector form we can express the time derivatives in terms of the independent time derivatives as follows:

\[
\begin{bmatrix}
\frac{ds_1}{dt} \\
\frac{ds_2}{dt} \\
\frac{ds_3}{dt} \\
\frac{ds_4}{dt} \\
\frac{ds_5}{dt}
\end{bmatrix}
= \begin{bmatrix}
1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
\frac{ds_1}{dt} \\
\frac{ds_2}{dt} \\
\frac{ds_3}{dt} \\
\frac{ds_4}{dt} \\
\frac{ds_5}{dt}
\end{bmatrix}
\]

or

\[
\begin{bmatrix}
\frac{ds_1}{dt} \\
\frac{ds_2}{dt} \\
\frac{ds_3}{dt} \\
\frac{ds_4}{dt} \\
\frac{ds_5}{dt}
\end{bmatrix}
= \begin{bmatrix}
1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
\frac{ds_1}{dt} \\
\frac{ds_2}{dt} \\
\frac{ds_3}{dt} \\
\frac{ds_4}{dt} \\
\frac{ds_5}{dt}
\end{bmatrix}
\]

This can be expressed even more compactly by combining the column vectors into a matrix, as follows:

\[
\begin{bmatrix}
\frac{ds_1}{dt} \\
\frac{ds_2}{dt} \\
\frac{ds_3}{dt} \\
\frac{ds_4}{dt} \\
\frac{ds_5}{dt}
\end{bmatrix}
= \begin{bmatrix}
1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
\frac{ds_1}{dt} \\
\frac{ds_2}{dt} \\
\frac{ds_3}{dt} \\
\frac{ds_4}{dt} \\
\frac{ds_5}{dt}
\end{bmatrix}
\]

By re-ordering the rows so as to bring those referring to independent time derivatives to the top, it is always possible to form an identity submatrix with dimension corresponding to the number of independent time derivatives:

\[
\begin{bmatrix}
\frac{ds_1}{dt} \\
\frac{ds_2}{dt} \\
\frac{ds_3}{dt} \\
\frac{ds_4}{dt} \\
\frac{ds_5}{dt}
\end{bmatrix}
= \begin{bmatrix}
1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
\frac{ds_1}{dt} \\
\frac{ds_2}{dt} \\
\frac{ds_3}{dt} \\
\frac{ds_4}{dt} \\
\frac{ds_5}{dt}
\end{bmatrix}
\]

(2.59)

It should be clear from eq. 2.59 that, without losing any information, the kinetic model can be expressed in terms of only the independent differential equations, casting it into a new form that is extremely important for metabolic modelling and control analysis:

\[
\begin{bmatrix}
\frac{ds_1}{dt} \\
\frac{ds_2}{dt} \\
\frac{ds_3}{dt} \\
\frac{ds_4}{dt} \\
\frac{ds_5}{dt}
\end{bmatrix}
= \begin{bmatrix}
1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
\frac{ds_1}{dt} \\
\frac{ds_2}{dt} \\
\frac{ds_3}{dt} \\
\frac{ds_4}{dt} \\
\frac{ds_5}{dt}
\end{bmatrix}
\begin{bmatrix}
v_1 \\
v_2 \\
v_3 \\
v_4 \\
v_5 \\
v_6
\end{bmatrix}
\]

(2.60)

Note that those rows of the original stoichiometric matrix that refer to the dependent metabolites have been deleted without losing any information, i.e., if we re-order the original stoichiometric matrix by switching rows 4 and 5 (to agree with the order in the vector of time derivatives) it can be expressed as a product of two matrices:

\[
\begin{bmatrix}
1 & -1 & 0 & 0 & 0 & -1 \\
0 & 2 & -1 & 0 & 0 & 0 \\
0 & 0 & 1 & -1 & 0 & 0 \\
0 & 0 & -1 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 & -1 & 0 \\
0 & 0 & 1 & 0 & -1 & 0
\end{bmatrix}
= \begin{bmatrix}
1 & 0 & 0 & 0 & 0 & -1 \\
0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 1 \\
0 & 0 & 0 & -1 & 0 & 0 \\
0 & 0 & 1 & -1 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
v_1 \\
v_2 \\
v_3 \\
v_4 \\
v_5 \\
v_6
\end{bmatrix}
\]

(2.61)

(Check your understanding of matrix multiplication by verifying the correctness of this equation).

We now give this special matrix that expresses the dependencies among the differential equations its own name and symbol: the link matrix, \( L \). The reduced stoichiometric matrix that results from deletion of the dependent metabolite rows is symbolised by \( N \) and the re-ordered stoichiometric matrix \( N \) so that \( N = LN \). In symbolic matrix form the kinetic model equation 2.12 can therefore be written as

\[
\frac{ds}{dt} = LNv
\]

(2.62)

It should be clear that the link matrix \( L \) will always contain an identity matrix with dimension corresponding to that of the number of independent metabolites.
This number (the number of independent equations in a set of equations) is so important that it also gets a special name and symbol, the rank, r of a matrix. In the example $r = 4$. Any link matrix $L$ can therefore be partitioned into two submatrices:

$$L = \begin{bmatrix} L_s & L_o \end{bmatrix}$$

(2.63)

where $L_o$ refers to those rows of $L$ that express the dependent differential equations in terms of the independent ones. $L_o$ has $m - r = 2$ rows (remember that $m$ is the number of metabolites).

Similarly, the vector of time derivatives can be split into rates of change of independent metabolites, $s_i$, and dependent metabolites, $s_d$:

$$\frac{ds}{dt} = \begin{bmatrix} \frac{ds_i}{dt} \\ \frac{ds_d}{dt} \end{bmatrix} = \begin{bmatrix} s_i \\ s_d \end{bmatrix}$$

Using the partitioned $s$-vector and $L$-matrix we can write eq. 2.62 as

$$\frac{d}{dt} \begin{bmatrix} s_i \\ s_d \end{bmatrix} = \begin{bmatrix} L_s \\ L_o \end{bmatrix} \hat{N}_v$$

(2.65)

or, as two separate equations:

$$\frac{ds_i}{dt} = I_s \hat{N}_v = N_s \hat{v}$$

(2.66)

$$\frac{ds_d}{dt} = L_o \hat{N}_v = L_o \frac{ds_i}{dt}$$

(2.67)

From eq. 2.67 it follows that

$$\frac{ds_d}{dt} - L_o \frac{ds_i}{dt} = \frac{d}{dt} (s_d - L_o s_i) = 0$$

(2.68)

which leads to the set of conservation equations:

$$s_d - L_o s_i = T$$

(2.69)

where $T$ is an $m - r = 2$-dimensional vector of conservation sums.

In explicit matrix form eq. 2.69 is written as

$$\begin{bmatrix} s_4 \\ s_5 \end{bmatrix} - \begin{bmatrix} 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ s_3 \\ s_5 \end{bmatrix} = \begin{bmatrix} T_2 \\ T_3 \end{bmatrix}$$

(2.70)

which, when expanded, can be seen to be equal to eqs. 2.40 and 2.42.

### 2.3 A general procedure for stoichiometric matrix analysis

You will probably agree that the above process of Gaussian elimination by hand seems quite cumbersome, even for this relatively simple model. Fortunately, it is here that matrix notation comes to the rescue to give a transparent and systematic view of all of the steps in the process. What we need is to represent Gaussian elimination in a way that clearly records the row manipulations. How this should be done is suggested by the following way of writing the kinetic model in which an identity matrix pre-multiplies the time derivative vector:

$$\begin{bmatrix} 1 & -1 & 0 & 0 & 0 & -1 \\ 0 & 2 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \frac{ds}{dt} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \frac{ds_i}{dt}$$

(2.71)

If row manipulations are performed on the stoichiometric matrix, they must simultaneously be performed on the identity matrix, thereby recording exactly which multiple of which row was added to which. This means that we only have to consider the manipulation of the following matrix (the dot-notation is shorthand for $d/dt$):

$$\begin{bmatrix} 1 & -1 & 0 & 0 & 0 & -1 \\ 0 & 2 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

(2.72)

This is called an augmented matrix obtained by placing side by side the stoichiometric matrix and an identity matrix of dimension equal to the number of metabolites (rows of $N$). Each column of the identity matrix refers to a different time derivative.

In terms of this matrix formulation, the objective of Gaussian elimination is to perform row manipulations on the whole augmented matrix until a special form of the stoichiometric matrix is obtained, i.e. the so-called row echelon form. A matrix in row echelon form has the following type of structure

$$\begin{bmatrix} * & * & * & * & * & * & * & * \\ 0 & * & * & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

(2.73)
The important feature is that the bottom block of zero entries form a kind of staircase pattern, or, alternatively, the starred entries form an upside down staircase. The circled entries on the corners of the stairs are called pivots; they are nonzero. The other starred entries may or may not be zero. Note that each stair of the upside down staircase must be one row high, but can be broader than one column.

How do we get the stoichiometric part of the augmented matrix in 2.72 into row echelon form? There is a systematic Gaussian elimination algorithm that computers find easy to understand, but put down in words makes for pretty incomprensible (and boring) reading. We humans have the advantage of, knowing what the end result should be, seeing directly how it can be achieved with the minimum fuss. It is obvious that the staircase pattern already exists for columns 1 and 2 (this is mainly because of fortuitous numbering of metabolites; one should always check whether renumbering the metabolites may not already do most of the job). The second obvious feature is that one can reduce row 6 to zero s by adding row 5 to it giving the following matrix (row operations are of course performed on whole rows of the augmented matrix):

\[
\begin{bmatrix}
S_1 & R_1 & R_2 & R_3 & R_4 & R_5 & R_6 & s_1 & s_2 & s_3 & s_4 & s_5 & s_6 \\
S_2 & 1 & -1 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 \\
S_3 & 0 & 2 & -1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
S_4 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
S_5 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
S_6 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 1 & 0 \\
\end{bmatrix}
\]

Note that the addition is automatically recorded as a 1 in row 6 column 5 of the identity matrix (as it is no longer an identity matrix we shall from now on refer to it as the righthand matrix). Turning now to column 3 we see that to form a corner of the staircase at the 1 in row 3 column 3 we have to eliminate −1 in row 5 by addition of row 3 to give:

\[
\begin{bmatrix}
S_1 & R_1 & R_2 & R_3 & R_4 & R_5 & R_6 & s_1 & s_2 & s_3 & s_4 & s_5 & s_6 \\
S_2 & 1 & -1 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 \\
S_3 & 0 & 2 & -1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
S_4 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
S_5 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
S_6 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 1 & 0 \\
\end{bmatrix}
\]

An extra 1 in row 5 column 4 of the righthand matrix records this last row operation.

An extra 1 in row 5 column 4 of the righthand matrix records this last row operation.

Let us review what we have accomplished. By Gaussian elimination we have transformed eq. 2.71 into the following equivalent equation:

\[
\begin{bmatrix}
1 & -1 & 0 & 0 & 0 & -1 & \frac{ds_1}{dt} \\
0 & 2 & -1 & 0 & 0 & 0 & \frac{ds_2}{dt} \\
0 & 0 & 1 & -1 & 0 & 0 & \frac{ds_3}{dt} \\
0 & 0 & 0 & 1 & -1 & 0 & \frac{ds_4}{dt} \\
0 & 0 & 0 & 0 & 1 & -1 & \frac{ds_5}{dt} \\
0 & 0 & 0 & 0 & 0 & 1 & \frac{ds_6}{dt} \\
\end{bmatrix}
\]

(This matrix equation is identical to the set of equations 2.50–2.55.) Note that during the elimination procedure we used only row additions. This ensured that the elements added to the identity matrix were all positive, which you will see in the end ensures that concentrations in the conservation equations are only added, never subtracted. If at all possible, try to stick to this procedure; the resulting conservation relationships are easier to relate to cycles in the reaction network (note, however, that there are situations (such as discussed in section 2.4.1) where one cannot avoid ending up with negative elements).

The above procedure constitutes the first essential step in the systematic analysis of stoichiometric matrices. From the transformed augmented matrix everything we need can be deduced, as discussed in the next two sections.

### 2.3.1 Steady-state flux relationships and the K-matrix

The first step is to identify the columns in the reduced stoichiometric matrix that contain pivots (the corners of the staircase, which have been circled in eq. 2.76). These fluxes \(J_1, J_2, J_3, J_4\) will be a valid set of dependent fluxes which can be expressed as linear combinations of the independent fluxes \(J_5\) and \(J_6\) (the fluxes referred to by columns that do not contain pivots). There are other valid choices of independent and dependent fluxes (remember that we chose \(J_5\) and \(J_6\) previously), but we want a systematic and generally applicable procedure for choosing such a set.
One of the useful features of Gaussian elimination is that after elimination the steady-state equation can be written by setting eq. 2.77 to zero, as all the time derivatives are zero in the steady state:

\[
\begin{bmatrix}
1 & -1 & 0 & 0 & 0 & -1 \\
0 & 2 & -1 & 0 & 0 & 0 \\
0 & 0 & 1 & -1 & 0 & 0 \\
0 & 0 & 0 & 1 & -1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
J_1 \\
J_2 \\
J_3 \\
J_4 \\
J_5 \\
J_6
\end{bmatrix}
= 
\begin{bmatrix}
0 \\
0 \\
0 \\
0 \\
0 \\
0
\end{bmatrix}
\]  

(2.78)

As this equation is valid only in the steady state, we replace rates with steady-state fluxes. From this form of the equation one can find, just by reading them off from the matrix, all the flux relationships (they are exactly those we found by inspection, i.e., eq. 2.24 and 2.25). Now it is a simple matter to express all the fluxes in terms of the independent fluxes \( J_1 \) and \( J_6 \):

\[
\begin{align*}
J_1 &= J_2 + J_6 = 0.5J_3 + J_6 \\
J_2 &= 0.5J_3 = 0.5J_5 \\
J_3 &= J_4 = J_5 \\
J_4 &= J_3 = J_5 \\
J_5 &= J_6 \\
J_6 &= J_6
\end{align*}
\]  

(2.79)

or in vector form

\[
\begin{bmatrix}
J_1 \\
J_2 \\
J_3 \\
J_4 \\
J_5 \\
J_6
\end{bmatrix}
= 
\begin{bmatrix}
0.5 & 0 & 0 & 0 & 0 & 1 \\
0.5 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
J_5 \\
J_6
\end{bmatrix}
\]

(2.80)

As before we combine the column vectors to form the \( K \)-matrix and re-order the rows so that the independent fluxes are at the top to form an identity submatrix in \( K \):

\[
\begin{bmatrix}
J_3 \\
J_5 \\
\cdots \\
J_1 \\
J_2 \\
J_4 \\
\end{bmatrix}
= 
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.5 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.5 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
J_5 \\
J_6 \\
\cdots \\
J_1 \\
J_2 \\
J_4
\end{bmatrix}
\]

(2.81)

It should be clear that this \( K \)-matrix is not unique. If another set of independent fluxes had been chosen, another \( K \) would have resulted (see eq. 2.34).

2.3.2 Conservation relationships and the \( \mathbf{L} \)-matrix

Now let us find the independent and dependent metabolites. The last two rows of the transformed augmented matrix in 2.76, which contain only zeros on the stoichiometric matrix side, express the linear dependencies among the differential equations on the righthand matrix side (which lead to the conservation equations). Let us call these \( m - r \) rows the **conservation matrix**:

\[
\begin{bmatrix}
\dot{s}_1 & \dot{s}_2 & \dot{s}_3 & \dot{s}_4 & \dot{s}_5 & \dot{s}_6 \\
0 & 0 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 1
\end{bmatrix}
\]

(2.87)

There are as many dependent metabolites as there are conservation relationships. The task at hand is to partition the metabolites into \( r \) (here 4) independent metabolites and \( m - r \) (here 2) dependent metabolites in such a way that the dependent metabolites are associated with one and only one of the above conservation rows. Often this can be accomplished by just re-ordering the columns: in the above example switching columns 4 and 5 leads to:

\[
\begin{bmatrix}
\dot{s}_1 & \dot{s}_2 & \dot{s}_3 & \dot{s}_4 & \dot{s}_5 & \dot{s}_6 \\
0 & 0 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 1
\end{bmatrix}
\]

(2.88)

Here the set of independent metabolites is \([s_1, s_2, s_5, s_4]\) and the dependent set is \([s_3, s_6]\). Note the righthand identity matrix formed by the \( s_4 \) and \( s_6 \) columns. It should be clear that each dependent metabolite is associated with one and only one of the above conservation rows.

We could have accomplished a similar result by switching columns 3 and 5:

\[
\begin{bmatrix}
\dot{s}_1 & \dot{s}_2 & \dot{s}_3 & \dot{s}_4 & \dot{s}_5 & \dot{s}_6 \\
0 & 0 & 1 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 1
\end{bmatrix}
\]

(2.89)

This shows a different but just as valid partitioning of the metabolites: the independent set is \([s_1, s_3, s_1, s_1]\) and the dependent set is \([s_2, s_6]\). When it is not possible to generate the righthand identity matrix just by re-ordering columns, the rows must be appropriately transformed by adding multiples of rows to other rows (this leaves the reduced stoichiometric matrix, which contains only zeros in these rows, undisturbed). This procedure can also be illustrated using the above example: subtracting the first row of eq. 2.87 from the second (adding \(-1\) times the first row to the second) gives:

\[
\begin{bmatrix}
\dot{s}_1 & \dot{s}_2 & \dot{s}_3 & \dot{s}_4 & \dot{s}_5 & \dot{s}_6 \\
0 & 0 & 1 & 1 & 0 & 0 \\
0 & 0 & -1 & 0 & 0 & 1
\end{bmatrix}
\]

(2.90)
Here the righthand identity matrix is generated automatically. In other cases it may be necessary to switch columns after transforming. Here we have a new partitioning of metabolites (the independent set is \([s_1, s_2, s_3, s_4]\); the dependent set is \([s_5, s_6]\)) and a transformed set of conservation equations.

This last form of the conservation equations would have been obtained automatically had we gone about the elimination procedure differently. Originally, we started off by getting rid of the last row. However, we could have started off by eliminating all the elements under the 1 in position \(c_{33}\) by adding row 3 to row 5 and subtracting row 3 from row 6 to give

\[
\begin{pmatrix}
R_1 & R_2 & R_3 & R_4 & R_5 & R_6 \\
1 & -1 & 0 & 0 & 0 & -1 \\
0 & 2 & -1 & 0 & 0 & 0 \\
0 & 0 & 1 & -1 & 0 & 0 \\
0 & 0 & 0 & 1 & -1 & 0 \\
0 & 0 & 0 & 0 & 1 & -1 \\
\end{pmatrix}
\begin{pmatrix}
\dot{s}_1 \\
\dot{s}_2 \\
\dot{s}_3 \\
\dot{s}_4 \\
\dot{s}_5 \\
\dot{s}_6 \\
\end{pmatrix}
\]

(2.91)

The last elimination step would then consist of adding row 4 to row 5 and subtracting row 4 from row 6 to give:

\[
\begin{pmatrix}
R_1 & R_2 & R_3 & R_4 & R_5 & R_6 \\
1 & -1 & 0 & 0 & 0 & -1 \\
0 & 2 & -1 & 0 & 0 & 0 \\
0 & 0 & 1 & -1 & 0 & 0 \\
0 & 0 & 0 & 1 & -1 & 0 \\
0 & 0 & 0 & 0 & 1 & -1 \\
\end{pmatrix}
\begin{pmatrix}
\dot{s}_1 \\
\dot{s}_2 \\
\dot{s}_3 \\
\dot{s}_4 \\
\dot{s}_5 \\
\dot{s}_6 \\
\end{pmatrix}
\]

(2.92)

The echelon form of the stoichiometric matrix (first four rows) is identical to before, but the last two rows of the righthand matrix are now the same as in eq. 2.90.

However, although the different partitionings of metabolites into independent and dependent sets, and the different forms of the conservation equations, are all valid and equivalent in the mathematical sense and for computational purposes, they are not all equivalent in terms of ease of physical interpretation. The most useful form of a conservation equation is one in which all the elements (weights) are positive. As noted above, these conservation equations can usually be mapped directly onto physical cycles in the reaction network scheme. If the elimination procedure generates negative elements in the conservation matrix, it is worthwhile to check if one cannot get rid of them by judicious recombination of the last \(m - r\) conservation rows (note that as long as only these rows are recombined, the echelon form remains unaltered). Of course, this type of recombination must be done before partitioning the concentrations into independent and dependent sets and constructing the L-matrix, the next step in our procedure.

In fact, we already know how to construct the L-matrix once the dependent and independent metabolites are known (eq. 2.59). First an \(r\)-dimensional identity matrix is written (each row of which refers to an independent metabolite). Then \(m - r\) rows are appended to the bottom. These rows are obtained by changing the sign of all elements in the first \(r\) columns of the transformed and column re-ordered conservation matrix (e.g. 2.88). These \(m - r\) rows of \(r\) elements form the \(L_0\) part of \(L\), which shows how the dependent time derivatives are expressed in terms of the independent time derivatives.

Here are the possible L-matrices for the glycolysis example. In principle, any of them can be used (the lefthand column of metabolites shows the numbering):

From eq. 2.88:

\[
\begin{pmatrix}
\dot{s}_1 \\
\dot{s}_2 \\
\dot{s}_3 \\
\dot{s}_4 \\
\dot{s}_5 \\
\dot{s}_6 \\
\end{pmatrix}
\begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 1 & 0 & 1 \\
0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 \\
\end{pmatrix}
\]

(2.93)

From eq. 2.89:

\[
\begin{pmatrix}
\dot{s}_1 \\
\dot{s}_2 \\
\dot{s}_3 \\
\dot{s}_4 \\
\dot{s}_5 \\
\dot{s}_6 \\
\end{pmatrix}
\begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \\
0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 \\
\end{pmatrix}
\]

(2.94)

In both of the above matrices the metabolite numbers do not always correspond with the row number. Once a suitable L-matrix has been found, this can of course be rectified by renumbering the metabolites in the scheme. This is unnecessary in the last L-matrix which follows from eq. 2.90:

\[
\begin{pmatrix}
\dot{s}_1 \\
\dot{s}_2 \\
\dot{s}_3 \\
\dot{s}_4 \\
\dot{s}_5 \\
\dot{s}_6 \\
\end{pmatrix}
\begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \\
0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 \\
\end{pmatrix}
\]

(2.95)

In the light of what was said above about obtaining a physically meaningful set of conservation equations in which all weights are positive, it is clear that this translates into all the elements of \(L_0\) being either zero or negative. This holds only for the first two of the above L-matrices.
2.4 Some complex examples

We now apply the general analytical procedure described in the previous section to reasonably complex networks to make sure that the procedure is understood. The first network has interweaved moiety-conserved cycles, while the second shows complex branching.

2.4.1 The photoreaction centre of *Rhodobacter sphaeroides*—a network of interweaved moiety-conserved cycles

The first example is a simplified model of the photoreaction centre of *Rhodobacter sphaeroides*. In terms of stoichiometric matrix analysis this model is interesting as it is totally closed with regard to mass. The free energy source that drives it away from equilibrium into a steady state is light. The model is deficient in that the energy sink has been left out; however, it serves as an interesting example of interweaved moiety-conserved cycles. Note that it would be messy to try to depict all of these cycles explicitly on the diagram by closing the C/C
couples.

![Diagram of the photoreaction centre of *Rhodobacter sphaeroides*](image)

Figure 2.3: A 3-state model of the photoreaction centre of *Rhodobacter sphaeroides*. Instead of numbered metabolites, we here use abbreviations: C and Ce are the oxidised and reduced forms of cytochrome c; Q and Qe2 are the oxidised and reduced forms of coenzyme Q; D, De and De2 are the three oxidation states of the photoreaction centre; hν represents light input.

The augmented stoichiometric matrix for this system would be

\[
\begin{bmatrix}
Q & Qe & D & De & De2 & C & Ce \\
Q & -1 & 0 & 0 & 0 & 0 & 0 \\
Qe & 0 & 0 & 0 & 0 & 0 & 0 \\
D & 0 & -1 & 0 & 1 & 0 & 0 \\
De & 0 & 0 & 0 & 0 & 0 & 0 \\
De2 & 0 & 0 & 0 & 0 & 0 & 0 \\
C & 0 & 0 & 0 & 0 & 0 & 0 \\
Ce & 0 & 0 & 0 & 0 & 0 & 0 \\
\end{bmatrix}
\]

(2.96)

The first elimination step adds row 1 to rows 2 and 3 and subtracts row 1 from row 5. We also see that it is possible to reduce row 7 to zeros by adding row 6 to it:

\[
\begin{bmatrix}
Q & Qe & D & De & De2 & C & Ce \\
Q & -1 & 0 & 0 & 1 & 0 & 0 \\
Qe & 0 & 0 & 0 & 0 & 0 & 0 \\
D & 0 & -1 & 0 & 1 & 1 & 0 \\
De & 0 & 0 & 0 & 0 & 0 & 0 \\
De2 & 0 & 0 & 0 & 0 & 0 & 0 \\
C & 0 & 0 & 0 & 0 & 0 & 0 \\
Ce & 0 & 0 & 0 & 0 & 0 & 0 \\
\end{bmatrix}
\]

(2.97)

Row 2 in the lefthand matrix has been reduced to zeros and will have to be shifted down. We leave this for later and continue eliminating the bottom elements in column 2 by adding row 3 to rows 4 and 6 to give:

\[
\begin{bmatrix}
Q & Qe & D & De & De2 & C & Ce \\
Q & -1 & 0 & 0 & 1 & 0 & 0 \\
Qe & 0 & 0 & 0 & 0 & 0 & 0 \\
D & 0 & -1 & 0 & 1 & 1 & 0 \\
De & 0 & 0 & 0 & 0 & 0 & 0 \\
De2 & 0 & 0 & 0 & 0 & 0 & 0 \\
C & 0 & 0 & 0 & 0 & 0 & 0 \\
Ce & 0 & 0 & 0 & 0 & 0 & 0 \\
\end{bmatrix}
\]

(2.98)

The last elimination step consists of adding row 4 to rows 5 and 6 to give:

\[
\begin{bmatrix}
Q & Qe & D & De & De2 & C & Ce \\
Q & -1 & 0 & 0 & 1 & 0 & 0 \\
Qe & 0 & 0 & 0 & 0 & 0 & 0 \\
D & 0 & -1 & 0 & 1 & 1 & 0 \\
De & 0 & 0 & 0 & 0 & 0 & 0 \\
De2 & 0 & 0 & 0 & 0 & 0 & 0 \\
C & 0 & 0 & 0 & 0 & 0 & 0 \\
Ce & 0 & 0 & 0 & 0 & 0 & 0 \\
\end{bmatrix}
\]

(2.99)

Now the rows are re-ordered by shifting row 2 down to the other zero rows:
The rank of the stoichiometric matrix is 3 (the number of rows of the lefthand matrix that contain some nonzero elements). The conservation matrix consists of four rows:

\[
\begin{bmatrix}
Q & Q_e & D & D_e & D_e^2 & C & C_e \\
1 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 & 1 & 0 & 0 \\
2 & 0 & 2 & 1 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 1 \\
\end{bmatrix}
\]

(2.101)

A 4-dimensional identity matrix can be obtained by subtracting row 3 from row 4 and shifting column 2 two positions to the right to become column 4:

\[
\begin{bmatrix}
Q & D & D_e & Q_e & D_e^2 & C & C_e \\
1 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 & 1 & 0 & 0 \\
2 & 2 & 1 & 0 & 0 & 1 & 0 \\
-2 & -2 & -1 & 0 & 0 & 0 & 1 \\
\end{bmatrix}
\]

(2.102)

The first row refers to conservation of \( Q \), the second to conservation of \( D \). The third can be interpreted as conservation of oxidised species; \( Q \) and \( D \) can accept two electrons, \( D_e \) and \( C \) one electron. The fourth row is the conservation of \( C \) minus the conservation equation for oxidised metabolites. In this example it is impossible to avoid the introduction of negative elements during the transformation of the conservation matrix into the required form.

The metabolites can therefore be partitioned into an independent set \([Q, D, D_e]\) and a dependent set \([Q_e, D_e^2, C, C_e]\). The \( L \)-matrix is

\[
\begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1 \\
-1 & 0 & 0 \\
0 & -1 & -1 \\
-2 & -2 & -1 \\
2 & 2 & 1 \\
\end{bmatrix}
\]

(2.103)

and the \( ds/dt = L ds/dt \) equation is

\[
\begin{bmatrix}
dQ/dt \\
dD/dt \\
dD_e/dt \\
dQ_e/dt \\
dD_e^2/dt \\
dC/dt \\
dC_e/dt \\
\end{bmatrix} =
\begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1 \\
-1 & 0 & 0 \\
0 & -1 & -1 \\
-2 & -2 & -1 \\
2 & 2 & 1 \\
\end{bmatrix}
\begin{bmatrix}
dQ/dt \\
dD/dt \\
dD_e/dt \\
dQ_e/dt \\
dD_e^2/dt \\
dC/dt \\
dC_e/dt \\
\end{bmatrix}
\]

(2.104)

We know that \( ds/dt = N v \). \( N \) is formed by deleting the rows that refer to the dependent metabolites from the original stoichiometric matrix, so that

\[
\begin{bmatrix}
dQ/dt \\
dD/dt \\
dDe/dt \\
dQe/dt \\
dD_e^2/dt \\
Ci/dt \\
Ce/dt \\
\end{bmatrix} =
\begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1 \\
-1 & 0 & 0 \\
0 & -1 & -1 \\
-2 & -2 & -1 \\
2 & 2 & 1 \\
\end{bmatrix}
\begin{bmatrix}
v_1 \\
v_2 \\
v_3 \\
v_4 \\
\end{bmatrix}
\]

(2.105)

There is only one general steady-state flux through this system. Although the form of the reduced stoichiometric matrix obtained above suggests \( J_4 \) as independent flux, any of the other three would serve as well. As all reaction fluxes are equal the steady state of this system, the \( K \)-matrix is a column vector of four 1’s.

### 2.4.2 The bacterial glutamine synthase system—a complexly branched network

In the glycolysis example we had only one simple branching of a flux in the reaction network. This example illustrates the construction of the \( K \)-matrix for a system in which there is complex branching. The model of bacterial nitrogen assimilation by glutamine synthase (Fig. 2.4) is based on that of Kahn and Westerhoff (JTB). There are no moiety-conserved cycles, although this is not obvious when just inspecting the network scheme.

![Figure 2.4: A model for bacterial nitrogen assimilation by glutamine synthase.](image)

The ammonia concentration is assumed to be fixed and is therefore a parameter of the system. Metabolites: Glu, glutamic acid; Gln, glutamine; 2OG, 2-oxoglutarate. Enzymes: 1. glutamine synthase (GS); 2. glutamate dehydrogenase (GDH); 3. glutamate-oxoglutarate aminotransferase (GOGAT). The other enzymes represent other supply and demand enzymes for the three metabolites.
The augmented stoichiometric matrix for this system is
\[
\begin{bmatrix}
R_1 & R_2 & R_3 & R_4 & R_5 & R_6 & R_7 & Glu & Gln & 2OG
\end{bmatrix}
\]

\[
\begin{bmatrix}
Glu & -1 & 2 & -1 & -1 & 0 & 0 & 0 & 1 & 0 & 0
\end{bmatrix}
\]

\[
\begin{bmatrix}
Gln & 1 & -1 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 1
\end{bmatrix}
\]

\[
\begin{bmatrix}
2OG & 0 & -1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0
\end{bmatrix}
\]

Gaussian elimination entails adding row 1 to row 2 and the new row 2 to row 3 to give
\[
\begin{bmatrix}
R_1 & R_2 & R_3 & R_4 & R_5 & R_6 & R_7 & Glu & Gln & 2OG
\end{bmatrix}
\]

\[
\begin{bmatrix}
Glu & -1 & 2 & -1 & -1 & 0 & 0 & 0 & 1 & 0 & 0
\end{bmatrix}
\]

\[
\begin{bmatrix}
Gln & 0 & 1 & -1 & -1 & 0 & 0 & 1 & 1 & 0 & 1
\end{bmatrix}
\]

\[
\begin{bmatrix}
2OG & 0 & 0 & 0 & -1 & 1 & 1 & -1 & 0 & 0 & 1
\end{bmatrix}
\]

The first step in constructing the \( K \) matrix is to express all seven fluxes in terms of the independent fluxes:

\[
J_1 = 2J_2 - J_3 - J_4 = J_3 + J_5 + J_6 - J_7
\]

\[
J_2 = J_3 + J_4 + J_5 = J_3 + J_6 - J_7
\]

\[
J_3 = J_1
\]

\[
J_4 = -J_3 + J_6 - J_7
\]

\[
J_5 = J_1
\]

\[
J_6 = J_5
\]

\[
J_7 = J_5
\]

In vector form these equations are written as

\[
\begin{bmatrix}
J_1 \\
J_2 \\
J_3 \\
J_4 \\
J_5 \\
J_6 \\
J_7
\end{bmatrix} =
\begin{bmatrix}
1 & 0 & 1 & -1 & 1 & -1
\end{bmatrix}
\begin{bmatrix}
J_1 \\
J_2 \\
J_3 \\
J_4 \\
J_5 \\
J_6 \\
J_7
\end{bmatrix}
\]

By re-ordering the rows of \( K \) and of course the corresponding elements of the lefthand flux vector to bring the independent fluxes to the top, an identity matrix is formed at the top of the \( K \) matrix. In this case no column re-ordering is necessary (remember that if columns are reordered the elements of the independent flux-vector have to be re-ordered correspondingly):

\[
\begin{bmatrix}
J_1 \\
J_2 \\
J_3 \\
J_4 \\
J_5 \\
J_6 \\
J_7
\end{bmatrix} =
\begin{bmatrix}
1 & 0 & 1 & 0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
J_1 \\
J_2 \\
J_3 \\
J_4 \\
J_5 \\
J_6 \\
J_7
\end{bmatrix}
\]

2.5 Analysis of the basic metabolic structures

In the last part of this chapter we apply what we have learnt to six simple metabolic systems that exemplify the basic structures of metabolism (Fig. 2.5). As far as possible, they were constructed to contain no more than three metabolites and three reactions; this means that it is possible to view constraints and relationships graphically in three dimensions and so obtain insight into the meaning of the \( L \) and \( K \) matrices. In order to do this we need to visualise a 3-dimensional space with lines and planes cutting through it. For this we use the unit box or unit room representations shown in Fig. 2.6.

2.5.1 A 3-enzyme linear system (Fig. 2.5A)

The augmented stoichiometric matrix of this linear system needs no elimination as it already is in row echelon form:

\[
\begin{bmatrix}
R_1 & R_2 & R_3 & s_1 & s_2 \\
S_1 & 1 & -1 & 0 & 1 \\
S_2 & 0 & 1 & -1 & 0 & 1
\end{bmatrix}
\]
Figure 2.5: Five basic metabolic structures. A. Linear system. B. Branched system. C. 3-Member moiety-conserved cycle. D. 2-Member moiety-conserved cycle with one extra freely variable metabolite. E. Stoichiometrically constrained system. F. Anti-parallel loop (substrate loop). The fixed input and output metabolites are not shown.

There is no moiety conservation and the L-matrix is a 2-dimensional identity matrix. The stoichiometric matrix suggests $J_3$ as independent flux ($J_1$ and $J_2$ refer to columns with pivots and are dependent). However, anyone of the three fluxes could serve as independent flux as they are all equal in steady-state. Using $J_3$ as independent flux, the flux-relationships can be written as

$$\begin{bmatrix} J_1 \\ J_2 \\ J_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} J_3$$

(2.120)

The K-matrix is a vector of 1’s. This means that all attainable sets of steady-state fluxes $J_1, J_2$ and $J_3$ are multiples of the vector $[111]^T$. All possible combinations of reaction rates $v_1$, $v_2$ and $v_3$ form the 3-dimensional rate space depicted in Fig. 2.7. In linear algebra a space consisting of all possible column vectors with $n$ components is called the vector space $\mathbb{R}^n$. In this example the rate space is the familiar 3-dimensional vector space $\mathbb{R}^3$. Note that zero and negative rates are also included;

if all rates are zero, the system is in equilibrium (the state corresponding to the origin of the graph), while negative rates imply that their reactions go in the reverse direction. It is obvious that attainable steady-state flux values lie on the line through the origin and the point $(1,1,1)$. This line forms a one-dimensional vector space within $\mathbb{R}^3$ and is called a subspace of $\mathbb{R}^3$. It is, however, a very special subspace, called the nullspace or kernel, which is the set of all possible solutions for the equation $\mathbf{Nv} = 0$, and therefore the set of steady-state fluxes. Any true subspace must always contain the origin $(0,0,0)$ where all elements are zero; it is easily verified that setting all rates to zero trivially satisfies $\mathbf{Nv} = 0$; setting all rates to 1 also satisfies the equation and so do all other points in the nullspace. All that is necessary to generate the 1-dimensional nullspace is knowledge of point on the line besides the origin (in the next example the nullspace is a 2-dimensional plane in 3-dimensional rate space; here one needs two points (column vectors) to generate this plane). Such vectors that generate a vector space are said to span the space; the column vectors in $\mathbf{K}$ therefore constitute a spanning set and thus form a basis for the nullspace.

From the above it is clear that $\mathbf{K}$ plays a very important role in our analysis of steady state behaviour. It generates the subspace of all attainable steady-state flux values within the full rate space. The next example builds on the picture obtained thus far.
Figure 2.7: The 3-dimensional rate space for the 3-enzyme linear system. The line through the origin and the point (1,1,1) represents the subset of all attainable steady-state flux values, and therefore forms a 1-dimensional subspace of the full rate space. This subspace is called the nullspace or kernel of the stoichiometric matrix (see text). Note that any point in the rate space is equivalent to a column vector with three elements.

2.5.2 A 3-enzyme branched system (Fig. 2.5B)

As in the previous example, the augmented matrix for this system is already in row echelon form and needs no elimination:

\[
\begin{bmatrix}
R_1 & R_2 & R_3 & \dot{s}_1 \\
S_1 & 1 & -1 & 1 \\
\end{bmatrix}
\] (2.121)

As there is only one row, any two of the three fluxes can be chosen as independent fluxes. Let us use what the matrix suggests, namely \( J_2 \) and \( J_3 \) (the \( J_1 \)-column contains the pivot and refers to the dependent flux). The flux-relationships can now be written as:

\[
\begin{bmatrix}
J_1 \\
J_2 \\
J_3 \\
\end{bmatrix} = \begin{bmatrix} J_2 + J_3 \\ J_2 \\ J_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} J_2 \\ J_3 \end{bmatrix}
\] (2.122)

Combining the column vectors and re-ordering the rows gives the standard form of \( K \) (this re-ordering is not strictly necessary, but is neat):

\[
\begin{bmatrix}
J_2 \\
J_3 \\
J_1 \\
\end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} J_2 \\ J_3 \end{bmatrix}
\] (2.123)

As before the columns of \( K \) span the nullspace, the set of steady-state flux-values. This means that all flux-values are linear combinations of the columns of \( K \), i.e.

\[
a \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + b \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}
\]

where \( a \) and \( b \) are (scalar) constants. This forms a 2-dimensional plane in 3-dimensional rate space as depicted in Fig. 2.8. The bold arrows represent the two spanning vectors: if \( J_2 = 0 \), the attainable flux-values lie on the line represented by \((1,0,1)\); if \( J_3 = 0 \), the attainable flux-values lie on the line represented by \((1,1,0)\); obviously, if both \( J_2 \) and \( J_3 \) are zero, \( J_1 \) must also be zero.

Figure 2.8: The 3-dimensional rate space for the branched system. Linear combination of the columns of \( K \) (represented by the two bold arrows) generates a 2-dimensional nullspace of all attainable steady-state flux values for this system. Points representing vectors are given as \((J_1,J_2,J_3)\). The shaded part of the plane represents all sets of positive fluxes; the plane of course extends behind the unit box.

In the next three examples the \( K \)-matrix is exactly the same as obtained for the 3-enzyme linear system in Fig. 2.7. The steady-state flux through all three enzymes is the same and the nullspace is a line in \( R^3 \) through the origin and the point \((1,1,1)\). The systems do differ in that all three have constraints on the concentrations of some or all metabolites. Not all are moiety-conservation constraints, however; in the last example we shall discover a new type of constraint. In analysing these three examples our objective is thus to discover a meaningful interpretation of the \( L \)-matrix.
2.5.3 A 3-membered moiety-conserved cycle (Fig. 2.5C)

The augmented stoichiometric matrix and its reduction by elimination is:

\[
\begin{pmatrix}
1 & -1 & 0 & 1 & 0 & 0 \\
0 & 1 & -1 & 0 & 1 & 0 \\
-1 & 0 & 1 & 0 & 1 & 1
\end{pmatrix}
\rightarrow
\begin{pmatrix}
1 & -1 & 0 & 1 & 0 & 0 \\
0 & 1 & -1 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 & 1 & 1
\end{pmatrix}
\]

As there is only one conservation relationship and all its elements are 1, any of the three metabolites can be chosen as the dependent metabolite. If \( S_1 \) is chosen as dependent, it follows that:

\[
\begin{pmatrix}
\dot{s}_1 \\
\dot{s}_2 \\
\dot{s}_3
\end{pmatrix}
= \begin{pmatrix}
-1 \\
1 \\
0
\end{pmatrix}
\begin{pmatrix}
\dot{s}_2 \\
\dot{s}_3 \\
\dot{s}_1
\end{pmatrix}
+ \begin{pmatrix}
-1 \\
0 \\
1
\end{pmatrix}
\dot{s}_3
\]

By re-ordering the rows and combining the columns the standard form of the \( L \)-matrix is obtained:

\[
\begin{pmatrix}
1 & 0 \\
0 & 1 \\
-1 & -1
\end{pmatrix}
\begin{pmatrix}
\dot{s}_2 \\
\dot{s}_3 \\
\dot{s}_1
\end{pmatrix}
\]

Just as the columns of \( K \) span and generate a subspace within the full rate space, so do the columns of \( L \) generate a subspace within the 3-dimensional space of possible time derivative values. As there are two columns in \( L \), this subspace is a 2-dimensional plane (Fig. 2.9). The steady state (and equilibrium) corresponds to the origin where all three time derivatives are zero.

How can we interpret this subspace? If all three metabolites were independent (as were the metabolites in the linear and branched systems), \( L \) would be a 3-dimensional identity matrix and the subspace generated by the columns of \( L \) would be the original 3-dimensional time derivative space (a subspace can be identical to the full vector space). This means that in principle each time derivative can have an arbitrary value, independent of the values of the other time derivatives. Where, as in the present system, the rank of the stoichiometric matrix is less than the number of metabolites, there are constraints on the values of some or all of the time derivatives. The subspace generated by the columns of \( L \) which embody these constraints, then contains the attainable values of these constrained time derivatives. In linear algebra the columns of \( L \) would be said to generate and span the column space of the stoichiometric matrix\(^5\).

Up to now we have concentrated on the relationship of \( L \) to the time derivatives. In general, the equation for this relationship is:

\[
\frac{ds}{dt} = L \frac{ds_i}{dt}
\]

\(^5\)There are also other methods in linear algebra for generating vectors that span the column space.

\[
\frac{ds}{dt} - L \frac{ds_i}{dt} = 0
\]

This can be written as:

\[
\frac{d}{dt} (s - Ls_i) = 0
\]

or

\[
s - Ls_i = T \quad \text{or} \quad s = T + Ls_i
\]

Where \( T \) is a vector containing the conservation constant(s). This is just an expanded form of eq. 2.69, the general expression for conservation equations:

\[
s_0 - Ls_i = T \quad \text{or} \quad s_0 = T + Ls_i
\]

Note that \( T \), the vector of conservation sums, differs in the two equations, \( T \) in eq. 2.130 being padded with a zero in every position that corresponds to an independent metabolite:

\[
\begin{pmatrix}
\dot{s}_2 \\
\dot{s}_3 \\
\dot{s}_1
\end{pmatrix}
= \begin{pmatrix}
1 & 0 \\
0 & 1 \\
-1 & -1
\end{pmatrix}
\begin{pmatrix}
\dot{s}_2 \\
\dot{s}_3 \\
\dot{s}_1
\end{pmatrix}
+ \begin{pmatrix}
0 \\
0 \\
T
\end{pmatrix}
\]
If the extra column vector containing $T$ were not present, the two column vectors of $L$ would have generated the same 2-dimensional plane of attainable concentrations in the 3-dimensional concentration space. The effect of the addition of $T$ is shown in Fig. 2.10. First, note that when any two of the three concentrations equal zero, the other concentration equals $T$. These three points $(T,0,0)$, $(0,T,0)$ and $(0,0,T)$ are depicted on the graph. In terms of the two columns of $L$, the addition of $T$ has shifted the point where they coincide from the origin (see Fig. 2.9) to the point $(T,0,0)$, a point where both $s_2$ and $s_3$ are zero. This means that the subspace of attainable concentrations does not go through the origin anymore; in fact, because of this it cannot be a true subspace and is now called an affine subspace. One can also think of the planar subspace shifting up an axis which consists of points $(c,c,c)$ up to the point where $c = T/3$. Another important difference between the concentration subspace and the time derivative subspace is that, whereas time derivatives can have positive or negative values, concentrations can only be positive. The concentration subspace is therefore confined to the positive octant to form a bounded triangular plane; this plane contains all the attainable values of the three concentrations. Note that this holds not only for the steady state, but for all system states.

Figure 2.10: The concentration space for the 3-membered moiety-conserved system. Linear combination of the columns of $L$ in conjunction with $T$ generates a 2-dimensional subspace of attainable concentrations for this system. Points representing vectors are given as $(s_1, s_2, s_3)$. As concentrations must be positive, only the bounded triangular plane in the positive octant has physical significance.

### 2.5.4 A 2-membered moiety-conserved cycle (Fig. 2.5D)

This example is interesting in that, besides conserved metabolites, there is also a freely variable metabolite. The augmented stoichiometric matrix and its reduction by elimination is:

$$
\begin{bmatrix}
R_1 & R_2 & R_3 & s_1 & s_2 & s_3 \\
S_1 & 1 & -1 & 0 & 0 & 0 \\
S_2 & 0 & -1 & 1 & 0 & 0 \\
S_3 & 0 & 1 & -1 & 0 & 1
\end{bmatrix} = \begin{bmatrix}
S_1 & 1 & -1 & 0 & 0 & 0 \\
S_2 & 0 & -1 & 1 & 0 & 1 \\
S_3 & 0 & 1 & -1 & 0 & 1
\end{bmatrix}
$$

(2.133)

Again there is one conservation relationship, this time for $s_2$ and $s_3$, being independent and freely variable. If $S_3$ is chosen as dependent then it follows that

$$
\begin{bmatrix}
\dot{s}_1 \\
\dot{s}_2 \\
\dot{s}_3
\end{bmatrix} = \begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & -1 \\
0 & 1 & -1
\end{bmatrix} \begin{bmatrix}
s_1 \\
s_2 \\
s_3
\end{bmatrix}
$$

(2.134)

Here no re-ordering is necessary and the columns are combined directly to form the standard $L$-matrix:

$$
\begin{bmatrix}
\dot{s}_1 \\
\dot{s}_2 \\
\dot{s}_3
\end{bmatrix} = \begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & -1 \\
0 & 1 & -1
\end{bmatrix} \begin{bmatrix}
s_1 \\
s_2 \\
s_3
\end{bmatrix}
$$

(2.135)

Figure 2.11: The 3-dimensional time derivative space for the 2-membered moiety-conserved system. Linear combination of the columns of $L$ generates a 2-dimensional subspace of all the attainable time derivative values for this system.

The columns of $L$ again generate a plane in the 3-dimensional space of possible time derivative values (Fig. 2.11), but now the $\dot{s}_3$-vector lies on the $\dot{s}_3$-axis (it is in
fact the unit vector); therefore, the column space contains the \( \hat{s}_1 \)-axis, which shows that there is no restriction on the value of \( \dot{s}_1 \). For any specified value of \( \dot{s}_1 \) one can draw a line in the column space which runs parallel to the \((0,1,-1)\) vector; this line represents the set of \((\dot{s}_2, \dot{s}_3)\)-values that can be attained with the chosen value of \( \dot{s}_1 \).

Just as in the previous example, the addition of a vector containing the conserved sum \( T \) transforms eq. 2.135 to one showing how the dependent concentrations are related to the independent concentrations.

\[
\begin{bmatrix}
  s_1 \\
  s_2 \\
  s_3
\end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \dot{s}_1 \\ \dot{s}_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ T \end{bmatrix} \quad (2.136)
\]

The affine subspace of concentrations for this system is shown in Fig. 2.12. In comparison with Fig. 2.11 it is clear that the addition of \( T \) shifts the point where the \((1,0,0)\) and \((0,1,-1)\) vectors co-inside from the origin down the \( s_3 \)-axis to the point \((0,0,T)\). This is equivalent to shifting the planar subspace down an axis which consists of points \((0,c,c)\) to the point where \( c = T/2 \). There is no bound on the (positive) values of \( s_1 \), but the \( s_2 \) and \( s_3 \) are constrained to lie on the plane in the positive octant parallel to the \( s_1 \)-axis.

Figure 2.12: The concentration space for the 2-membered moiety-conserved system. Linear combination of the columns of \( L \) in conjunction with \( T \) generates a 2-dimensional subspace of attainable concentrations for this system.

2.5.5 A stoichiometrically constrained system (Fig. 2.5E)

This example illustrates a situation hitherto unmet. Reduction of the augmented stoichiometric matrix shows that there is a constraint on the time derivatives, and therefore the concentrations:

\[
\begin{bmatrix}
  s_1 \\
  s_2 \\
  s_3
\end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} \dot{s}_1 \\ \dot{s}_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ T \end{bmatrix} \quad (2.137)
\]

We can choose either \( s_1 \) or \( s_3 \) as dependent (they correspond to a \(+1\) in the conservation matrix). If \( s_1 \) is chosen as dependent then it follows that

\[
\begin{bmatrix}
  \dot{s}_1 \\
  \dot{s}_2 \\
  \dot{s}_3
\end{bmatrix} = \begin{bmatrix} \dot{s}_1 \\ \dot{s}_2 \\ \dot{s}_3 \\
\end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad \begin{bmatrix} \dot{s}_1 + \dot{s}_2 \\ \dot{s}_2 \\ 1 \end{bmatrix} \quad (2.138)
\]

The subspace generated by the columns of \( L \) (Fig. 2.13) differs in an interesting way from that of the previous two examples (Figs. 2.9 and 2.11).

Figure 2.13: The 3-dimensional time derivative space for the stoichiometrically constrained system. Linear combination of the columns of \( L \) generates a 2-dimensional subspace of all the attainable time derivative values for this system.

Whereas for a true moiety-conserved cycle it is impossible for all the time derivatives of the moiety-conserved metabolites to have the same sign (which makes physical sense), it is possible in the present system (for example, all three time derivatives are positive in the part of the subspace that cuts through the interior of the unit room). This already indicates that this system does not have true moiety conservation. There is, however, a clear constraint on the concentrations, which follows when we transform eq. 2.138 as before to the concentration form:

\[
\begin{bmatrix}
  s_1 \\
  s_2 \\
  s_3
\end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} \dot{s}_1 \\ \dot{s}_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ T \end{bmatrix} \quad (2.139)
\]
Again, as before, the addition of $T$ shifts the point where the $(1, 0, -1)$ and $(0, 1, 1)$ vectors meet from the origin down the $s_3$-axis to the point $(0, 0, T)$ (see Fig. 2.14). An interesting point arises when one compares the three systems with conservation constraints with respect to how the subspace of attainable concentrations changes when $T$ is reduced to zero. In Fig. 2.10 the subspace shrinks to a point (at the origin), while in Fig. 2.12 it shrinks to a line (the $s_3$-axis). In this system, however, it becomes a triangular plane in positive concentration space with one point at the origin; the plane is bounded on two sides, but unbounded in the third direction. In physical terms this can be understood as follows: for a true moiety-conserved cycle a conserved sum of zero implies that all conserved concentrations must be zero, as they are all positive. This emphasises the catalytic nature of the moiety; in its absence nothing happens. However, in this example $T = 0$ only imposes the constraint that $s_1 + s_3 = s_2$. This constraint can be thought of as due to the strict stoichiometric coupling present in this system and will therefore be called a stoichiometric constraint. This type of constraint can be seen, for example, to ensure the conservation of charge. Consider $S_1$ and $S_3$ as positively charged, and $S_2$ as negatively charged. The constraint ensures that, if the system is started with balanced charge, neutrality is maintained.

Figure 2.14: The concentration space for the stoichiometrically constrained system. Linear combination of the columns of $L$ in conjunction with $T$ generates a 2-dimensional subspace of attainable concentrations for this system.

Another way of understanding the difference between moiety-conservation constraints and stoichiometric constraints is to assign realistic chemistries to our examples as in Fig. 2.15. In the true cyclic structure (scheme A) the moiety $M$ cycles between its three forms without net production or consumption of $M$. The only net conversion is $A$ to $B$. In the stoichiometrically constrained system the moiety $M$ flows through the system and is not conserved; it is immaterial whether there is free $M$ to start off with or not, but the stoichiometric structure of the system ensures that $m$ always differs from $a + b$ by a constant $T$ (which can be zero). Here there is again net conversion from $A$ to $B$, but now while coupled to carrier $M$.

2.5.6 A system with a loop (Fig. 2.5F)

The last example is analysed not so much to illustrate the physical interpretation of $K$ and $L$ matrices (no new insights to be obtained here), but to show how the behavioural type of a metabolic structure can depend on the kinetics of the component enzymes.

The augmented stoichiometric matrix needs no elimination:

$$
\begin{array}{cccccc}
R_1 & R_2 & R_3 & R_4 & s_1 & s_2 \\
S_1 & 1 & 0 & 1 & 0 & 1 & 0 \\
S_2 & 0 & 1 & -1 & -1 & 0 & 1
\end{array}
$$

(2.140)

With $J_1$ and $J_2$ as dependent fluxes the flux-relationships are:

$$
\begin{bmatrix}
J_3 \\
J_4 \\
J_1 \\
J_2
\end{bmatrix} =
\begin{bmatrix}
1 & 0 \\
0 & 1 \\
0 & 1 \\
1 & 1
\end{bmatrix}
\begin{bmatrix}
J_5 \\
J_6
\end{bmatrix},
$$

(2.141)

The attainable steady-state fluxes therefore lie on a 2-dimensional plane in 4-dimensional rate space; this cannot be visualised graphically in the manner used for the other systems, but nevertheless the interpretation remains the same. This system also confirms that although there is a type of cyclical structure (which we prefer to call a loop) in the network, this is not a moiety-conservation cycle. However, and this
is the interesting point, there is potentially a moiety-conservation cycle in this network. If the input and output fluxes $J_1$ and $J_4$ are negligibly small compared to the loop fluxes $J_2$ and $J_3$, the system behaves as a 2-membered moiety-conserved cycle, i.e., the augmented stoichiometric matrix before and after elimination reduces to:

$$
\begin{bmatrix}
R_2 & R_3 \\
S_1 & -1 & 1 & 0 \\
S_2 & 1 & -1 & 0
\end{bmatrix} \rightarrow
\begin{bmatrix}
R_2 & R_3 & \delta_1 & \delta_2 \\
S_1 & -1 & 1 & 0 \\
S_2 & 0 & 0 & 1
\end{bmatrix}
$$

(2.142)

This illustrates a more realistic physiological scenario where the moiety in a moiety-conserved cycle is synthesized and degraded \textit{de novo} at much lower rates than that of interconversion of the different forms of the moiety. As the rates of synthesis and degradation become comparable to the interconversion rate the cycle loses its moiety-conserved nature and becomes a metabolic loop.

As in this example, most metabolic systems have more than three reactions and three metabolites, but nevertheless the insight gained by analysing these simple systems and visualising the relationships and constraints can easily be extended to systems of higher dimensionality. Such $n$-dimensional hyperspaces and the hyperplanes that cut through them are difficult, usually impossible, to visualise, but the physical interpretation of the $K$ and $L$ matrices that generate them remain unchanged.

### 3 Metabolic Control Analysis

The aim of metabolic control analysis is twofold: First, we want to understand how sensitive the steady-state variables (the fluxes and concentrations) are to variations in the parameters. As these sensitivities depend on all the interactions within a metabolic system, we call them \textit{global} or \textit{systemic} properties. Second, we want to relate these global sensitivities to local properties of the reactions (enzymes) themselves. How these local and global properties are expressed remains to be seen.

The reason that this type of analysis is called 'metabolic control analysis' arises from a specific meaning that we attach to the concept of \textit{control}. The term 'metabolic control' is often used rather loosely and qualitatively, and we need to define it in a precise and quantitative way. Generally, 'to control' is understood to mean 'to be able to influence something'. When we say that an enzyme controls the flux through a pathway, we mean that a change in the activity of the enzyme results in a change of the flux. Unfortunately, the word 'control' seems to have acquired an 'all or none' connotation: something exerts either complete control or no control at all. This qualitative meaning has been taken over in the concept of a 'rate-limiting' enzyme, which would (completely) control the flux. We shall see that, on the contrary, control of flux is often shared by all the enzymes in a metabolic system; in fact, control of any steady-state variable is shared by all the enzymes. From the following treatment it will become clear exactly how control can be defined quantitatively in this context. Remember that we made a careful distinction between parameters and variables in metabolic systems. This is important, as the parameters must be regarded as the 'controllers', the variables as the 'controlled'.

The aim of this chapter is to give more insight into what control analysis tells us about metabolic systems, and also to show the valuable insight that can be gained about metabolic behaviour by performing 'thought experiments'.

### 3.1 Quantifying metabolic control

As we have seen in Chapter 1 the relationship between the steady-state variables (fluxes and concentrations) and parameters that affect the activities of the enzymes in a system is non-linear and cannot be obtained analytically, except in the simplest cases. At present, no theory will therefore be able to predict quantitatively what the change in, say, flux would be for a large change in enzyme activity (unless we make the simplifying and dangerous assumption that flux responds linearly or logarithmically to enzyme concentration). In order to arrive at a mathematically
We start off our exploration of the world of control analysis by focusing on the functional steps of the system where we regard each functional step as being catalysed by an enzyme, but this is not strictly necessary; any reaction or transport step can be treated in this way. As we have seen in our kinetic models, each enzyme is directly connected to specific metabolites, i.e., its substrates, products and effectors (which can be either fixed external inhibitors or activators, or variable internal factors which can be either fixed external inhibitors or activators, or variable internal factors which can be altered by the system, e.g., pH, ionic strength, protein concentration, etc.). First, it should be clear that if we unfreeze the enzyme, while keeping all the other metabolites constant at their prevailing concentrations, we would obtain a curve around the original initial rate of the reaction. If we were to plot the apparent kinetic order against other concentrations such as the substrate $s$, we would obtain a curve around the original initial rate of the reaction. If we were to plot the apparent kinetic order against other concentrations such as the substrate $s$, we would obtain a curve around the original initial rate of the reaction. If we were to plot the apparent kinetic order against other concentrations such as the substrate $s$, we would obtain a curve around the original initial rate of the reaction. If we were to plot the apparent kinetic order against other concentrations such as the substrate $s$, we would obtain a curve around the original initial rate of the reaction. If we were to plot the apparent kinetic order against other concentrations such as the substrate $s$, we would obtain a curve around the original initial rate of the reaction. If we were to plot the apparent kinetic order against other concentrations such as the substrate $s$, we would obtain a curve around the original initial rate of the reaction. If we were to plot the apparent kinetic order against other concentrations such as the substrate $s$, we would obtain a curve around the original initial rate of the reaction. If we were to plot the apparent kinetic order against other concentrations such as the substrate $s$, we would obtain a curve around the original initial rate of the reaction.

Mathematically, we can express elasticity coefficients as:

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

This means that we can also define the elasticity coefficient as:

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

For those who like a more formal mathematical definition, we must consider the limit where the changes become infinitesimally small, i.e., when $\delta v \rightarrow 0$. Then the elasticity coefficient of a rate $v$ at a specified value $s$, that directly interacts with $v$, is equivalently defined as:

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

Elasticity coefficients can be described in five ways:

1. The ratio of the fractional change $\delta v/v$ of the local rate of enzyme and the fractional change $\delta s/s$ of the concentration of a metabolite that interacts directly with the enzyme at a specified value $s$ of $s$. This description implies that an elasticity coefficient is a variable enzymological property; like all other enzymic properties ($K_M$, $V_{max}$, etc.) its value is constant for a prescribed set of conditions but will vary as conditions vary.

2. Multiplying both fractional changes by 100 will give the percentage change in $s$ and $v$, respectively. We call this scaled slope the elasticity coefficient:

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

3. The product of $\varepsilon$ and the slope of the tangent $\delta v/\delta s$ at a specified value of $s$.

4. An even more elegant way of expressing elasticity coefficients takes into account the number of other metabolites that interact with the enzyme and with each other. We can now also define the elasticity coefficient as:

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

We now answer the following question: How can we describe the sensitivity of the local rate through an enzyme to small changes in these directly-connected metabolites? Note that, to answer this question, we must regard each enzyme separately, even though they act in concert and that their kinetic models are interdependent. To simplify the discussion, we regard the enzyme that catalyses reaction as if it were isolated (in fact, although we answer this question by conducting a thought experiment, it could be studied experimentally with the isolated enzyme in a test tube, using the well-known procedures of steady-state enzymology).
The $s_k$ and $s_l$ outside the brackets indicate that any other metabolites that interact with the enzyme are kept constant at the prevailing steady-state values; that is why we define an elasticity coefficient as a partial derivative.

Note that elasticity coefficients describe the properties of the 'isolated' functional units in the system and, individually, are uninformative about system behaviour. Nevertheless we shall see that they enter in important ways into control analysis.

Above we considered the elasticity coefficient of an enzyme with respect to a substrate. An enzyme will of course have elasticity coefficients with respect to all other metabolites that interact with it directly (i.e. that appear in its rate equation). For example, keeping the substrate constant, we can now vary the product concentration around its steady-state value to obtain:

$$
\varepsilon_v^p = \frac{\delta v / \nu}{\delta [p]/p} > 0
$$

It should also be clear that there are other entities that directly affect the rate of an enzyme reaction, e.g., the enzyme concentration $e$, so that

$$
\varepsilon_v^e = \frac{\delta v / \nu}{\delta e/e}
$$

When the rate equation for the reaction is known, an explicit expression for an elasticity coefficient can be obtained as the scaled partial derivative of the rate function towards an interacting metabolite. For instance, the net-rate equation for the reaction $S \leftrightarrow P$ could be that of a reversible Michaelis-Menten mechanism:

$$
\nu = \frac{V_f}{K_c} \left( \frac{s - p}{K_e} \right) \left( 1 + \frac{s}{K_s} + \frac{p}{K_p} \right)
$$

where $V_f$ denotes the limiting forward velocity, $K_e$ the equilibrium constant and $K_s$ and $K_p$ the respective Michaelis constants for $S$ and $P$. Partial differentiation with respect to $s$ or $p$ and scaling of the partial derivative with $s/\nu$ and $p/\nu$ respectively gives the elasticities towards $s$ and $p$:

$$
\varepsilon_v^s = \frac{1}{1 - \frac{s}{K_e}} \left( 1 + \frac{s}{K_s} + \frac{p}{K_p} \right)
$$

$$
\varepsilon_v^p = \frac{1}{1 - \frac{p}{K_p}} \left( 1 + \frac{s}{K_s} + \frac{p}{K_p} \right)
$$

where $\Gamma$ is the mass-action ratio $p/s$.

An enzymic rate equation is usually a proportional to enzyme concentration $e$, i.e., $e$ is a multiplier of some function of $s$, $p$, etc. This can be seen in the above reversible Michaelis-Menten rate equation if one remembers that $V_f = k_{cat} e$. If this is so then partial differentiation with respect to $e$ and scaling with $e/\nu$ gives

$$
\varepsilon_v^e = 1
$$

### 3.1.2 Global coefficients: Response and control coefficients

Let us now perform another thought experiment, this time not with an isolated enzyme, but with the intact system in steady state. We already know that this steady state is completely determined by the parameters of the system. Changing any parameter will affect the steady-state to some degree. However, the parameters of interest in metabolic studies, i.e., pathway substrates and products, external inhibitors and activitors, all exert their effects through specific steps in the pathways, usually acting on the enzymes that catalyse these steps. Enzyme concentrations are
of course also very important and they certainly affect the steady-state through the steps that they catalyse.

In the light of this we now ask a fundamental question: **How can we describe the sensitivity of a steady-state variable to a perturbation of the local rate (activity) through a specific step caused by a change in a parameter that affects this step directly?** To make things more real we shall consider as the steady-state variable a flux $J$ and as the parameter an external effector $X$ with concentration $x$ that acts only on the enzyme that catalyses step $i$. Our argument will hold, however, for any other steady-state variable, e.g. an internal concentration, and any other parameter that only affects step $i$.

Now we perform the thought experiment to answer our question. First we make a small change $\delta x$ in the parameter $X$. What happens? Well, a whole chain of things happen. The change in $x$ causes a change in the rate of step $i$, which in turn affects the concentrations of its substrates and products, which then affect the other rates with which these metabolites interact, and so the effect of the perturbation in $x$ spreads out via changes in the metabolites that link the reactions in the network. The initial change in the rate of step $i$ therefore reverberates through the whole system which goes into a transient state. Eventually though, the system will again settle into a steady state, but this steady state will differ from the original steady state because one of the parameters has changed. Therefore we should be able to measure a difference $\delta J$ between the old and new steady-state flux-values. Using fractional changes as before (and for the same reasons), we can summarise by saying that a fractional parameter change $\delta x/x$ caused a fractional steady-state flux change $\delta J/J$. Therefore, although the parameter acted locally on step $i$, the change we observe is systemic in the sense that the properties of all steps in the system contributed to the observed change in the steady state.

The ratio of these fractional changes we call a **response coefficient**, equivalently defined as:

$$ R_J^i = \frac{\delta J/J}{\delta x/x} = \frac{\delta J}{\delta x} \cdot \frac{x}{J} = \frac{\delta \ln J}{\delta \ln x} $$

(3.10)

Note, that although the type of mathematical function in this definition is the same as that in the definition of an elasticity coefficient, the two types of coefficient are conceptually and operationally quite different, elasticity coefficients being local (isolated) enzymic properties and response coefficients being systemic (global) properties.

We have, however, still not answered our original question. We now know how to describe the response of a steady-state variable to a change in a parameter, but we actually want to describe the response of a steady-state variable to a change in the rate through a step. Well, seeing that the change in $x$ initially affected only step $i$ before anything else happened, we could ask by how much $v_i$ changed. This we can already answer: The effect of $\delta x$ on the local rate $v_i$ seen in isolation is described by the elasticity coefficient

$$ \varepsilon_{v_i} = \frac{\delta v_i/v_i}{\delta x/x} $$

(3.11)

so that

$$ \frac{\delta v_i}{v_i} = \varepsilon_{v_i} \frac{\delta x}{x} $$

(3.12)

Now, if we take the ratio of the systemic effect of $\delta x/x$ on $J$ and the local effect of $\delta x/x$ on $v_i$ we get

$$ \left( \frac{\delta J/J}{\delta x/x} \right) \cdot \left( \frac{\delta v_i/v_i}{\delta x/x} \right) = \frac{R_J^i}{\varepsilon_{v_i}} $$

(3.13)

As the value of $\delta x/x$ is the same in both the numerator and the denominator it cancels from the expression, leaving

$$ \frac{\delta J/J}{\delta v_i/v_i} = \frac{R_J^i}{\varepsilon_{v_i}} $$

(3.14)

It should be clear that this ratio is the key to the answer to our original question: **How can we describe the sensitivity of a steady-state variable to a perturbation of the local rate (activity) through a specific step caused by a change in a parameter?**
that affects this step directly? It is this ratio that we call the control coefficient of step $i$, more specifically the flux-control coefficient of step $i$:

$$
C_i^p = \frac{\delta J / J}{\delta v_i / v_i}
$$

(3.15)

In general, therefore, the control coefficient of steady-state variable $y$ with respect to the rate of step $i$ is equivalently defined as

$$
C_i^y = \frac{\delta y / y}{\delta v_i / v_i} = \frac{\delta y}{\delta v_i} \frac{v_i}{y} = \frac{\delta \ln y}{\delta \ln v_i}
$$

(3.16)

Note that we usually just use a subscript $i$ in the symbol of a concentration control coefficient. $v$ refers to a steady-state concentration, and the control coefficient is called a concentration-control coefficient.

A very important general relationship arises from eq. 3.14:

$$
X / E \rightarrow X projecting from $X$ to $E$
$$

(3.17)

where $x$ is any parameter that acts on step $i$. A response coefficient is therefore always a product of a control coefficient and an elasticity coefficient. This last relationship is called the combined response relationship.

In most of the early papers on control analysis (and even in some recent ones) a control coefficient is defined not in terms of $v$, but in terms of enzyme concentration $e$, eq. 3.19 shows that in mny cases the two definitions are equivalent. However, most workers in the field now accept the more general definition in eq. 3.16. However, although accepting the conceptual difference between $R_i^y$ and $C_i^y$, many papers still call both of these entities control coefficients. This is understandable in terms of the nomenclature recommendations made in 1985 [TIBS 10, 16], but it is our hope that response and control coefficients will become formally accepted as distinctive concepts. Calling two different things by the same name can only cause confusion.

$$
R_i^p = \sum_{i=1}^{n} C_i^p \epsilon_i^{e_p}
$$

(3.20)

### 3.2 Control Properties

Now that we know what control, response and elasticity coefficients are, we come to the bit that makes metabolic control analysis so powerful: the relationships between the various types of coefficients. We have already discovered the combined response property that expresses a response coefficient as a product of a control coefficient and an elasticity coefficient. Here we discover the summation and connectivity properties by means of thought experiments.

We shall explore the relationships between the various coefficients of control analysis by analysing a specific metabolic system, shown in Fig. 3.3, which is in a stable steady state.

Figure 3.3: A 3-enzyme linear system with a feedback loop from $S_2$ onto $E_1$, and an external effector, $X_e$, that interacts with $E_1$.

### 3.2.1 Summation properties of control coefficients

Thought experiment: What would happen if we simultaneously made the same fractional change, $\alpha$, in the local rates of all the steps in the system?, i.e. if

$$
\frac{\delta v_1}{v_1} = \frac{\delta v_2}{v_2} = \frac{\delta v_3}{v_3} = \alpha
$$

(3.21)

Answer: The flux, $J$, must increase fractionally by $\alpha$, but, since all the rates increased in the same proportion, the concentrations of the variable metabolites $s_1$ and $s_2$ remain unchanged.

The combined effect of all the changes in local rates on the systemic variables $J$, $s_1$ and $s_2$ can be written as the sum of the individual effects caused by each change in local rate. For the flux $J$:

$$
\frac{\delta J}{J} = C_1^j \frac{\delta v_1}{v_1} + C_2^j \frac{\delta v_2}{v_2} + C_3^j \frac{\delta v_3}{v_3}
$$

(3.22)

$$
\alpha = \alpha(C_1^j + C_2^j + C_3^j)
$$

(3.23)

Therefore

$$
C_1^j + C_2^j + C_3^j = 1
$$

(3.24)
Similarly, for $s_1$

\[
\frac{\delta s_1}{s_1} = C_1^{v_1} \delta v_1 + C_2^{v_2} \delta v_2 + C_3^{v_3} \delta v_3
\]

\[
0 = \alpha (C_1^{v_1} + C_2^{v_2} + C_3^{v_3})
\]

Therefore

\[
C_1^{v_1} + C_2^{v_2} + C_3^{v_3} = 0
\]

Similarly, for $s_2$

\[
C_1^{v_2} + C_2^{v_2} + C_3^{v_2} = 0
\]

These two properties hold in general. The first, namely that for any steady-state flux in the system $J_m$ the $J_m$-control coefficients of all the steps of the metabolic system sum to one, is called the summation property of flux-control coefficients and can be generalized as:

\[
\sum_{i=1}^{n} C_i^{v_m} = 1
\]

where $n$ denotes the number of enzymes in the system.

The important conclusion to be reached from the flux-summation property is that enzymes share the control of a flux.

The second property, namely that with respect to any steady-state concentration $s_j$, the $s_j$-control coefficients of all the steps of the metabolic system sum to zero, is called the summation property of concentration-control coefficients and can be generalized as:

\[
\sum_{i=1}^{n} C_i^{v_j} = 0
\]

From the summation to zero it should be clear that whereas some enzymes act to decrease metabolite concentrations, others increase them.

These summation equations show how global control properties of metabolic pathways are related. The summation property for flux-control coefficients played an important role in dispelling the established notion that flux-control resides in one so-called ‘rate-limiting’ enzyme. On the contrary, as this summation property shows, flux-control can be shared by all enzymes in the system. Furthermore, as the conditions change the distribution of control between the enzymes of the system will vary within the confines of the summation property.

### 3.2.2 Connectivity properties

The combined response property describes the relationship between local and systemic coefficients for external metabolites. Are there similar relationships between control coefficients and elasticities towards variable metabolites?

**Thought experiment:** How can we bring about a small change in the concentration of a variable metabolite concentration, say $\delta s_j$, without changing the flux $J$ or any of the other metabolite concentrations (here $s_j$)?

**Answer:** The immediate effect of the change $\delta s_j$ is to change the local flux rates of all the reactions to which it is directly connected, in this case reaction 3 (for which it is a substrate), reaction 2 (for which it is a product), and reaction 1 (for which it is an effector). If we could cancel these local rate changes by making compensating changes in some parameter, such as enzyme concentration, of each of the directly affected reactions, then the flux through each of these rates would remain constant, thereby ensuring that no other steady-state concentrations in the system change. So, we make these compensating rate changes by $\delta e_1$, $\delta e_2$, and $\delta e_3$ (any three parameters each acting specifically on one reaction will do, but enzyme concentration is the most convenient).

That $\delta e_1$ compensates for the effect of $\delta s_j$ on $v_1$ is expressed by

\[
\frac{\delta v_1}{v_1} = \frac{e^{v_1} \delta e_1}{e_1} + \frac{e^{v_2} \delta s_j}{s_2} = 0
\]

Similarly,

\[
\frac{\delta v_2}{v_2} = \frac{e^{v_2} \delta e_2}{e_2} + \frac{e^{v_3} \delta s_j}{s_3} = 0
\]

\[
\frac{\delta v_3}{v_3} = \frac{e^{v_3} \delta e_3}{e_3} + \frac{e^{v_3} \delta s_j}{s_3} = 0
\]

It follows that

\[
e^{v_1} \delta e_1 = - \frac{e^{v_1} \delta s_j}{s_2} \quad \text{or} \quad \frac{\delta e_1}{e_1} = - \frac{e^{v_1} \delta s_j}{s_2}
\]

\[
e^{v_2} \delta e_2 = - \frac{e^{v_2} \delta s_j}{s_2} \quad \text{or} \quad \frac{\delta e_2}{e_2} = - \frac{e^{v_2} \delta s_j}{s_2}
\]

\[
e^{v_3} \delta e_3 = - \frac{e^{v_3} \delta s_j}{s_3} \quad \text{or} \quad \frac{\delta e_3}{e_3} = - \frac{e^{v_3} \delta s_j}{s_3}
\]

Although the system is in a new steady state (some parameters have changed), only the value of $s_2$ is different; $J$ and $s_1$ are still the same. Therefore

\[
\frac{\delta J}{J} = R_{e_1}^J \frac{\delta e_1}{e_1} + R_{e_2}^J \frac{\delta e_2}{e_2} + R_{e_3}^J \frac{\delta e_3}{e_3} = 0
\]

Substituting the above expressions for $\delta e_i/e_i$ yields

\[
-R_{e_1}^J \frac{e^{v_1} \delta s_j}{s_2} - R_{e_2}^J \frac{e^{v_2} \delta s_j}{s_2} - R_{e_3}^J \frac{e^{v_3} \delta s_j}{s_2} = 0
\]
From eq. 3.16 we recognise the response coefficient divided by the elasticity coefficient as a control coefficient, so that

\[-C_J^i e^i_{j2} \frac{\delta s^i_{j2}}{s^i_{j2}} - C_J^i e^i_{j1} \frac{\delta s^i_{j1}}{s^i_{j1}} - C_J^i e^i_{j3} \frac{\delta s^i_{j3}}{s^i_{j3}} = 0 \]  

(3.39)

Dividing by \(-\delta s^i_{j2}/s^i_{j2}\) yields

\[C_J^i e^i_{j2} + C_J^i e^i_{j1} + C_J^i e^i_{j3} = 0 \]  

(3.40)

Similarly, as \(s^i_j\) did not change, it follows that

\[\frac{\delta s^i_j}{s^i_j} = \frac{R^i_1 \delta e^i_1}{e^i_1} + \frac{R^i_2 \delta e^i_2}{e^i_2} + \frac{R^i_3 \delta e^i_3}{e^i_3} = 0 \]  

(3.41)

so that, by the same manipulations as above

\[C_J^i e^i_{j2} + C_J^i e^i_{j1} + C_J^i e^i_{j3} = 0 \]  

(3.42)

However, \(\delta s^i_{j2}/s^i_{j2} \neq 0\), so that

\[\frac{\delta s^i_{j2}}{s^i_{j2}} = \frac{R^i_1 \delta e^i_1}{e^i_1} + \frac{R^i_2 \delta e^i_2}{e^i_2} + \frac{R^i_3 \delta e^i_3}{e^i_3} \]  

(3.43)

Substitution as above yields

\[C_J^i e^i_{j2} + C_J^i e^i_{j1} + C_J^i e^i_{j3} = -1 \]  

(3.44)

The same thought experiment done for \(S_1\) would yield

\[C_J^i e^i_{j1} + C_J^i e^i_{j2} = 0 \]  

(3.45)

\[C_J^i e^i_{j1} + C_J^i e^i_{j2} = -1 \]  

(3.46)

\[C_J^i e^i_{j1} + C_J^i e^i_{j2} = 0 \]  

(3.47)

Note that since \(S_1\) does not interact with \(E_3\) in any way, there is no term that contains \(e^i_{j1}\) in any of these three expressions.

These so-called connectivity properties can be generalized as the

- **Connectivity between flux-control coefficients and elasticities**

\[\sum_{i=1}^n C_J^i e^i_{j} = 0 \]  

(3.48)

where \(S_j\) is a variable metabolite pool and \(f_m\) a specific system flux.

- **Connectivity between concentration-control coefficients and elasticities**

\[\sum_{i=1}^n C_J^i e^i_{j} = -\delta_{jk} \]  

(3.49)

where \(\delta_{jk}\) is the Kronecker delta (\(\delta_{jk} = 1\) when \(j = k\); \(\delta_{jk} = 0\) when \(j \neq k\)).

---

4 The Algebra of Metabolic Control Analysis

4.1 Introduction

Like the evolution of life, the development of metabolic control analysis can be likened to a process of tinkering\(^1\). What now stands as the theoretical body of control analysis is the result of a piecemeal addition to and refinement of theorems presented in the original papers of Kacser and Burns [40, 41] and Heinrich and Rapoport [22]. A landmark paper on the formalisation of control analysis is by Reder [54], although others have provided some formal description from first principles [5, 6, 17, 18, 19, 72]. At present the most complete formalisation can be found in [24]. One may therefore rightly question the need for another treatment. However, re-inventing the metaphorical wheel often yields new insights, and it is in this spirit that this paper is offered as a journey of discovery through the algebraical landscape of metabolic control analysis.

We start at the very beginning with the general kinetic model for a network of chemical reactions, and then proceed step by step, doing our best to avoid those unexplained jumps which, although seemingly obvious to experienced mathematicians, leaves us lesser mortals feeling woefully inadequate. Nevertheless, the reader is at least assumed to be acquainted with introductory differential calculus. The great strength and elegance of symbolic matrix algebra is utilised throughout, but there is nothing mysterious about it. ... matrices can only be multiplied if the number of columns of the first matrix equals the number of rows of the second (an \(m \times n\) matrix can only be multiplied by a \(k \times n\) matrix; the product will have dimensions \(m \times k\)), and (ii) matrix multiplication is not commutative, i.e., the product \(AB\) is usually not equal to \(BA\). For those that feel more comfortable with explicit matrix equations, the example provided in Appendix 4.7 serves as a starting point.

4.2 The kinetic model

The kinetic model for any (metabolic) network of coupled chemical reactions and transport processes can be written as a set of nonlinear differential equations (see e.g., [54]):

\[
\frac{ds}{dt} = Nv[s, p] \tag{4.1}
\]

where, for a system of \(n\) coupled reactions that inter-convert \(m\) substances (from here on called ‘metabolites’), \(s\) is an \(m\)-dimensional column vector of metabolite concentrations, \(N\) is an \(m \times n\)-dimensional matrix of stoichiometric coefficients (the stoichiometric matrix), \(v\) is an \(n\)-dimensional column vector of reaction rates, and \(p\) is a \(p\)-dimensional column vector of parameters. Only variable metabolite concentrations are included in \(s\); metabolite concentrations which are buffered externally and can therefore be regarded as constant are considered to be included in the parameter vector \(p\).

In any systemic state the reaction rates \(v\) are functions of both metabolite concentrations \(s\) and parameters \(p\) such as kinetic constants and fixed external concentrations. This is expressed in eqn. 4.1 by the functional relationship \(v = v[s, p]\).

The structure or topology of the reaction network is embodied in the stoichiometric matrix \(N\). An element \(c_{ij}\) of \(N\) is the stoichiometry, usually an integer, with which metabolite \(S_i\) participates in reaction \(j\) (if \(S_i\) is a reactant, \(c_{ij} < 0\); if a product, \(c_{ij} > 0\); otherwise, \(c_{ij} = 0\)). Two invariant properties can be extracted from \(N\), namely (i) the conservation relationships that arise when the differential equations are not all linearly independent, and (ii) the steady-state flux relationships [54]. Here we discuss the first; the second will follow once the steady state has been treated.

By Gaussian elimination to row echelon form (see, e.g., [67]) we can determine whether the rows of \(N\) (and, therefore, the differential equations themselves) are linearly independent (see Appendix 4.7 for an example). If they are independent then \(r = m\), where \(r\) is the rank of \(N\) (the number of independent equations). If \(r < m\) then there are \(m - r\) dependencies among the differential equations. Eliminating \(m - r\) dependent rows of \(N\) leaves a reduced stoichiometric matrix, \(N_r\), with \(r\) independent rows. \(N\) and \(N_r\) can be related by constructing a link matrix \(L\) with dimensions \(m \times r\) so that \(N = LN_r\) [54]. If \(N\) is re-arranged so that the independent rows come first, then \(L\) and the concentration vector \(s\) that corresponds to the rows of the re-arranged \(N\) have the structure

\[
L = \begin{bmatrix} L_r & L_o \end{bmatrix} \quad \text{and} \quad s = \begin{bmatrix} s_i \cr s_d \end{bmatrix} \tag{4.2}
\]

where \(L_r\) is an \(r\)-dimensional identity matrix and \(L_o\) an \((m - r) \times r\)-dimensional matrix that expresses the dependent time derivatives in terms of the independent time derivatives (see eqn. 4.6 below); \(s_i\) refers to independent and \(s_d\) to dependent concentrations.

Using these relationships the kinetic model in eqn. 4.1 can be written as

\[
\frac{ds}{dt} = LN_r v[s_i, s_d, p] \tag{4.3}
\]

where the functional relationship \(v = v[s_i, s_d, p]\) emphasises the fact that the dependencies among the differential equations allows the partitioning of \(s\) into \(r\) independent concentrations \(s_i\) and \(m - r\) dependent concentrations \(s_d\). Now the kinetic model can be expanded into

\[
\frac{d}{dt} \begin{bmatrix} s_i \\ s_d \end{bmatrix} = LN_r v[s_i, s_d, p] = \begin{bmatrix} L_r & L_o \end{bmatrix} N_r v[s_i, s_d, p] \tag{4.4}
\]

which can be split into two equations:

\[
\frac{d}{dt} s_i = N_r v[s_i, s_d, p] \tag{4.5}
\]

\[
\frac{d}{dt} s_d = L_o N_r v[s_i, s_d, p] = L_o \frac{ds_i}{dt} \tag{4.6}
\]

which, when combined, give

\[
\frac{ds}{dt} = L \frac{ds_i}{dt} \tag{4.7}
\]

It is clear that if \(L_o\) is known we need only consider the kinetics as expressed by eqn. 4.5, as eqn. 4.6 allows the expression of the linear dependencies between the rates of change of metabolite concentrations:

\[
\frac{d}{dt}(s_d - L_o s_i) = 0 \tag{4.8}
\]

where \(0\) is a null vector (a vector of zeros). This implies that

\[
s_d = L_o s_i + T \tag{4.9}
\]

where \(T\) is an \((m - r)\)-dimensional vector of constant (conserved) sums of concentrations. The full concentration vector \(s\) can therefore be expressed as a function of \(s_i\) and \(T\):

\[
s = \begin{bmatrix} s_i \\ s_d \end{bmatrix} = \begin{bmatrix} L_r & L_o \end{bmatrix} s_i + \begin{bmatrix} 0 \\ T \end{bmatrix} = L s_i + \begin{bmatrix} 0 \\ T \end{bmatrix} \tag{4.10}
\]

where \(0\) represents an \(r\)-dimensional subvector of zeros.

4.2.1 Functional relationships in the steady state

In the steady state the kinetic model \(ds/dt = 0\), and the equations simplify to a system of non-linear equations of the form

\[
N_r v[s_i, s_d, p] = 0 \tag{4.11}
\]
When there are no conservation relationships (when \( r = m \)), the equation system reduces to a slightly simpler form

\[
N v(s, p) = 0
\]

(4.12)

but we shall only consider eqn. 4.11 as it is more general. The solution to eqn. 4.11 is a vector of independent concentrations

\[
s_i = s_i[T, p]
\]

(4.13)

Note that the concentrations are now steady-state concentrations (as the context is clear when one refers to steady-state concentrations). Furthermore, the solution can be expanded to a vector of dependent steady-state concentrations, which depends through eqn. 4.9 on \( s_i \) and \( T \),

\[
s_d = s_d[s_i, T]
\]

(4.14)

and a steady-state reaction rate vector

\[
J = v[s_i, s_d, p]
\]

(4.15)

for which we reserve the special name flux vector.

Usually, we cannot solve for the steady-state concentrations and fluxes analytically, although the powerful symbolic manipulators available today (e.g., Mathematica, Maple, Reduce) enlarge the scope of what is possible. Except for the simplest cases, analytical solutions are in any case extremely difficult to interpret. The central question asked by metabolic control analysis is how the steady-state variables change when the steady-state changes in response to a perturbation in one or more parameters. In order to answer this question it is necessary to differentiate the steady-state equations with respect to the parameters, and for this we must have an accurate picture of the functional relationships in these equations. A diagrammatic representation makes these nested functional relationships more transparent:

\[
\begin{align*}
& \text{v} \\
& \quad \text{s_i} \\
& \quad \quad \quad \text{s_d} \\
& \quad \quad \quad \quad \quad \text{p} \\
& \quad \quad \quad \quad \quad \text{T} \\
& \quad \text{J} \\
& \quad \quad \text{T}
\end{align*}
\]

It is clear that the steady-state concentrations \( s_i \) and \( s_d \) as well as the steady-state fluxes \( J \) ultimately depend only on the parameters \( p \) and the conservation sums \( T \). Nevertheless, the intermediary levels of functional dependencies are important when the steady-state equations are differentiated with respect to \( p \) and \( T \). However, before we turn to this topic we complete our structural analysis of \( N \) by considering the relationships that exist between fluxes in the steady state.

\subsection{Flux-relationships in the steady state}

We showed above that linear dependencies among the rows of \( N \) can be captured in the link matrix \( L \). Similarly, in the steady state when \( N v = 0 \) (or \( N v = 0 \), if there are no conservation relationships) their exist dependencies among the columns of \( N \) (or \( N v \)) that can be expressed as

\[
NK = 0 \quad \text{or} \quad Nv = 0
\]

(4.16)

where \( K \) is the kernel (or nullspace) of \( N \) \[^{54}\]. Each column of \( K \) is a particular solution to eqn. 4.16, and the set of columns is linearly independent and therefore span the nullspace. Because each column represents an independent flux, it follows that

\[
J = KJ_i
\]

(4.17)

where \( J \) is an \( n \)-dimensional column vector of all the steady-state fluxes, and \( J_i \) is an \( (n - r) \)-dimensional column vector of independent fluxes (recall that \( r \) is the rank of the stoichiometric matrix). \( K \) therefore has dimensions \( n \times (n - r) \). If \( K \) is re-arranged so that the \( n - r \) rows that correspond to independent fluxes come first, then the flux vector \( J \) is partitioned into \( n - r \) independent fluxes \( J_i \) and \( r \) dependent fluxes \( J_d \), and eqn. 4.17 becomes

\[
\begin{bmatrix}
J_1 \\
J_d
\end{bmatrix} =
\begin{bmatrix}
I_{n-r} \\
K_0
\end{bmatrix}
J_i
\]

(4.18)

where \( I \) is an \((n - r)\)-dimensional identity matrix and \( K_0 \) an \( r \times (n - r) \)-dimensional matrix that expresses the dependent fluxes in terms of the independent fluxes, \( J_d = K_0J_i \).

\subsection{Differentiation of the steady-state equation}

Because the parameters \( p \) and \( T \) determine the steady-state, any change in these parameters can potentially change the steady state. If the parameter perturbation is small enough, the change from steady state \((s_i, J_i)\) to steady state \((s_d, J_d)\) is approximated by

\[
s_2 = s_1 + \frac{\partial s}{\partial p}(p_2 - p_1)
\]

(4.19)

and

\[
J_2 = J_1 + \frac{\partial J}{\partial p}(p_2 - p_1)
\]

(4.20)

the first two terms in the so-called multivariate Taylor expansion. Similar equations for perturbations in \( T \) can be obtained by replacing \( p \) by \( T \). The matrices of partial derivatives \( \frac{\partial s}{\partial p}, \frac{\partial J}{\partial p}, \frac{\partial s}{\partial T}, \text{ and } \frac{\partial J}{\partial T} \) are thus of great interest and will be obtained next.
Readers will be familiar with the differentiation of an explicit mathematical equation of the form \( y = f(x) \) with respect to \( x \), or if, as is the case here, \( y \) is a function of more than one variable \( y = f(x, z) \), partial differentiation with respect to either \( x \) or \( z \) or both at the same time. However, we are working not with simple scalar variables, but with vector variables. In addition, equations such as eqn. 4.11 are implicit functions. Fortunately, the extension to vector variables and implicit functions is not difficult at all. Let us first get a feel for this process by differentiating eqn. 4.14, the explicit function \( \mathbf{s}_d = \mathbf{L}_0 \mathbf{s}_i + \mathbf{T} \), with respect to \( \mathbf{s}_d \) and \( \mathbf{p} \) or, as is the case here, \( \mathbf{s}_d \) and \( \mathbf{z} \) and \( \mathbf{v} \), then one obtains

\[
\frac{\partial \mathbf{s}_d}{\partial \mathbf{s}_i} \mathbf{L}_0 + \frac{\partial \mathbf{s}_d}{\partial \mathbf{z}} \mathbf{T} + \frac{\partial \mathbf{s}_d}{\partial \mathbf{v}} \mathbf{I}
\]

(4.21)

When we differentiate with respect to \( \mathbf{T} \) at constant \( \mathbf{p} \) there are two routes from \( \mathbf{s}_d \) to \( \mathbf{T} \), one via \( \mathbf{s}_i \) and one direct:

\[
\left( \frac{\partial \mathbf{s}_d}{\partial \mathbf{z}} \right)_p + \left( \frac{\partial \mathbf{s}_d}{\partial \mathbf{v}} \right)_p \mathbf{I}
\]

(4.22)

Inserting eqn. 4.21 we obtain

\[
\left( \frac{\partial \mathbf{s}_d}{\partial \mathbf{z}} \right)_p = \mathbf{L}_0 \left( \frac{\partial \mathbf{s}_d}{\partial \mathbf{s}_i} \right)_p \mathbf{L}_0 + \mathbf{I}
\]

(4.23)

Note how we use the chain rule to follow the different branches of the tree structure of the diagram of functional relationships. Because we are differentiating with respect to vectors, all the dimensions have to be consistent; this is why an identity matrix \( \mathbf{I} \) appears rather than a 1. Finally, we differentiate with respect to \( \mathbf{p} \) at constant \( \mathbf{T} \):

\[
\frac{\partial \mathbf{s}_d}{\partial \mathbf{p}} \mathbf{L}_0 = \mathbf{L}_0 \left( \frac{\partial \mathbf{s}_d}{\partial \mathbf{z}} \right)_p + \mathbf{L}_0 \left( \frac{\partial \mathbf{s}_d}{\partial \mathbf{v}} \right)_p \mathbf{L}_0
\]

(4.24)

We shall need all three of these equations when we next proceed to differentiate the kinetic model in steady state. We first differentiate with respect to \( \mathbf{p} \) at constant \( \mathbf{T} \), and then with respect to \( \mathbf{T} \) at constant \( \mathbf{p} \).

Case 1: \( d\mathbf{p} = 0, d\mathbf{T} = 0 \)

Eqn. 4.11 is an implicit equation of the form \( f(x, y, z) = 0 \). In general, to obtain, say, \( \frac{\partial y}{\partial x} \) one could in principle solve for \( y \) and partially differentiate with respect to \( x \) while keeping \( z \) constant. This is often impossible, but there is, fortunately, a much simpler way around this, namely implicit differentiation

\[
df = \left( \frac{\partial f}{\partial x} \right)_{y,z} dx + \left( \frac{\partial f}{\partial y} \right)_{x,z} dy + \left( \frac{\partial f}{\partial z} \right)_{x,y} dz = 0
\]

(4.25)

where \( df \) is called the total differential. If only one variable, say \( x \), is considered to change at constant \( y \) and \( z \), then one obtains

\[
\left( \frac{\partial f}{\partial x} \right)_{y,z} \frac{\partial x}{\partial x} + \left( \frac{\partial f}{\partial y} \right)_{x,z} \frac{\partial y}{\partial x} + \left( \frac{\partial f}{\partial z} \right)_{x,y} \frac{\partial z}{\partial x} = 0
\]

(4.26)

To differentiate eqn. 4.11 with respect to \( \mathbf{p} \) we use this technique in combination with the chain rule to traverse the three routes from \( \mathbf{v} \) to \( \mathbf{p} \) on the diagram of functional relationships:

\[
\mathbf{N}_d \left[ \frac{\partial \mathbf{v}}{\partial \mathbf{s}_d} \right] \frac{\partial \mathbf{s}_d}{\partial \mathbf{p}} = 0
\]

(4.27)

where \( \mathbf{N}_d \) is a null matrix. Inserting eqn. 4.21 and collecting the first two terms in the square brackets we obtain

\[
\mathbf{N}_d \left[ \frac{\partial \mathbf{v}}{\partial \mathbf{s}_d} \right] \frac{\partial \mathbf{s}_d}{\partial \mathbf{p}} + \mathbf{N}_d \frac{\partial \mathbf{v}}{\partial \mathbf{p}} = 0
\]

(4.28)

By definition the column vector of matrices is \( \mathbf{L} \). The partitioned matrix \( \frac{\partial \mathbf{s}_d}{\partial \mathbf{s}_i} \) is the matrix \( \frac{\partial \mathbf{v}}{\partial \mathbf{s}_d} \frac{\partial \mathbf{s}_d}{\partial \mathbf{p}} \) of partial derivatives of reaction rate functions with respect to the individual concentrations in \( \mathbf{s} \). In control analysis these partial derivatives are called elasticity coefficients, defined as \( E_{xj} = \frac{\partial \mathbf{v}_x}{\partial \mathbf{s}_j} \). The above matrix of elasticity coefficients is symbolised with \( \bar{\xi} \). Similarly, \( \frac{\partial \mathbf{v}}{\partial \mathbf{p}} \bar{\xi} \) is a matrix of elasticity coefficients with respect to \( \mathbf{p} \), symbolised by \( \bar{\xi} \). The bar in \( \bar{\xi} \) reminds us that these are ordinary partial derivatives, not the normalised (scaled) partial derivatives which are usually used in control analysis and which we shall encountered further on (and for which we shall use the unbarred symbol; this distinction between barred (non-normalised) and unbarred (normalised) symbols is made throughout the paper for all the coefficients of control analysis and their matrices). Using this symbolism eqn. 4.28 is written as

\[
\mathbf{N}_d \bar{\xi} \mathbf{L} \frac{\partial \mathbf{s}_d}{\partial \mathbf{p}} + \mathbf{N}_d \bar{\xi} \mathbf{p} = 0
\]

(4.29)

The steady-state eqn. 4.15 for fluxes is explicit. Differentiation with respect to \( \mathbf{p} \) yields:

\[
\frac{\partial \mathbf{j}}{\partial \mathbf{p}} = \left( \frac{\partial \mathbf{v}}{\partial \mathbf{s}_d} \right)_{s, \mathbf{p}} \frac{\partial \mathbf{s}_d}{\partial \mathbf{p}} + \left( \frac{\partial \mathbf{v}}{\partial \mathbf{s}_d} \right)_{\mathbf{s}, \mathbf{p}} \frac{\partial \mathbf{s}_d}{\partial \mathbf{p}} + \frac{\partial \mathbf{v}}{\partial \mathbf{p}}
\]

(4.30)

This expression is identical to the sum of terms within the square brackets of eqn. 4.27. Substituting and collecting terms as before gives

\[
\frac{\partial \mathbf{j}}{\partial \mathbf{p}} = \bar{\xi} \mathbf{L} \frac{\partial \mathbf{s}_d}{\partial \mathbf{p}} + \bar{\xi} \mathbf{p}
\]

(4.31)
Case 2: \( dp = 0, dT = 0 \)

Now we implicitly differentiate eqn. 4.11 with respect to \( T \) at constant \( p \). On the diagram of functional relationships there are two routes from \( v \) to \( T \); one of them branches at \( s_d \) into two subroutes to \( T \). We differentiate in two steps to avoid getting things mixed up—first we differentiate to the level of \( s_i, T, p \):

\[
\mathbf{N}_s \left[ \frac{\partial \mathbf{v}}{\partial s_i}_{s_i,T,p} \left( \frac{\partial s_i}{\partial T}_{s_i,T} \right)_p + \left( \frac{\partial \mathbf{v}}{\partial s_d}_{s_i,T,p} \frac{\partial s_d}{\partial T}_{s_i,T} \right)_p \right] = \mathbf{0} \tag{4.32}
\]

Now we must take care of the two routes from \( s_d \) to \( T \). In fact, we have already done this in eqn. 4.22 so we just substitute:

\[
\mathbf{N}_s \left[ \frac{\partial \mathbf{v}}{\partial s_i}_{s_i,T,p} \left( \frac{\partial s_i}{\partial T}_{s_i,T} \right)_p + \left( \frac{\partial \mathbf{v}}{\partial s_d}_{s_i,T,p} \frac{\partial s_d}{\partial T}_{s_i,T} \right)_p \left[ I_0 \left( \frac{\partial s_i}{\partial T}_{s_i,T} \right)_p + I_{m-r} \right] \right] = \mathbf{0} \tag{4.33}
\]

Multiplying out and collecting the first two terms as before gives

\[
\mathbf{N}_s \left[ \frac{\partial \mathbf{v}}{\partial s_i}_{s_i,T,p} \frac{\partial s_i}{\partial T}_{s_i,T} \right] + \mathbf{N}_s \left. \frac{\partial \mathbf{v}}{\partial s_i}_{s_i,T,p} \frac{\partial s_i}{\partial T}_{s_i,T} \right] \left( \frac{\partial s_i}{\partial T}_{s_i,T} \right)_p + \mathbf{N}_s \left. \frac{\partial \mathbf{v}}{\partial s_d}_{s_i,T,p} \frac{\partial s_d}{\partial T}_{s_i,T} \right] \left( \frac{\partial s_d}{\partial T}_{s_i,T} \right)_p = \mathbf{0} \tag{4.34}
\]

\[
\begin{align*}
\frac{\partial \mathbf{v}}{\partial s_i}_{s_i,T,p} &= \left( \frac{\partial \mathbf{v}}{\partial s_i}_{s_i,T,p} \frac{\partial s_i}{\partial T}_{s_i,T} \right) \left( \frac{\partial s_i}{\partial T}_{s_i,T} \right)_p + \left( \frac{\partial \mathbf{v}}{\partial s_d}_{s_i,T,p} \frac{\partial s_d}{\partial T}_{s_i,T} \right) \left( \frac{\partial s_d}{\partial T}_{s_i,T} \right)_p \\
\end{align*}
\]

This concludes the differentiation of the steady-state equations.

4.4 Metabolic control analysis

We are now in a position to derive the basic definitions of and relationships between all the different matrices of the coefficients of metabolic control analysis (the boxed equations in the rest of this section).

4.4.1 Concentration response with respect to \( p \)

The matrix product \( \mathbf{N}_s \xi_s \mathbf{L} \), which appears in eqn. 4.29, is the so-called Jacobian matrix, which we symbolise with \( \mathbf{M} \). The nature and significance of the Jacobian matrix is explained in Appendix A. Here we just note that if a steady state exists the Jacobian matrix is invertible. From here on, for the sake of brevity, we leave out the subscripts that indicate which vectors remain constant during differentiation. In this spirit, eqn. 4.29 is now written as

\[
\frac{\partial \xi_s}{\partial p} = -\mathbf{M}^{-1} \mathbf{N}_s \xi_p \tag{4.37}
\]

From eqn. 4.24 it follows that

\[
\frac{\partial \xi_d}{\partial p} = -\mathbf{L} \xi_d - \mathbf{L} \mathbf{M}^{-1} \mathbf{N}_s \xi_p \tag{4.38}
\]

Combining the eqns. 4.37 and 4.38 gives

\[
\left[ \frac{\partial \xi_s}{\partial p} \right] = - \left[ \mathbf{I}_r \right] \mathbf{M}^{-1} \mathbf{N}_s \xi_p \tag{4.39}
\]

which simplifies to

\[
\frac{\partial \xi_s}{\partial p} = (-\mathbf{M}^{-1} \mathbf{N}_s) \xi_p \tag{4.40}
\]

We have therefore obtained the first of the matrices of partial derivatives that we seek. The elements of this matrix are concentration-response coefficients, defined as \( K_{jk}^v = \partial s_j / \partial p_k \); they quantify the steady-state response in a metabolite concentration to a perturbation in parameter \( p_k \). The matrix \( \partial \xi_s / \partial p \) will be symbolised by \( \xi_p^v \).

4.4.2 Concentration-control coefficients

What are the elements of the matrix \( -\mathbf{M}^{-1} \mathbf{N}_s \mathbf{v} \) or in more explicit form, \( -\mathbf{L} (\mathbf{N}_s \xi_s \mathbf{L})^{-1} \mathbf{N}_s \mathbf{v} \)? Besides \( \mathbf{L} \) and \( \mathbf{N}_s \), which are integer matrices, this matrix contains only partial derivatives of the rates with respect to the steady-state concentrations, i.e., elasticity constants; nothing in the matrix depends explicitly on \( p \) or \( T \). Now, consider a set of parameters such that each uniquely affects a single reaction, i.e., in the elasticity matrix \( \xi_p^v \),

\[
\frac{\partial v_k}{\partial p_l} = 0 \quad \text{and} \quad \frac{\partial v_k}{\partial p_k} = 0 \quad \text{for} \quad k \neq l \tag{4.41}
\]

This means that \( \xi_p^v \) is now a diagonal matrix (only the diagonal elements \( \partial v_k / \partial p_k \) are non-zero). The inverse of a diagonal matrix is again a diagonal matrix but with all diagonal elements in reciprocal form. Therefore, by multiplying both sides of eqn. 4.40 with this inverse leads to a matrix expression for \( -\mathbf{M}^{-1} \mathbf{N}_s \mathbf{v} \) in which an element in row \( j \) and column \( k \) is

\[
\frac{\partial s_j}{\partial p_k} / \frac{\partial v_k}{\partial p_k} \tag{4.42}
\]

This is, in fact, the fundamental definition of a concentration-control coefficient:

\[
\delta c_{jk}^r = \frac{\partial s_j}{\partial p_k} / \frac{\partial v_k}{\partial p_k} \tag{4.43}
\]
The matrix \(-LM^{-1}N_a\) is therefore the matrix of concentration-control coefficients

\[
\tilde{C}^s = -LM^{-1}N_a
\]  
(4.44)

The matrix \(\frac{\partial s}{\partial p}\) can be recognised as a matrix of concentration-response coefficients \(\tilde{R}_p\). Eqn. 4.40 is therefore a statement of the partitioned concentration-response property of metabolic systems:

\[
\tilde{R}_p^s = \tilde{C}^s \tilde{\xi}_p
\]  
(4.45)

### 4.4.3 Flux-response and control coefficients

Differentiation of the steady-state flux equation with respect to \(p\) led to eqn. 4.31. Inserting eqn. 4.37 we obtain

\[
\frac{\partial J}{\partial p} = \tilde{\xi}_s \left( -LM^{-1}N_a \right) \tilde{\xi}_p + \tilde{\xi}_p
\]  
(4.46)

From eqn. 4.44 we recognise the bracketed term as \(\tilde{C}^s\), so that

\[
\frac{\partial J}{\partial p} = (\tilde{\xi}_s \tilde{C}^s + I_n) \tilde{\xi}_p
\]  
(4.47)

Similar to the previous section, an element in the \(i\)th row and \(k\)th column of \(\tilde{C}^s \tilde{\xi}_p + I_n\) can be seen to be a flux-control coefficient of reaction \(k\):

\[
\tilde{C}^s_{ik} = \frac{\partial J_i}{\partial p_k} \frac{\partial v_k}{\partial p_k}
\]  
(4.48)

so that the matrix of flux-control coefficients is defined as

\[
\tilde{C}^s = \tilde{\xi}_s \tilde{C}^s + I_n
\]  
(4.49)

The matrix \(\frac{\partial J}{\partial p}\) can be recognised as a matrix \(\tilde{R}_p^s\) of flux-response coefficients, individually defined as \(\tilde{R}^s_{ik} = \frac{\partial J_i}{\partial p_k}\), that quantify the steady-state response in a flux \(J_i\) to a perturbation in parameter \(p_k\). Eqn. 4.47 is therefore a statement of the partitioned flux-response property of metabolic systems:

\[
\tilde{R}_p^s = \tilde{C}^s \tilde{\xi}_p
\]  
(4.50)

### 4.4.4 Concentration-response with respect to \(T\)

Differentiation of eqn. 4.11 with respect to \(T\) led to eqn. 4.34, which, if \(\frac{\partial s}{\partial T}\) containing concentration-response coefficients with respect to the conservation sums is symbolised by \(\tilde{R}_T^s\), can be written as

\[
N_a \tilde{\xi}_s L \tilde{R}_T^s + N_a \tilde{\xi}_s = 0
\]  
(4.51)

which, using \(N_a \tilde{\xi}_s L = M\) and re-arranging gives

\[
\tilde{R}_T^s = -M^{-1}N_a \tilde{\xi}_s
\]  
(4.52)

Eqn. 4.23 can be written as

\[
\tilde{R}_T^s = L \tilde{\xi}_s
\]  
(4.53)

which, inserting eqn. 4.52, is

\[
\tilde{R}_T^s = -L M^{-1}N_a \tilde{\xi}_s + I_{m-r}
\]  
(4.54)

Combining eqns. 4.52 and 4.54 gives

\[
\begin{bmatrix}
\tilde{R}_T^s \\
\tilde{R}_T^c
\end{bmatrix}
= -
\begin{bmatrix}
L_r \\
L_o
\end{bmatrix} M^{-1}N_a \tilde{\xi}_s +
\begin{bmatrix}
0 \\
I_{m-r}
\end{bmatrix}
\]  
(4.55)

which reduces to

\[
\tilde{R}_T^c = -L M^{-1}N_a \tilde{\xi}_s +
\begin{bmatrix}
0 \\
I_{m-r}
\end{bmatrix}
\]  
(4.56)

As \(-LM^{-1}N_a = \tilde{C}^s\) we get

\[
\tilde{R}_T^c = \tilde{C}^s \tilde{\xi}_s +
\begin{bmatrix}
0 \\
I_{m-r}
\end{bmatrix}
\]  
(4.57)

Finally, we prefer to write this equation in terms of the full elasticity matrix. This can be done if we realise that

\[
\tilde{\xi}_s =
\begin{bmatrix}
\tilde{\xi}_{s_1} \\
\tilde{\xi}_{s_2} \\
\vdots \\
\tilde{\xi}_{s_n}
\end{bmatrix}
\begin{bmatrix}
0 \\
I_{m-r}
\end{bmatrix}
\]  
(4.58)

Therefore

\[
\tilde{R}_T^c = (\tilde{C}^s \tilde{\xi}_s + I_m) 
\begin{bmatrix}
0 \\
I_{m-r}
\end{bmatrix}
\]  
(4.59)

The righthand matrix singles out the dependent metabolites that are each unique to a different conservation equation, ensuring that each conservation sum is perturbed independently.

### 4.4.5 Flux-response with respect to \(T\)

Differentiation of eqn. 4.15 with respect to \(T\) led to eqn. 4.36, which, if \(\frac{\partial s}{\partial T}\) containing flux-response coefficients with respect to the conservation sums is symbolised by \(\tilde{R}_T^s\), can be re-written as

\[
\tilde{R}_T^s = \tilde{\xi}_s L \tilde{R}_T^s + \tilde{\xi}_{s_1}
\]  
(4.60)

We can now insert the expression for \(\tilde{R}_T^s\) in eqn. 4.52 to give

\[
\tilde{R}_T^s = \tilde{\xi}_s (-LM^{-1}N_a) \tilde{\xi}_s + \tilde{\xi}_s
\]  
(4.61)
As $\mathbf{C}^* = -\mathbf{L}^{-1} \mathbf{N}_a$ we obtain

$$\mathbf{R}^i_T = (\mathbf{C}^* + \mathbf{I}_n) \mathbf{\check{\xi}_m}$$  \hspace{1cm} (4.62)

The bracketed term is the defining expression for $\mathbf{C}^i$ (eqn. 4.49). Therefore,

$$\mathbf{R}^i_T = \mathbf{C}^i \mathbf{\check{\xi}_m}$$  \hspace{1cm} (4.63)

As before, we rather write this in terms of the full elasticity matrix

$$\mathbf{R}^i_T = \mathbf{C}^i \mathbf{\check{\xi}}$$ \hspace{1cm} (4.64)

It is possible to express $\mathbf{R}^i_T$ in terms of $\mathbf{\check{R}}^i_T$. Inserting eqn. 4.49 into eqn. 4.64 gives

$$\mathbf{R}^i_T = (\mathbf{C}^* + \mathbf{I}_n) \mathbf{\check{\xi}_s}$$ \hspace{1cm} (4.65)

Multiplying the first RHS product out and recollecting terms gives

$$\mathbf{R}^i_T = \mathbf{\check{\xi}_s} (\mathbf{C}^* \mathbf{\check{\xi}_s} + \mathbf{I}_m)$$ \hspace{1cm} (4.66)

Using eqn. 4.59 we finally get [1]

$$\mathbf{R}^i_T = \mathbf{\check{\xi}_s} \mathbf{\check{R}}^i_T$$ \hspace{1cm} (4.67)

### 4.4.6 Normalising the central equations

In control analysis the use of the dimensionless normalised form of the control and elasticity coefficients is generally preferred [24, 27]. With one trivial exception, the basic equations developed in the previous sections look exactly the same in normalised form, provided that the $\mathbf{K}$, $\mathbf{L}$, $\mathbf{N}_a$ and $\mathbf{M}$-matrices and are scaled appropriately. To do this we define the diagonal matrices $\mathbf{D}^i$ and $\mathbf{D}^e$ which respectively have the steady-state fluxes and concentrations on their diagonal (just as with the coefficient matrices the flux and concentration vectors are arranged so that the independent variables come first, the dependent variables second). Their inverses ($\mathbf{D}^i\mathbf{I}$ and $\mathbf{D}^e\mathbf{I}$) have inverse fluxes and inverse steady-state concentrations on their diagonals. Similarly, we define $\mathbf{D}^p$, $\mathbf{D}^s$, and $\mathbf{D}^T$. Using these diagonal matrices, the matrices that occur in the control-matrix equation are scaled as follows (note that the absence of a bar denotes normalised matrices):

$$\mathbf{C}^i = (\mathbf{D}^i)^{-1} \cdot \mathbf{\check{C}}^i \cdot \mathbf{D}^i$$ \hspace{1cm} (4.68)

$$\mathbf{C}^* = (\mathbf{D}^e)^{-1} \cdot \mathbf{\check{C}}^* \cdot \mathbf{D}^e$$ \hspace{1cm} (4.69)

$$\mathbf{\check{\xi}_s} = (\mathbf{D}^s)^{-1} \cdot \mathbf{\check{\xi}_s} \cdot \mathbf{D}^s$$ \hspace{1cm} (4.70)

$$\mathbf{\check{\xi}_m} = (\mathbf{D}^T)^{-1} \cdot \mathbf{\check{\xi}_m} \cdot \mathbf{D}^T$$ \hspace{1cm} (4.71)

$$\mathbf{R}_p^i = (\mathbf{D}^i)^{-1} \cdot \mathbf{\check{R}}_p^i \cdot \mathbf{D}^p$$ \hspace{1cm} (4.72)

$$\mathbf{R}_s^i = (\mathbf{D}^e)^{-1} \cdot \mathbf{\check{R}}_s^i \cdot \mathbf{D}^s$$ \hspace{1cm} (4.73)

$$\mathbf{R}_T^i = (\mathbf{D}^s)^{-1} \cdot \mathbf{\check{R}}_T^i \cdot \mathbf{D}^T$$ \hspace{1cm} (4.74)

$$\mathbf{L} = (\mathbf{D}^i)^{-1} \cdot \mathbf{L} \cdot \mathbf{D}^s$$ \hspace{1cm} (4.75)

$$\mathbf{K} = (\mathbf{D}^i)^{-1} \cdot \mathbf{K} \cdot \mathbf{D}^e$$ \hspace{1cm} (4.76)

$$\mathbf{N}_a = (\mathbf{D}^s)^{-1} \cdot \mathbf{N}_a \cdot \mathbf{D}^s$$ \hspace{1cm} (4.77)

$$\mathbf{M} = (\mathbf{D}^s)^{-1} \cdot \mathbf{M} \cdot \mathbf{D}^s$$ \hspace{1cm} (4.78)

The equations central to control analysis derived above (the boxed equations) are now summarised in the normalised format:

Matrix definition of concentration-control coefficients

$$\mathbf{C}^* = -\mathbf{L}^{-1} \mathbf{N}_a$$ \hspace{1cm} (4.79)

Matrix definition of flux-control coefficients

$$\mathbf{C}^i = \mathbf{\xi}_s \mathbf{C}^a + \mathbf{I}_n$$ \hspace{1cm} (4.80)

Partitioned concentration-response property with respect to parameters $\mathbf{p}$:

$$\mathbf{R}_p^i = \mathbf{C}^i \mathbf{\xi}_p$$ \hspace{1cm} (4.81)

Partitioned flux-response property with respect to parameters $\mathbf{p}$:

$$\mathbf{R}_p^e = \mathbf{C}^e \mathbf{\xi}_p$$ \hspace{1cm} (4.82)

The partitioned response properties with respect to $\mathbf{T}$ differ slightly from the non-normalised eqns. 4.59 and 4.64 in that the identity submatrix in the righthand matrix is replaced by a matrix containing inverse mole fractions of the dependent metabolites on its diagonal. The partitioned concentration-response property with respect to $\mathbf{T}$ is:

$$\mathbf{R}_p^i = (\mathbf{C}^s \mathbf{\xi}_s + \mathbf{I}_m) \left[ \begin{array}{c} 0 \\ \mathbf{D}^e \mathbf{\check{\xi}_s} \end{array} \right]$$ \hspace{1cm} (4.83)

Partitioned flux-response property with respect to $\mathbf{T}$:

$$\mathbf{R}_p^e = \mathbf{C}^e \mathbf{\xi}_s$$ \hspace{1cm} (4.84)

Relationship between $\mathbf{R}_p^i$ and $\mathbf{\check{R}}_T^i$:

$$\mathbf{R}_p^i = \mathbf{\check{\xi}_p} \mathbf{\check{R}}_T^i$$ \hspace{1cm} (4.85)
4.4.7 Summation theorems

The summation equations for flux and concentration control coefficients follow directly from the definitions of \( C^k \) (eqn. 4.79) and \( C^l \) (eqn. 4.80) and the relationship \( N_a \epsilon_a K = 0 \) (the normalised form of \( N_a K = 0 \)). The first is called the summation theorem for concentration-control coefficients:

\[
C^k \epsilon_k = -L(N_a \epsilon_a L)^{-1}N_a \epsilon_a L = 0 \tag{4.86}
\]

and the second the summation theorem for flux-control coefficients:

\[
C^l \epsilon_L = (\epsilon_s C^s + I_n) \epsilon_L = K \tag{4.87}
\]

4.4.8 Connectivity theorems

The flux and concentration connectivity equations follows from the invertibility of the Jacobian matrix \( N_a \epsilon_a L \) (the normalised form of \( N_a \epsilon_a L \)). Multiplying \( C^k \) and \( C^l \) by \( \epsilon_s L \) gives, first, the connectivity theorem for flux-control coefficients:

\[
C^k \epsilon_s L = -L(N_a \epsilon_a L)^{-1}N_a \epsilon_a L = -L \tag{4.88}
\]

and, second, the connectivity theorem for flux-control coefficients:

\[
C^l \epsilon_s L = (\epsilon_s C^s + I_n) \epsilon_s L = 0 \tag{4.89}
\]

Together, the summation and connectivity theorems allows the expression of control coefficients in terms of elasticity coefficients. This is arguably the most powerful feature of metabolic control analysis and is treated next.

4.4.9 The control-matrix equation

It is possible to combine the summation and connectivity theorems into a generalised matrix form, which we call the control-matrix equation. Quite a few permutations of such an equation have been suggested \([5, 6, 15, 16, 17, 42, 54, 59, 60, 66, 70, 71]\), but the one that follows arises naturally from the formalism developed in this paper \([31, 35]\). Furthermore, it simplifies to the form \( CE = I \) (see below), which shows explicitly how the matrix expressing independent systemic properties, \( C \), and the matrix expressing structural and local properties, \( E \), are inverses of each other (if the product of two square matrices equals the identity matrix, then they are inverses of each other). This means that control coefficients can be calculated from elasticity coefficients, \( C = E^{-1} \), and vice versa, \( E = (C^l)^{-1} \) (the last case being strictly true only if there are no conservation equations; see next section). The result will once again stress the fundamental role of the \( K \) and \( L \)-matrices in control analysis.

The control-matrix equation is formed by combining eqns. 4.86–4.89 as \([31, 35]\):

\[
\begin{bmatrix}
C^k \\ C^l
\end{bmatrix}
\begin{bmatrix}
K \\ -\epsilon_s L
\end{bmatrix}
= \begin{bmatrix}
K & 0 \\ 0 & L
\end{bmatrix}
\tag{4.90}
\]

The matrices can be partitioned in terms of independent and dependent variables to give

\[
\begin{bmatrix}
C^k \\ C^l
\end{bmatrix}
\begin{bmatrix}
K \\ -\epsilon_s L
\end{bmatrix}
= \begin{bmatrix}
I_{n-r} & 0 \\ 0 & K_{0,0} \\ 0 & L_0
\end{bmatrix}
\tag{4.91}
\]

Extracting the equations for the independent variables \( J_i \) and \( S_i \) gives:

\[
\begin{bmatrix}
C^k \\ C^l
\end{bmatrix}
\begin{bmatrix}
K \\ -\epsilon_s L
\end{bmatrix}
= \begin{bmatrix}
I_{n-r} & 0 \\ 0 & L_0
\end{bmatrix}
\tag{4.92}
\]

which, if \( C^l = [C^h C^s]^T \) and \( E = [K - \epsilon_s L] \), reduces to the particularly elegant form:

\[
C^l E = I_n \tag{4.93}
\]

Both \( C^l \) and \( E \) are square invertible \( n \times n \) matrices \([24, 31]\), i.e., the equation can also be written as \( CE = I \), which expresses flux-control coefficients in terms of concentration-control and elasticity coefficients. These equations are completely general and hold for any network of reactions.

4.4.10 The inverse problem

We have seen that \( C^l = E^{-1} \); if all the elasticity coefficients have been determined (either experimentally or by calculation as the normalised partial derivatives of the rate equations), the control coefficients with respect to the independent concentrations and fluxes can be calculated by inverting \( E \). The control coefficients with respect to the dependent variables are calculated using the relationships:

\[
\begin{align*}
C^u &= E \cdot C^h \\
C^h &= K^{-1} \cdot C^k
\end{align*}
\tag{4.94}
\tag{4.95}
\]

which follow from eqns. 4.2 and 4.18.

However, consider the inverse problem, i.e., calculating the elasticity coefficients from experimentally determined control coefficients. Using \( E = (C^l)^{-1} \) we can calculate \( E \) by inverting \( C^l \). If \( L = I \), i.e., if there are no conservation constraints, the task is accomplished—the right hand \( r \) columns of \( E \) form the elasticity matrix \( -\epsilon_s \), and therefore contain the values of the elasticity coefficients. However, if \( L \neq I \), some elements in the right hand \( r \) columns of \( E \) contain linear functions of elasticity coefficients, and more information is needed to solve for the individual elasticities. This extra information can only be obtained by perturbing the conservation sums.
in the column vector $T$ and measuring the resulting steady-state changes in all the fluxes and concentrations. We augment on the left both sides of eqn. 4.85:

$$R_1 = \varepsilon sR_1$$

(4.96)

with the matrix $\varepsilon sL$ to give

$$[\varepsilon sL] R_1 = [\varepsilon sL] R_1^1$$

(4.97)

which can be re-arranged to solve for $\varepsilon s$:

$$\varepsilon s = [\varepsilon sL R_1^1][L R_1^1]^{-1}$$

(4.98)

The $n \times n$ matrix $[L R_1^1]$ has been proved to be invertible [1].

4.5 Discussion

This paper set out to provide, in a nutshell, the complete formal basis for metabolic control analysis in a way that leaves as little as possible unexplained. In particular, care has been taken to show how the functional relationships in the steady-state equations hang together, thereby proscribing how the steady-state equations should be differentiated. For a more extensive exposition of much of the material covered in this paper the reader is referred to the excellent monograph by Heinrich and Schuster [24], which is a treasure trove of information on biochemical modelling and control analysis. However, there are aspects covered here which are either absent from their treatment (the response to $T$ in Sections 4.4.4 and 4.4.5, and the inverse problem in 4.4.10) or different (Sections 4.4.7, 4.4.8, and 4.4.9, where full scaling is used instead of the partial scaling used in [24]).

Metabolic control analysis has been applied to many types of systems, which has led to interesting extensions of the theory, for example, multi-level or hierarchical systems [38, 43], modular systems [56, 63], signal transduction pathways [45], time-dependent phenomena [23, 24], transition times [49], oscillating systems [12], channelled systems [44], and group-transfer pathways [46].

Co-response analysis [31] is an extension built on the control-matrix equation described in this paper. It not only has useful experimental implications (control analysis requiring neither kinetic knowledge of the component reactions nor quantitative information about the magnitudes of the effects of perturbations on individual enzyme activities), but also forms the basis for the analysis of regulatory aspects of metabolism (for example, supply-demand analysis [27, 30, 34]).

4.6 Appendix: The Jacobian matrix

In general, any dynamical system $dx/dt = f(x)$, where the $f$ are nonlinear functions of $x$, can be linearised around any current state $x^o$ (transient or steady) to give $dx^o/dt = Mx^o$, where $M$ is the Jacobian matrix $2f/2x_i$ i.e., the matrix of partial derivatives of $f$ with respect to $x$ evaluated at state $x^o$. If a steady state is considered, a necessary condition for its existence is that the Jacobian matrix be invertible (its determinant be non-zero), and for its asymptotic stability that the eigenvalues of the Jacobian matrix have negative real parts (see, e.g., [24]). The Jacobian matrix therefore characterises the local behaviour around the steady state. We now show that for the kinetic model discussed in this paper $M = N_0 \varepsilon sL$ is the Jacobian matrix.

Assume that for the kinetic model

$$\frac{ds}{dt} = Nv[s, p]$$

(4.99)

the current state is symbolised by the vector of concentrations $s^o$. As $p$ is assumed to be constant in the following, we simplify the representation of the kinetic model to

$$\frac{ds^o}{dt} = Nv[s^o]$$

(4.100)

When the current concentrations are perturbed by $\delta s$ so that

$$s(t) = s^o + \delta s(t)$$

(4.101)

the kinetic model becomes:

$$\frac{ds^o}{dt}(s^o + \delta s) = Nv[s^o + \delta s]$$

(4.102)

From multivariate calculus we know that Taylor’s theorem enables us to approximate to any degree of accuracy the function $v[s^o + \delta s]$ by the expansion

$$v[s^o] + \frac{\partial v}{\partial s} \cdot (\delta s) + \frac{1}{2!} \left( \frac{\partial^2 v}{\partial s^2} \right)_{s^o} (\delta s)^2 + \text{higher order terms}$$

(4.103)

For small deviations $\delta s$ the first two terms suffice to approximate the function, so that

$$\frac{ds^o}{dt}(\delta s) = Nv[s^o] + N \left( \frac{\partial v}{\partial s} \right) \cdot (\delta s)$$

(4.104)

which amounts to a linearisation around state $s^o$. By definition the matrix of partial derivatives $\partial v/\partial s$ is the matrix of non-normalised elasticity coefficients $\varepsilon s$, evaluated at the state characterised by $s^o$. Using eqn. 4.100 we obtain the linearised form of the kinetic model at state $s^o$:

$$\frac{d}{dt}(\delta s) = N_0 \varepsilon sL \cdot \delta s$$

(4.105)

If the number of independent concentrations is less than the number of metabolites then by definition $N = LN_0$ and $s = Ls_0$, or equivalently, $\delta s = L\delta s_0$. Using the argument of eqns. 4.4-4.6 it follows that

$$\frac{d}{dt}(\delta s_0) = N_0 \varepsilon sL \cdot \delta s_0$$

(4.106)
From the general definition of the Jacobian matrix given in the first paragraph of this section we see that

\[ M = N \hat{e}_v L \]  \hspace{1cm} (4.107)

is the Jacobian matrix.

The Jacobian matrix can be normalised as follows: Define the diagonal matrices \( D_v \), \( D_s \) and \( L \) which respectively have the reaction rates, concentrations and independent concentrations obtaining at state \( s^o \) on their diagonal (the rate and concentration vectors are arranged so that the independent variables come first, the dependent variables second). Their inverses \( (D_v)^{-1} \), \( (D_s)^{-1} \) and \( (D_i)^{-1} \) have the inverse rates and inverse concentrations on their diagonals. Using the identities

\[ \delta \ln s_i = (D_i)^{-1} \delta s_i \]  \hspace{1cm} (4.108)

\[ N \xi_v = (D_v)^{-1} \cdot N_v \cdot D_v \]  \hspace{1cm} (4.109)

\[ \xi_v = (D_v)^{-1} \cdot \hat{e}_v \cdot D_v \]  \hspace{1cm} (4.110)

\[ L = (D_v)^{-1} \cdot L \cdot D_v \]  \hspace{1cm} (4.111)

eqn. 4.106 can be written as

\[ \frac{d}{dt} (\delta \ln s_i) = N \xi_v L \cdot \delta \ln s_i \]  \hspace{1cm} (4.112)

The kinetic model for perturbations from state \( s^o \) has therefore been transformed to logarithmic space. From this formulation the normalised Jacobian matrix is seen to be

\[ \mathcal{M} = N \hat{e}_v L \]  \hspace{1cm} (4.113)

which, if there are no dependent metabolites \( L = I \) simplifies to

\[ \mathcal{M} = N \xi_v \]  \hspace{1cm} (4.114)

4.7 Appendix: An explicit example

Fig. 4.1 represents a simple reaction network containing both a branched flux and a moiety-conserved cycle [36]. Here we show how the \( K \) and \( L \)-matrices can be constructed from an analysis of its stoichiometric matrix. Once these matrices are available it is a simple matter to formulate the control matrix equation \( CE = I \) explicitly (for a numerical solution of this example see [31]).

The first step is to write down the stoichiometric matrix \( N \) for this system, labelling the rows and columns (the left-hand matrix in eqn. 4.115). \( N \) is then augmented with an identity matrix in which each column represents a time derivative (the right-hand matrix in eqn. 4.115). Note that only variable metabolites \( S_i \) are represented. The terminal X-metabolite concentrations must be constant (at non-equilibrium values) for a steady state to exist, and are therefore considered part of the parameter set.

\[ \begin{array}{cccccccc}
R_1 & R_2 & R_3 & R_4 & \delta_1 & \delta_2 & \delta_3 \\
S_1 & 1 & 0 & -1 & -1 & 1 & 0 & 0 \\
S_2 & -1 & 1 & 0 & 0 & 0 & 1 & 0 \\
S_3 & 1 & -1 & 0 & 0 & 0 & 0 & 1 \\
\end{array} \]  \hspace{1cm} (4.115)

Next the augmented matrix is subjected to Gaussian elimination to row echelon form\(^2\):

\[ \begin{array}{cccccccc}
R_1 & R_2 & R_3 & R_4 & \tilde{s}_1 & \tilde{s}_2 & \tilde{s}_3 \\
S_1 & 1 & 0 & -1 & -1 & 1 & 0 & 0 \\
S_2 & 0 & 1 & -1 & -1 & 1 & 1 & 0 \\
S_3 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
\end{array} \]  \hspace{1cm} (4.116)

The rank of the stoichiometric matrix is 2 and there is one conservation relationship \( \tilde{s}_2 + \tilde{s}_3 = T \), where \( T \) is the conserved sum. There are two independent fluxes and two independent metabolites.

Choosing \( J_3 \) and \( J_4 \) (the \( R_3 \) and \( R_4 \)-columns without pivots in the reduced stoichiometric matrix \( N_a \)) in eqn. 4.116) as the independent fluxes, the \( K \)-matrix follows from the flux relationships \( J = K_j \):

\[ \begin{bmatrix}
J_3 \\
J_4 \\
J_3 + J_4 \\
J_3 + J_4 \\
\end{bmatrix} = \begin{bmatrix}
J_3 \\
J_4 \\
J_3 + J_4 \\
J_3 + J_4 \\
\end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
1 & 1 & 0 & 0 \\
1 & 1 & 0 & 0 \\
\end{bmatrix} \begin{bmatrix}
J_3 \\
J_4 \\
J_3 + J_4 \\
J_3 + J_4 \\
\end{bmatrix} \]  \hspace{1cm} (4.117)

\(^2\text{Simple systems can be analysed by hand, more complicated systems with one of the numerous computer tools that do this automatically, such as Mathematica, Matlab, etc. or dedicated metabolic simulators (e.g., \([51, 57, 69]\), Metatool [ftp://mudshark.brookes.ac.uk/pub/software/bmpc/metatool])}
Either $S_3$ or $S_2$ can be chosen as the dependent metabolite. We choose $S_3$. The $L$ matrix follows from the relationships in eqn. 4.116:

$$\dot{S_3} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\dot{S_2} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\dot{S_1} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The matrix product $\mathbf{K} \mathbf{L}^{-1}$ as follows:

$$\mathbf{K} = \begin{bmatrix} K_{11} & K_{12} & \cdots & K_{1n} \\ K_{21} & K_{22} & \cdots & K_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ K_{n1} & K_{n2} & \cdots & K_{nn} \end{bmatrix}$$

$$\mathbf{L}^{-1} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

$$\mathbf{K} \mathbf{L}^{-1} = \begin{bmatrix} K_{11} & K_{12} & \cdots & K_{1n} \\ K_{21} & K_{22} & \cdots & K_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ K_{n1} & K_{n2} & \cdots & K_{nn} \end{bmatrix} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

To solve the inverse problem ($\mathbf{L}^{-1}$ from eqn. 4.116.1), we need to construct eqn. 4.116.1. There is only one conserved sum $S_1 + S_3$. The matrix product $\mathbf{K} \mathbf{L}^{-1}$ is
5 Metabolic Regulation: Supply-Demand Analysis

5.1 Introduction

How highly would one rate an economic analysis of a factory that ignored the consumer demand for its products? Ludicrous as it may sound, this is precisely what most metabolic studies of the past century have been doing. If this seems far-fetched, consider for example that we have yet to find a textbook analysis of, say, biosynthetic flux to an amino acid that takes into account the rate of protein synthesis. This state of affairs is perhaps understandable: faced with the huge complexity of the cellular reaction network the only way to proceed was to chop it up into manageable parts and study the parts separately in terms of stoichiometric structure, enzymes, and transporters. However, although all these parts are undoubtedly connected, the current view of metabolism and its regulation still seems to be that the parts behave the same whether in isolation or in cellular context. A telling example is the continued insistence of modern biochemistry textbooks on the purported rate-limiting role of the kinases in glycolysis, despite clear evidence that over-expression of these (and other) glycolytic enzymes either on their own or in combination has no effect on the carbon flux in vivo from glucose to ethanol in yeast [62, 28].

Here we outline a quantitative theory called metabolic supply-demand analysis that addresses this problem by allowing the integration of the different parts of metabolism with each other and with other intracellular processes. Within this framework the concepts of metabolic regulation and function acquire a clear and quantitative meaning. In addition, a number of concepts central to the classical view of metabolic regulation are shown to be fallacious.

5.2 Metabolic regulation, organisation and function

We consider metabolic regulation to be inextricably linked to function: to say a system is regulated is to mean that its intrinsic properties have been moulded by evolution to fulfill specific functions [30, 27]. Because mass action is the intrinsic driving force for self-organisation of reaction networks, we broadly define metabolic regulation as the alteration of reaction properties to augment or counteract the mass-action trend in a network of reactions [27, 55]. A corollary to this definition is that regulatory performance should always be measured in terms of a specified function. Reaction properties can be regulated by altering the concentrations and the catalytic and binding properties of enzymes; a host of such regulatory mechanisms have evolved [27, 8]. Enzymes lift the metabolic network from the underlying network of thermodynamically feasible reactions onto a different timescale and therefore act as the primary 'handles' through which evolution can create function.

Central to any understanding of metabolic function is our knowledge of the organisation of the metabolic network. Its core consists of a catabolic block that provides phosphorylation and reducing power plus carbon skeletons, a biosynthetic block that makes building blocks for macromolecular synthesis, and a 'growth' block that makes and maintains the cellular structure and the gene and enzyme machinery. These blocks are coupled by either one common intermediate (e.g., an amino acid or nucleotide) or a pair of common intermediates that form a moiety-conserved cycle in which the sum of the cycle members remains constant (e.g., NADP(H)-NADP(H), ATP-ADP or, in the presence of adenylate kinase, ATP-ADP-AMP [2]). To remind ourselves that the living process is intrinsically a molecular economy (cf. [7]) we call the producing block in these linkages the supply and the consuming block the demand (Fig. 5.1). Although in this article we restrict the discussion to the simplest case (Fig. 5.1) the same general approach applies with almost no differences to supply-demand systems that involve moiety-conserved cycles [28].

![Figure 5.1: A metabolic supply-demand system.](image)

The metabolic network is an open system that can exist in either a transient or a steady state. Although equilibrium is excluded as a possible state for living systems, it is an important reference state: the distance \( \rho \) of any reaction or reaction block from equilibrium, defined as \( \Gamma/K_{eq} \) where \( \Gamma \) is the mass-action ratio and \( K_{eq} \) is the equilibrium constant, is an important factor in determining the behaviour of any reaction network.

Metabolic function is a multi-level concept. At the lowest level the function of an enzyme is to catalyse a reaction. At the level of the integrated system of coupled enzyme-catalysed reactions its function may be to control a steady-state metabolite concentration. Enzymes are regulated to perform these higher-level systemic functions, namely: (i) the determination of the steady-state itself, (ii) control over the steady-state fluxes and intermediate concentrations, (iii) the steady state response to a perturbation in some system parameter or a fluctuation in some intermediate concentration (structural and dynamic stability), (iv) the time of transition from one steady state to another [49], and (v) the dynamic form of the transient or steady

---

states (e.g., point, monotonic, oscillatory, trigger, chaotic [24, 20]). Here we only consider the first three functions, although the others are also important in a complete supply-demand analysis.

5.3 Quantitative analysis of supply-demand systems

We now describe a theory that allows a visual and quantitative analysis of how the properties of the supply and demand blocks determine the behaviour and control of the steady-state flux and concentration of P. As our main tools we use rate characteristics [27] and control analysis [41, 21].

\[
\begin{align*}
\epsilon_{\text{supply}} & = \frac{\partial \ln v_{\text{supply}}}{\partial \ln p} \\
\epsilon_{\text{demand}} & = \frac{\partial \ln v_{\text{demand}}}{\partial \ln p}
\end{align*}
\]

Figure 5.2: The rate characteristics of a supply-demand system plotted in double logarithmic space.

The graph of combined rate characteristics (Fig. 5.2) is a powerful tool for visualising how the steady state in a supply-demand system is formed, and how the distribution of flux and concentration control depends on the properties of the supply and demand blocks. On the graph the natural logarithms of the supply and demand rates are plotted as a function of the natural logarithm of the concentration variable that links them. If the supply and demand were catalysed by single enzymes these curves would represent, for example, the familiar Michaelis-Menten or Hill responses of a rate with respect to a product or a substrate. In general, however, the supply and demand are reaction blocks, so that the rate curves actually represent the variation in the local steady-state fluxes of the isolated supply and demand blocks as they respond to variation in the concentration of P. The use of logarithmic rather than linear scales has a number of advantages [27], the most important being that it allows direct comparison of the magnitude of steady state responses to perturbations at different positions of the rate and concentration scale.

The intersection of the supply and demand rate characteristic represents the steady state, which is characterised by a flux, \( J \), and concentration of P, \( \bar{p} \). From the graph it should be clear that the response in the steady state to small perturbations in the activities of supply or demand depend completely on the elasticity coefficients, i.e., the slopes of the tangents to the double logarithmic rate characteristics at the steady-state point.

Figure 5.3: How the steady state (0) responds to a small increase \( d \ln v \) in the activities of either supply (leading to a new steady state at 1) or demand (leading to a new steady state at 2) or both supply and demand (leading to a new steady state at 3).

Fig. 5.3 shows how flux and concentration control can be quantified [50]. Consider a small increase \( d \ln v \) in the activity of the supply, caused by, say, an increase in the concentrations of the supply enzymes. The system moves from the original steady state 0 to a new steady state 1; flux increases by \( d \ln J_1 \) and \( \bar{p} \) by \( d \ln \bar{p} \). Similarly, if the demand activity is increased by \( d \ln v \) the system moves from steady state 0 to 2 with a flux increase of \( d \ln J_2 \) and a decrease in \( \bar{p} \) of \( d \ln \bar{p} \). The degrees to which supply and demand control \( J \) and \( \bar{p} \) are given by the flux-control coefficients

\[
C^J_{\text{supply}} = \frac{d \ln J_1}{d \ln v_{\text{supply}}}, \quad C^J_{\text{demand}} = \frac{d \ln J_2}{d \ln v_{\text{demand}}}
\]

and the concentration-control coefficients

\[
C^p_{\text{supply}} = \frac{d \ln \bar{p}}{d \ln v_{\text{supply}}}, \quad C^p_{\text{demand}} = \frac{-d \ln \bar{p}}{d \ln v_{\text{demand}}}
\]

If both supply and demand are both increased by \( d \ln v \) the system moves to steady state 3 in which the flux has increased by \( d \ln J_2 + d \ln J_3 = d \ln v \) while \( \bar{p} \) remains unchanged. Using the definition of control coefficients given in eqns. eq:ccJdef and eq:ccpdef it follows that

\[
C^J_{\text{supply}} + C^J_{\text{demand}} = 1 \quad (5.3)
\]

\[
C^p_{\text{supply}} + C^p_{\text{demand}} = 0 \quad (5.4)
\]
These are specific cases of the so-called summation theorems of control analysis [41]. Furthermore, using the definitions of the elasticities of supply and demand given in Fig. 5.2 the connectivity theorems [41] can also be derived:

\[ C^J_{\text{supply}} \varepsilon_p^{\text{supply}} + C^J_{\text{demand}} \varepsilon_p^{\text{demand}} = 0 \]

\[ (5.5) \quad \text{and} \]

\[ C^J_{\text{supply}} \varepsilon_p^{\text{supply}} + C^J_{\text{demand}} \varepsilon_p^{\text{demand}} = -1 \]

\[ (5.6) \]

The summation and connectivity theorems provide enough information to express the control coefficients in terms of elasticities of supply and demand [30]. The flux-control coefficients are

\[ C^J_{\text{supply}} = \frac{\varepsilon_p^{\text{demand}}}{\varepsilon_p^{\text{demand}} - \varepsilon_p^{\text{supply}}} \]

\[ (5.7) \quad \text{and} \]

\[ C^J_{\text{demand}} = \frac{\varepsilon_p^{\text{supply}}}{\varepsilon_p^{\text{demand}} - \varepsilon_p^{\text{supply}}} \]

\[ (5.8) \]

and the concentration-control coefficients:

\[ C^p_{\text{supply}} = -C^p_{\text{demand}} = \frac{1}{\varepsilon_p^{\text{demand}} - \varepsilon_p^{\text{supply}}} \]

\[ (5.9) \]

Note that \( \varepsilon_p^{\text{supply}} \) is typically a negative quantity, i.e., product inhibits supply. The ratio of elasticities determines the distribution of flux-control between supply and demand (if \( |\varepsilon_p^{\text{supply}}/\varepsilon_p^{\text{demand}}| > 1 \) the demand has more control over the flux than the supply; if \( |\varepsilon_p^{\text{supply}}/\varepsilon_p^{\text{demand}}| < 1 \) the demand has less control over the flux than the supply). With regard to \( \bar{p} \) it is not the distribution of \( \bar{p} \)-control that is of interest \( (C^p_{\text{supply}} \) always being equal to \( -C^p_{\text{demand}} \) no matter what the values of the elasticities), but what determines the magnitude of the variation in \( \bar{p} \) (and, therefore, its homeostatic maintenance): the larger \( \varepsilon_p^{\text{supply}} - \varepsilon_p^{\text{demand}} \), the smaller the absolute values of both \( C_{\text{supply}} \) and \( C_{\text{demand}} \). This algebraic analysis is clearly illustrated by the different configuration of rate characteristics around the steady state shown in Fig. 5.4.

Fig. 5.4A shows a situation where the elasticities of supply and demand are equal, so that the functions of flux and concentration control are equally distributed: the same percentage change in the activity of either supply or demand causes the same change in the flux \( (C^p_{\text{supply}} = C_{\text{demand}} = 0.5) \). The magnitude of the variation in \( \bar{p} \) is determined to the same degree by supply and demand.

In Fig. 5.4B the elasticity of demand in decreased to zero (the demand becomes saturated with \( \bar{p} \)); it is clear that the demand now has complete control over the flux, while the supply has none. However, the elasticity of supply now completely determines the magnitude of the variation in \( \bar{p} \) \((C^p_{\text{supply}} = -C_{\text{demand}} = -1/\varepsilon_p^{\text{supply}})\). The steeper the slope of the supply characteristic, the narrower the band of variation in \( \bar{p} \) and, therefore, the better the homeostatic maintenance of \( \bar{p} \). The opposite would obtain if the supply elasticity were zero whereas the demand elasticity remained finite: supply would completely control flux, and the elasticity of demand would completely determine the magnitude of variation in \( \bar{p} \).

In Fig. 5.4C not only is the elasticity of demand zero (as in Fig. 5.4B) but in addition that of supply is \( -\infty \). The homeostatic maintenance of \( \bar{p} \) in the face of changes in the maximal activity of either supply or demand is now perfect; the only way in which \( \bar{p} \) can change is if the half-limiting concentration \( C_{1/2} \) of the supply block changes, as in the bottom half of Fig. 5.4C.

Supply-demand analysis therefore shows that the functions of flux and concentration control are mutually exclusive in the sense that if one block controls the flux it loses any influence over the magnitude of variation in the linking product \( \bar{p} \): this becomes the sole function of the other block. This finding has profound consequences for any view of metabolic regulation.

Up to now the analysis has been limited to the response of the steady state to small variations in the activity of supply or demand without considering either the form of the full rate characteristics or the position of the steady state in relation to equilibrium. We now expand the picture to obtain an overall view of the limits within which the system can fulfil its functions.

Like any factory, a supply pathway must be able to fulfil two primary functions: to meet increasing demand for its product at least up to some limit, and to cope with low demand in such a way that its product and intermediate metabolite concentrations do not tend towards their equilibrium concentrations (most biosynthetic pathways have huge equilibrium constants so that near-equilibrium conditions would cause fatal high accumulation of supply pathway intermediates and product [2]). Textbook wisdom has it that allosteric feedback inhibition of supply
by its product is responsible for satisfying demand, while it has little to say about low demand. What can supply-demand analysis teach us?

![Figure 5.5: The steady-state behaviour of a supply-demand system with (solid) and without (dashed) inhibition of supply by its product $P$. The grey lines represent different demand activities. The four marked steady states are discussed in the text. The rate characteristics were generated with Gepasi [51] for the supply-demand system described in [30] using the reversible Hill [33] and reversible Michaelis-Menten rate equations with realistic parameter values.](image)

Fig. 5.5 gives a bird’s-eye view of a hypothetical set of supply-demand rate characteristics spanning the full range of $p$ to its equilibrium value (assuming that the substrate for the supply pathway is buffered and therefore constant). For the supply to be able meet a specific range of variation in demand it cannot have any flux control in that range. Focussing for the moment on the solid supply curve, it is clear that only in the shaded band between steady states 2 and 3 will the supply be able to meet the variation in demand while keeping $p$ reasonably constant. When demand becomes higher that 2 it loses control over the flux (steady state 1) with a concomitant sharp decrease in $\bar{p}$. An increase in the maximal activity of the supply (the plateau at 1) would extend the range in which the supply can meet the demand. However, it is also clear that the presence of allosteric feedback inhibition is not a prerequisite for flux control by demand: in the shaded band on the right demand also controls the flux in the absence of allosteric feedback (the dashed supply characteristic), and the supply is equally effective in keeping $p$ homeostatic. The dramatic difference between the two situations is in the concentration where $P$ is homeostatically maintained: without feedback inhibition it can only be near equilibrium (with all the accompanying disadvantages), whereas with feedback inhibition it can be maintained orders of magnitude away from equilibrium (at a concentration around the $p_{0.5}$ of the allosteric enzyme). Clearly, therefore, when demand controls flux the functional role of feedback inhibition is homeostatic maintenance of $p$ at a concentration far from equilibrium.

In general each elasticity coefficient is the sum of a thermodynamic term that depends only on $\Gamma/K_{eq}$ and a kinetic term that is determined by the binding properties of the enzyme. The thermodynamic term in the supply elasticity approaches 0 in conditions far from equilibrium and $-\infty$ near equilibrium, where it completely swamps the kinetic term, which typically varies between 0 and the Hill coefficient [27]. Kinetic effects such as allosteric feedback inhibition can therefore only play any regulatory role far from equilibrium where the thermodynamic term in negligible. This is also shown by the solid curve in Fig. 5.5: there is a lower limit (around 3) to the range in which $p$ can be kinetically regulated; below this limit $p$ jumps to the region where the thermodynamic term dominates the supply elasticity.

5.4 Discussion

The central regulatory problem of metabolism is to be able to satisfy a varying demand for its products from low to high values while maintaining these products within narrow concentration ranges far from equilibrium. Supply-demand analysis shows that these two functions are inextricably linked: the more control either block has over flux, the less it determines the degree of homeostasis and the distance from equilibrium where homeostasis is maintained, which becomes the function of the other block. A common solution to this design problem in living cells is that the flux is largely controlled by the demand block, whereas the supply block determines homeostasis of the linking metabolite. Direct experimental evidence for control by demand exists (see, for example, [62, 64, 65, 47, 4]), while it can be deduced for many systems on the basis of known kinetics (in general, for example, aminoacyl-tRNA transferases have $K_m$-values for their amino-acid substrates at least an order of magnitude lower than the intracellular concentrations of amino acids, thereby ensuring that protein synthetic demand is saturated, giving a demand elasticity of zero [25]). By identifying the elasticities of supply and demand as the keys to a quantitative understanding of the integrated cellular process, supply-demand analysis provides a framework for further experimentation. A number of experimental strategies for measuring block elasticities are already available [32, 3].

Supply-demand analysis also has major implications for biotechnology [28, 10] and biomedicine and drug design [13, 9] because it shows that what were thought to be ‘rate-limiting’ steps catalysed by allosteric enzymes actually have nothing to do with flux control, but are responsible for homeostasis of metabolites. It opens a new window on our understanding of metabolic design and regulation [37].
6 Metabolic engineering of branched systems: redirecting the main pathway flux

6.1 Summary

Fluxes into secondary metabolism are usually much smaller than the primary metabolic flux they branch off from. Using metabolic control analysis general principles are derived that result from this feature. Control coefficients that quantify the control over this branch flux are expressed in elasticity coefficients for the branching metabolite, showing that the enzymes in the branch have a control of 1 on the branch flux (i.e., the flux is proportional to the enzyme activities in this branch). We show that, paradoxically, this does not make the branching enzymes necessarily the best target for increasing of the flux into the branch. It may be more effective to modulate enzymes in the main pathway. The control of the enzymes in the main pathway on the branch flux is dependent on the relative magnitude of the elasticity of the main pathway enzymes for the branching metabolite compared to the sensitivity of the enzymes in the branch. Typically when the enzymes in the main pathway have a low sensitivity for the branch-point metabolite it is better to manipulate outside the branch. The approach is illustrated with core and detailed kinetic models and shown experimentally for the branches around pyruvate in the lactic acid bacterium, *Lactococcus lactis*.

6.2 Introduction

In many biotechnological applications the flux towards the product of choice is negligible compared to the major catabolic routes. For instance in secondary metabolism and biosynthetic pathways fluxes are orders of magnitude lower than the glycolytic flux. Invariably the secondary metabolic routes depend on a metabolite in or of a primary metabolic pathway. This can be the Carbon or Nitrogen substrate, or the free energy or redox source. Pyruvate, as the substrate for diacetyl synthesis in *Lactis* may serve as an example. Because the flux through the secondary pathways is small compared to the flux through the ‘main’ pathway it branches from, it would seem that the concentration of the common metabolite can be hardly affected by modulations of the branch pathway. Here we examined whether this phenomenon could have implications for the possibilities of changing the flux through the biotechnologically relevant pathway. We found some general principles that might be profitably watched when devising strategies aimed at increasing the flux through a branch that has a small flux as compared to the main pathway it branches from. In this paper, we will first develop the new theoretical principles. Subsequently we use core models to show how the new principles can be applied. Finally we demonstrate the potential importance of the new principles in a detailed kinetic model of glycolysis in *Lactococcus lactis*.

6.3 Theory

The simplest branched system consists of three enzymes with one common intermediate, as shown in Figure 6.1. Metabolic engineering involves making changes in enzyme properties, such as the concentration of an enzyme or the affinity of an enzyme for its substrate, to improve system properties, for instance increasing the steady state flux through a branch. To relate the characteristics of an enzyme to system properties, which are a function of the properties of all the enzymes together, one needs a robust theoretical framework. A number of such system frameworks, e.g., biochemical systems theory [61], and metabolic control analysis...
MCA [40, 21, 29] (see Chapter 4) have been developed and can be used as rational tools in metabolic engineering. Here we will focus on MCA, showing some of its strengths in designing a metabolic pathway. A control coefficient of an enzyme \( i \) on the steady state flux \( J \) is defined as:

\[
C_{Ji} = \frac{dJ}{dv_i} \cdot \frac{v_i}{J}
\]

(6.1)

Via connectivity theorems the control coefficients are linked to local enzyme characteristics, the elasticity coefficients. These elasticity coefficients quantify the sensitivity of any enzyme for any variable in the system, e.g. the substrate and product concentration or allosteric regulators. An elasticity coefficient is a property of an enzyme in isolation from the system, it quantifies the percentage change in instantaneous enzyme activity upon a percentage change in the effector. The elasticity of enzyme \( i \) for effector \( x \) is defined as:

\[
\varepsilon_{xi} = \frac{\partial v_i}{\partial x} \cdot \frac{x}{v_i}
\]

(6.2)

The links between elasticity and control coefficients, between control coefficients themselves, and between control coefficients and flux ratios, make it possible to express the one in terms of the other. Thus, using MCA one can not only quantify the importance of an enzyme for a steady state property but also understand why the enzyme has that importance on the basis of its local characteristics. For metabolic engineering purposes MCA is very useful; no matter how complex a system the magnitude of the control coefficients immediately points at the enzyme of choice for genetic manipulation. In linear systems the usefulness of control coefficients is even more pronounced; since all flux control coefficients sum up to 1, one does not need to determine all of the flux control coefficients to know which of the enzymes is the most important one; a control of 0.3 means that the enzyme has 30% of all control. In branched systems, due to the existence of negative control coefficients, this simple rule does not hold (although of course all flux control coefficients for any specific flux still sum up to 1). In general, enzymes in a branch will tend to have a negative control on the flux through the other branch. Even finding a flux control coefficient of 1 for an enzyme does not mean that this is the enzyme with the highest control. Certainly with branched systems where fluxes through the branches differ greatly in magnitude [39], or with enzymes that show ultrasensitivity [48], flux control coefficients (amplification factors in [48]) can easily attain values higher than 1. The existence of negative control coefficients seems to rule out the simplifying strategies that can be applied in linear pathways, but does it really? To express control coefficients into elasticity coefficients one usually sets up a matrix of control coefficients and a matrix of elasticity coefficients such that the product of the two equals the identity matrix [72, 66]. A general method to do this is reviewed in [29] (see Chapter 4). Subsequently components of the one matrix can be expressed in components of the other via a simple matrix inversion. The control and elasticity matrix for the system shown in Figure 6.1 are listed in Appendix 1 and there it is shown how the control coefficients can be expressed in elasticity coefficients. Although these expressions look complex at first, in the specific case we study here, i.e. the flux through the branch of choice is orders of magnitude lower that the main flux (and the additional assumption that enzyme 2 is not in equilibrium, i.e. that the elasticity of that enzyme is not extremely high) the expressions can be greatly simplified (Appendix 1). The strongest simplifications can be made with respect to the control of enzyme 2; \( C_{J1} \) and \( C_{J3} \) are 0, i.e. enzyme 2 has no control on the major pathway flux and no control on the steady state concentration of \( x \). Enzyme two however does have a control on its own flux, \( C_{J2} = 1 \). Using simple principles we have come at an important conclusion; the flux control of the enzyme(s) in a branch with a flux that is negligible compared to the main flux, sums up to 1, i.e. the branch flux is proportional to the branch’s enzyme concentration. But is this now the highest flux control coefficient in the system, i.e. is this the enzyme that is the best candidate for metabolic engineering of the branch flux? To answer this question we need to look at the flux control coefficients of enzyme 1 and 3 (see Appendix 1).

\[
C_{J1}^2 = \frac{\varepsilon_{x1}^2}{\varepsilon_{x1}^2 + \varepsilon_{x3}^2}
\]

(6.3)

\[
C_{J2}^2 = \frac{-\varepsilon_{x2}^2}{-\varepsilon_{x1}^2 + \varepsilon_{x3}^2}
\]

(6.4)

Let us first consider whether these control coefficients will be positive or negative. \( \varepsilon_{x1}^2, \varepsilon_{x2}^2 \) and \( \varepsilon_{x3}^2 \) will respectively have a negative, positive and positive sign, i.e. the elasticity of a step with respect to its substrate will be positive and with respect to its product negative. Thus, \( C_{J1}^2 \) will be positive and \( C_{J2}^2 \) negative. This corresponds to the intuitive expectation that an increase in the upper part of the main pathway should enhance the branch flux whereas activation of its lower part should reduce the branch flux. For the value of the control coefficients we need to consider the magnitude of the elasticity of the enzyme in the branch with respect to \( x \) (numerator in eq. 6.3 and 6.4) relative to the magnitude of the sum of the absolute values of the elasticities of the enzymes outside the branch (denominator in eq. 6.3 and 6.4). The elasticity of a reaction for a metabolite is numerically equal to the normalized partial derivative of the rate equation with respect to that metabolite (i.e. the tangent in a log-log plot of the enzyme activity against the metabolite concentration). This value is equal to the order of the reaction with respect to the metabolite that is considered. For the Michaelis Menten rate equation this means that the value of the elasticity coefficient for the substrate is always between 1 and 0, (i.e. 1 at low substrate concentrations relative to the \( K_m \), 0 at high substrate concentrations relative to the \( K_m \)). In case of reactions that show cooperativity the
maximal value of the elasticity is numerically equal to the Hill coefficient, for low substrate concentrations, whereas for substrate concentrations above the $K_m$ value it approaches zero again. In case of reversible reactions, the elasticity is not only dependent on kinetics but also on thermodynamics; the thermodynamic component is more important for reactions that are in equilibrium, whereas the kinetic component is more important for reactions that are not in equilibrium. Thus, although elasticity coefficients can vary between $-\infty$ and $\infty$, for reactions that are in equilibrium, for most reactions the values range between 0 and 1 for substrates and 0 and -1 for products. The implication of these elasticity properties for the control of the enzymes in the main pathway on the branch flux is that the former control should be high when (i) the elasticity of the branch pathway for its substrate (X) is high and (ii) the elasticities of the reactions in the main pathway for X are small in the first condition and (iii) when the upper part of the main pathway is product insensitive and [X] exceeds the (apparent) $K_m$ of the lower part of that main pathway. Until now we have considered the most simple branched system, that of three reactions with a common intermediate. How does this effect our analysis? A method to analyze such a system is to group all of the reactions in each of the branches and then determine the so-called overall elasticity of each of the branches with respect to the common intermediate. The parameter specifications are given in Appendix 2.

### 6.4 Core Models

Metabolic control analysis points at the enzymes of choice for metabolic engineering via its control coefficients. MCA can be used to calculate what the resulting steady state values of variables will be upon an infinitesimal small perturbation, but will be less precise when larger perturbations are considered. This is because...
0.03 for \( v_1, v_2 \) and \( v_3 \). Enzyme 2 still has a high elasticity while enzyme 1 and 3 are almost insensitive for changes in \( x \). The low elasticity of enzymes 1 and 3 as compared to enzyme 2 is reflected in the flux control coefficients; enzymes 1 and 3 have now a very high flux control of 16 and -16 respectively on the flux through the branch. The complete control matrix is given in Appendix 2. Let us now compare how the system reacts to a change in the concentration of enzyme 1 and 2, as could be made by genetic engineering. We here focus on perturbations of enzyme 1 to make changes outside the branch. In principle the same effects could be brought about by decreasing the concentration of enzyme 3 but since the control of enzyme 1 and 3 is very similar, (in absolute magnitude, but of opposing sign), we only illustrate the effect of changing one of them. Under 'sensitive' conditions (condition 1), when overexpressing the activity of enzyme 2 (increasing its Vmax from 0.1 to 2.0 units/mg protein, i.e. 20 fold) and solving for steady state we obtain the results as shown in Figure 6.2. The flux through the branch (\( J_2 \)) increases proportionally with increasing activity of enzyme 2. Even at the higher activity enzyme 2 maintains its control of 1 on the flux through the branch. At low activities of enzyme 1 the enzyme has a control on \( J_2 \) similar to that of enzyme 2 but at higher activities enzyme 1 quickly loses its control. The enzyme of choice to manipulate the flux through the branch under 'saturating' conditions a very different result is obtained (figure 6.3). For enzyme 2 the situation is the same as under 'sensitive' conditions, i.e. the enzyme has a control of 1 and maintains this control. The strong positive control of enzyme 1 under 'saturating' conditions is clearly seen in the positive effect that increasing the Vmax of this enzyme has on the branch flux, surpassing the positive effect of enzyme 2. Clearly with the parameter set leading to the 'saturating' conditions the enzyme of choice to engineer would be in the main pathway. It should be stressed here that the specific plots of Figure 6.2 and 6.3 are dependent on the parameter values in the models. However a number of conclusions can be drawn that are general and independent of the parameter choice: 1) the control of enzyme 2 equals 1 independent of the kinetics, as long as the branch is not in equilibrium and the flux is low and, 2) under 'saturating' conditions the main control over the branch flux lies with the enzymes in the main pathway and this will be so over a large scan range if enzyme 1 has a high equilibrium constant.

### 6.5 Detailed models

The core model of the simple branched system can be used to illustrate the theory and test its applicability outside the immediate steady state range for which it was derived. Whether or not the theory can be used for specific experimental systems can be tested using detailed kinetic models. Such a model describes the experimental system in great detail, using parameter values specific for the system that is simulated. We will here use a detailed model for glycolysis and its branch points around pyruvate of *L. lactis*. The detailed model can be used for simulations via...
L. lactis is important for the dairy industry in milk fermentations. An important flavour compound in these fermentations is diacetyl. This arises from the进行代谢 pathway of lactate dehydrogenase and acetolactate synthase (ALS). In addition to these two pathways, there are two additional branches at the level of pyruvate: pyruvate formate lyase (PFL) and pyruvate dehydrogenase complex (PDHc). The steady-state concentrations of the metabolites are important for the control of the flux through Pyruvate branch.

The magnitude of the control coefficients is indicated by the size of the arrows in Figure 6.4. The flux control coefficient for the ALS branch is very small compared to the glycolytic flux and the flux to lactate but significantly higher than the flux via the ALS. Using these control coefficients, one can calculate the estimated control coefficients for Glycolysis, LDH, and the ALS branches. The estimated control coefficients are 2.9, -2.8, and 1.0 for Glycolysis, LDH, and the ALS branches, respectively.
ATP, ADP and phosphate interaction is not very important: with these variables "free" (in addition to NADH), the control of LDH on the ALS branch flux remains 1.9. To keep the complexity and length of this chapter manageable we will not go into detail on how to deal with systems in which there is more than one metabolite communicating between the branches, but instead give a little more detail on how the block elasticities (module response coefficients) are related to normal elasticities and show the simulation results of manipulations within the branch (ALS manipulation) and outside the branch (LDH manipulation). Overall elasticities treat a group of enzymes as if they are one enzyme and quantify the sensitivity of this "enzyme" for the metabolite under study [72]. For our system we use overall elasticities of the different branches for pyruvate. For instance the overall elasticity of glycolysis for pyruvate is defined as the percentage change in glycolytic flux upon a 1% change in the pyruvate concentration. The reference pyruvate concentration is the one obtained at steady state in the full model. Subsequently one clamps the pyruvate concentration at a 1% higher concentration and determines the effect this has on the steady state glycolytic flux. Importantly one needs to clamp all other metabolites via which the different branches communicate (in this case NADH, ADP, ATP and phosphate). By clamping a metabolite one essentially isolates the different branches in the model from each other and the clamped metabolite can be considered equivalent to an external metabolite (such as glucose). With respect to external metabolites a different coefficient is defined (the response coefficient, R) which is equivalent to our block elasticity. Just as the response coefficient can be related to an elasticity and a control coefficient, also the block elasticity can be related to the elasticity of the enzyme that is sensitive to the metabolite and the control of this enzyme. For simple branches with only one enzyme this is straightforward, e.g. the LDH branch only consists of one enzyme and the control of this enzyme on its flux (with pyruvate clamped), will be 1. Thus the block elasticity is equal to the normal elasticity. However if the branch consists of more than one enzyme, such as the ALS branch then the block elasticity of this branch for pyruvate equals the elasticity of the ALS for pyruvate multiplied by the control of the ALS on the flux through the ALS branch, \( \hat{\varepsilon}_{JALS}^{pyr} = \varepsilon_{vALS}^{pyr} \cdot c_{JALS}^{vALS} \). For glycolysis the situation is a little more complicated as pyruvate is sensed by two enzymes in the pathway, the pyruvate kinase (PK) and the phospho-enol-pyruvate dependent phosphotransferase system for glucose (PTS). Thus, \( \hat{\varepsilon}_{JGlycol}^{pyr} = \varepsilon_{vPK}^{pyr} \cdot c_{JGlycol}^{vPK} + \varepsilon_{vPTS}^{pyr} \cdot c_{JGlycol}^{vPTS} \). When we discuss the control of glycolysis on the flux through the ALS branch we assume as if glycolysis is one enzyme that can be perturbed, in reality one can approximate this by perturbing an enzyme and divide the resulting effect by the enzyme’s flux control on the glycolytic flux (with pyruvate clamped). Our analysis has shown that with wild type conditions the enzyme of choice to manipulate would be LDH, a decrease in this enzyme’s activity would result in a larger effect on ALS flux than an increase in the ALS activity. To test whether this holds also for larger changes we made a scan in ALS and LDH activities and followed the resulting flux through the ALS branch (Figure 6.5). Clearly decreasing the LDH activity leads to a greater increase in ALS flux than increasing the ALS activity (note the log scale on the y-axis).

![Figure 6.5: Branch flux via ALS as a function of the Vmax of LDH and ALS in Lactococcus lactis](image)

Both the Vmax value of the LDH and of the ALS were varied from very low activities to five fold increased activities compared to wild type activities (3300 mmoles/L cytosol/min for LDH and 600 mmoles/L cytosol/min for ALS). The flux via the ALS is plotted on logarithmic scale. Note that at the lower LDH activities the flux via the ALS is roughly 1000 fold higher as compared to the ALS flux at highest ALS activities.

### 6.6 Experiment

The detailed model as discussed in the last section consists of 28 reactions, with a full description of glycolysis and the pyruvate branches. It is a rather daunting task to start making experimental manipulations in such a large system. The aim of this chapter is to give a few general guidelines to look at before starting to overexpress any of the enzymes. In the specific case where the branch of choice has a very small flux as compared to the main pathway flux, the system can be importantly reduced in complexity using the method as we developed above in the theory section. We have shown that a description of the control of each of the branches on the branch of choice can be given via three sensitivities. These so-called block elasticities for each of the branches for the communicating metabolite were shown to be related to the elasticity of the enzyme that is sensitive for the metabolite times its control
on the branch flux this enzyme resides in. How can we use kinetic information that is available for the system, e.g. K_m values of the enzymes in the different branches for the branching metabolite, the order of these reactions with respect to the branching metabolite and the Keq values of these steps? If we have this information (or can make reasonable estimates) and have limited information on the in vivo system, such as the concentration of the branching metabolite we can make an estimate of the elasticity of the branches for the branching metabolite. Let us use the pyruvate branches in L. lactis as an example. The K_m value of the ALS for pyruvate is 50 mM. This is much higher than the K_m values of the enzymes in the main pathway PTS, and LDH, (Km values for pyruvate of 2 and 1 mM respectively) with the exception of the Km of the PK for pyruvate which is rather high too (21 mM). The highest block elasticity for pyruvate is however observed for the ALS branch, here due to the low affinity of the enzyme, and the positive cooperativity that is observed for pyruvate, a response coefficient of 2.4 is calculated (equal to the Hill coefficient). This together with a control of 1 for the ALS on the flux through its branch makes this branch highly responsive towards changes in pyruvate. With our simplified expressions we can deduce that the main control over the ALS branch flux should reside outside the pathway. Although the control of the ALS on the flux through its own branch will be 1 we can calculate that the sum of the block elasticities of glycolysis and LDH are lower than the block elasticity of the ALS branch. Thus on basis of these simple principles, and assuming that pyruvate is the main metabolite via which these branches interact, we can predict that the control on the ALS flux will mainly reside outside the branch. A number of metabolic engineering strategies have been directed at increasing the flux through the ALS branch[52, 53, 68, 11]. The most intuitive approach to follow would be to increase the activity of the ALS via overexpression of the als gene. Since this enzyme has a control of 1 on the branch flux one would expect to see a significant increase in the pathway flux upon increasing the activity of the enzyme. Indeed the pathway flux does increase upon overexpressing the als gene but even upon a 20 fold increased activity the pathway flux still is insignificant compared to the main pathway flux [53]. As was expected from our analysis a much bigger effect was observed upon downregulation of the LDH [53, 26]. In an LDH minus strain a significant portion of the main flux was rerouted via the ALS pathway (114 mmol acetoin/L internal volume/min of 290 mmol glucose/L internal volume/min), [Hofnung et al., 2002].

The kinetic model predicted that a knock-out in the LDH would greatly impede the glycolytic flux (73% reduction), which was observed experimentally albeit to a lesser extent (11% reduction). The next step in the genetic manipulation to increase the flux via the ALS pointed at the NADH oxidase and culturing under aerobic conditions. These experiments fall beyond the scope of this chapter and are related to there being more than one metabolite via which the branches interact.

### 6.7 Conclusions

Here our goal was to illustrate the importance of theory, modelling and experiment for a rational approach to metabolic engineering of biological systems. We specifically addressed the problem of how to increase very low fluxes in branched pathways. Using MCA we could derive that the flux control within such a branch is always 1 on its own flux. The extent to which flux control outside the branch exceeds this control was shown to be dependent on the relative overall elasticities of the branches for the common branching metabolite. If the sum of the block elasticities of the main branches is lower than the block elasticity of the branch under study than the majority of the control lies outside the pathway. This general theory was illustrated using a core model of a branched pathway and a detailed model of L. lactis glycolysis and pyruvate branches. The latter system was also experimentally shown to be in agreement with the general principles derived here.

### 6.8 Appendix 1

For the system depicted in Figure 6.1 the control (C) and elasticity (E) matrix, set up according to the method of [72] and reviewed in [29] (see Chapter 4) are:

\[
C = \begin{bmatrix}
C_{v_1}^{C_1}
& C_{v_1}^{C_2}
& C_{v_1}^{C_3}

C_{v_2}^{C_1}
& C_{v_2}^{C_2}
& C_{v_2}^{C_3}

C_{v_3}^{C_1}
& C_{v_3}^{C_2}
& C_{v_3}^{C_3}
\end{bmatrix}
\]

\[
E = \begin{bmatrix}
1 & 0 & -\varepsilon_{E_{v_1}}^{C_1}
0 & 1 & -\varepsilon_{E_{v_2}}^{C_2}
\frac{J_1}{J_2} & \frac{J_2}{J_1} & -\varepsilon_{E_{v_3}}^{C_3}
\end{bmatrix}
\]

The matrix multiplication C - E yields the identity matrix I. The control matrix can be expressed in elasticities via a matrix inversion of E:

\[
C = \begin{bmatrix}
J_1 J_2 + J_1 J_3
& J_1 J_2 + J_2 J_3
& J_1 J_2 + J_3 J_1

-J_1 J_2 + J_1 J_3
& J_1 J_2 - J_2 J_3
& J_1 J_2 - J_3 J_1

-J_1 J_3 + J_2 J_1
& J_1 J_3 - J_2 J_1
& J_1 J_3 - J_3 J_1
\end{bmatrix}
\]
The expression of the control coefficients in elasticities appears to be rather complex but if we realize that 1) the specific pathway that we are interested in has a flux $J_2$ orders of magnitude smaller than $J_1$ and $J_3$, and 2) that $J_1$ can be expressed as $J_1 = J_2 J_3$, and if we assume that 3) the elasticity of enzyme 2 is not extremely large (i.e. the enzyme is not close to equilibrium) then $C$ reduces significantly:

$$C \approx \begin{bmatrix} \frac{\epsilon x_3}{x_3 J_2^2} & 0 & \frac{\epsilon x_3}{x_3 J_2 J_3} \\ \frac{\epsilon x_2}{x_2 J_1 + x_2 J_3} & 1 & \frac{\epsilon x_2}{x_2 J_2} \\ \frac{\epsilon x_1}{x_1 J_1 + x_1 J_3} & 0 & \frac{\epsilon x_1}{x_1 J_1} \end{bmatrix}$$

6.9 Appendix 2

A kinetic model for the system depicted in Figure 6.1 was constructed. The system was described with the following ordinary differential equation:

$$\frac{d(x)}{dt} = v_1 - v_2 - v_3$$

The enzymes were modelled using reversible Michaelis Menten type of reactions:

$$v_i = \frac{V_{\text{max},i} \frac{1}{K_{\text{m},i}} (s - \frac{p}{K_e})}{(1 + \frac{x_i}{K_{\text{m},i}} + \frac{p}{K_e})}$$

with $s$ and $p$ the substrate and product concentrations of the respective reactions $i$. For the 'sensitive' conditions (condition 1) the following parameter values were used: $s = 1; p_i = 0; \; p_2 = 0;$ $V_{\text{max},1} = V_{\text{max},3} = 4.5606; \; V_{\text{max},2} = 0.1; \; K_1s = 1; \; K_1x = K_3x = 0.178; \; K_2x = 100; \; K_2p = K_3p = 1; \; K_{\text{eq}1} = 0.1; \; K_{\text{eq}2} = 1000; \; K_{\text{eq}3} = 10$; For 'saturating' conditions (condition 2) the following parameter values were used: $s = 10; p_i = 0; \; p_2 = 0; \; V_{\text{max},1} = V_{\text{max},3} = 1.03221; \; V_{\text{max},2} = 0.1; \; K_1s = 0.01; \; K_1x = K_3x = 0.0016; \; K_2x = 100; \; K_2p = K_3p = 1; \; K_{\text{eq}1} = 1000; \; K_{\text{eq}2} = 1000; \; K_{\text{eq}3} = 10$; Metabolite concentrations in mM, binding constant for substrates and products in mM and Vmax values in units/mg cell protein. Steady state concentration of X under the two conditions was 0.05 mM and the flux values: $J_1 = 1.00, J_2 = 1.00 \times 10^{-4}, J_3 = 1.00$ units mg$^{-1}$ cell protein. The control matrix of the system under 'sensitive' conditions is:

$$C = \begin{bmatrix} C_{v_1}^J & C_{v_2}^J & C_{v_3}^J \\ 0.4102 & 0.0000 & 0.5898 \\ C_{v_1}^J & C_{v_2}^J & C_{v_3}^J \\ 0.5251 & 1.0000 & -0.5251 \\ C_{v_1}^X & C_{v_2}^X & C_{v_3}^X \\ 0.5254 & -0.0000 & -0.5254 \end{bmatrix}$$


Transcriptome meets metabolome: hierarchical and metabolic regulation of the glycolytic pathway

Benno H. ter Kuile$^{a,*}$, Hans V. Westerhoff$b$

$^a$Leiden Institute of Chemistry, P.O. Box 9502, 2300 RA Leiden, The Netherlands
$^b$Molecular Cell Physiology and Mathematical Biochemistry, BioCentrum Amsterdam, De Boelelaan 1087, 1081 HV Amsterdam, The Netherlands

Received 3 May 2001; revised 11 June 2001; accepted 13 June 2001
First published online 22 June 2001

Edited by Judit Ovádi

Abstract The fact that information flows from DNA to RNA to protein to function suggests that regulation is ‘hierarchical’, i.e. dominated by regulation of gene expression. In the case of dominant regulation at the metabolic level, however, there is no quantitative relationship between mRNA levels and function. We here develop a method to quantitate the relative contributions of metabolic and hierarchical regulation. Applying this method to the glycolytic flux in three species of parasitic protists, we conclude that it is rarely regulated by gene expression alone. This casts strong doubts on whether transcriptome and proteome analysis suffices to assess biological function.

Key words: Metabolic control; Hierarchical control; Metabolic flux; Regulation of transcription

1. Introduction

The usually unspoken assumption underlying all genomic studies is that the expression of a gene at the mRNA level is a quantitative predictor for function. In short, an $X$-times higher signal at the Northern blot or gene chip level means $X$-times more protein and $X$-times more function. In particular, in the case of enzymatic reactions that are part of a metabolic pathway, this assumption may not reflect reality. At any time the rate of an enzymatic reaction is a function of the substrate(s), product(s), modifier(s), so-called metabolic properties and of gene expression, changing the concentration of the enzyme in question. Here we define ‘metabolic’, as all changes caused by concentrations of substrate(s), product(s) and modifier(s). The ‘hierarchical’ changes are those caused by change in enzyme concentration, via alterations in mRNA sequestration and intracellular localization and/or rates of whichever process, be it transcription, translation or degradation. Below, we consider regulation of steady-state fluxes through enzymes by either class of properties. We divide the change in the rate of an enzyme-catalyzed reaction, that is not directly affected by the external regulator, by the change in flux through the pathway to which the enzyme belongs. At steady state, the resulting ratio is 1, and distributed over hierarchical and metabolic regulation:

$$l = \rho_h + \rho_m$$

(1)

$\rho_b$ and $\rho_m$ are the products of co-response coefficients and elasticity [1] quantifying the hierarchical (gene expression) and the metabolic component of the regulation, respectively. The hierarchical regulation coefficient is defined as the relative change in enzyme-catalyzed rate divided by the relative change in enzyme concentration (this factor usually equals 1 as expressed by the $\approx$ sign below) multiplied by the ratio of the change in enzyme concentration to the change in flux. Although this is not essential, all changes are here taken relative and small, such that the corresponding mathematical equation for the hierarchical regulation coefficient of an enzyme $i$ is:

$$\rho_h = \frac{\delta \ln v_i}{\delta \ln e_i} \frac{d \ln e_i}{d \ln J} \approx \frac{d \ln e_i}{d \ln J}$$

(2)

$J$ represents the pathway flux, $v_i$ is the concentration of any enzyme $i$ through which the flux runs, at a rate $v_i$. The metabolic component of the regulation is given by the change in enzyme rate divided by the change in concentration ($\delta$) of any of the varying metabolites around it, multiplied by the ratio of change in that concentration to the change in flux. This should be summed over all metabolites that change:

$$\rho_m = \sum_X \frac{\delta \ln v_i}{\delta \ln X} \frac{\delta \ln X}{\delta \ln J}$$

(3)

Graphically, the slope of a double logarithmic plot of the enzyme concentration (activity) versus the flux, which corresponds to a co-response coefficient [1], will give the hierarchical regulation coefficient, $\rho_h$. By virtue of the above relationship, the metabolic component can be obtained by subtraction of the $\rho_h$ from 1 (100%). For full derivation see [2]. The regulation of each enzyme of a pathway can thus be compared with that of the others to provide a sophisticated picture of the regulation of the entire pathway. This type of analysis is complementary to metabolic control analysis [3,4], which considers pathways at constant gene expression or in long-term steady states after manipulation by genetic means.

2. Materials and methods

The species used and the methodology are described in references

---

Corresponding author. Fax: (31)-71-5274537.
E-mail: b.terkuile@chem.leidenuniv.nl
E-mail: hw@bio.vu.nl

0014-5793/01/$20.00 \copyright$ 2001 Federation of European Biochemical Societies. Published by Elsevier Science B.V. All rights reserved.

PII: S0014-5793(01)02613-8
Further analysis was performed using Excel (Microsoft) software.

3. Results

This approach was applied to an experimental example that is particularly suited for such analysis: a study of well defined steady states concerning regulation and adaptation of the glycolytic pathway of three species of parasitic protists, *Trypanosoma brucei*, *Leishmania donovani* and *Trichomonas vaginalis*, growing on glucose [5,6]. Each species was grown in chemostats at approximately 0.2, 0.4, 0.6, 0.8 and 0.95 times the maximum growth rate under glucose limitation. At each steady state, the magnitudes of fluxes and the activities of most enzymes of the pathway were documented.

3.1. Enzyme activity and flux

Fig. 1 is a double logarithmic plot of enzyme activities versus the accompanying flux. ‘Enzyme activity’ is defined throughout as the in vitro rate at saturating substrate concentrations. The enzymes shown were selected to illustrate possible ways of regulation. Glycerol-3-phosphate dehydrogenase of *T. brucei*, and phosphoglucoseisomerase (PGI) of *L. donovani* showed completely metabolic regulation: whereas flux changed more than an order of magnitude, there was little change in activity of the enzyme. Therefore the $b_h$ amounted to 0.0, hence $b_m$ to 1.0. Apparently, the higher flux through these metabolic steps was entirely due to changes in metabolite concentrations; changes in expression level were irrelevant for the regulation of flux through this enzyme with increasing growth rate. This was also the case when glucose was supplied in excess and growth rate was limited by another nutrient (data not shown). Completely hierarchical regulation ($\rho_h \approx 1.0$) occurred in the case of malic enzyme of *L. donovani*.

Regulation can also be neither totally hierarchical nor totally metabolic. The extent to which metabolic flux through an enzyme is regulated hierarchically or metabolically differs between organisms and between enzymes. Regulation of the flux through the PGI and inorganic pyrophosphate-dependent phosphofructokinase (PPi-PFK) steps of *T. vaginalis* illustrates this. The average $\rho_h$ of PPi-PFK (as calculated from the slope in the ln-ln plot) was close to two-thirds, one-third remaining for metabolic regulation. PGI in *T. vaginalis* was regulated hierarchically for only 25%, while other enzymes showed different distributions between 0 and 1 (Table 1).

3.2. Shifts in regulation

Biology is complex and methods of analysis must take this into account: within a single organism, for a single enzyme, the method of regulation can shift as regulation proceeds. This was observed for glyceraldehyde-3-phosphate dehydrogenase (GAPDH) of *T. brucei*; a shift from metabolic regulation...
at low dilution rates (average \( \rho_b \) close to 0) to overcompensating hierarchical control at high dilution rates (average \( \rho_b \) exceeding 1) was observed. It is as if the organism uses metabolic regulation to upregulate the GAPDH reaction until that mechanism is no longer sufficient and then invokes hierarchical regulation to achieve even higher fluxes through this step. The switch between the two types of regulation appeared to occur over a limited range of glucose consumption rates.

Thus far, hierarchical regulation analysis helped us to appreciate the relative importance of metabolic and hierarchical regulation. However, living systems can be more surprising than this. Negative hierarchical regulation at an average \( \rho_b \) of −0.4, hence a metabolic regulation of 1.6, was observed in the case of regulation of the flux through the PFK step in *L. donovani*. Negative hierarchical control was also noticed for several more enzymes of *T. brucei* and *L. donovani*. This may reflect the phenomenon that in the latter species, synthesis rates of some glycolytic enzymes are independent of growth rate. The lower activities at higher growth rates are caused by the enzyme being ‘diluted’ by other cell material [6].

The other extreme of regulation was also found, exemplified by GAPDH in *T. brucei* at high glucose consumption rates. Here \( \rho_b \) becomes as high as 1.8, the metabolic regulation becoming negative at approximately −0.8. This we interpret as extensive positive adaptation by the organism; when it sees more glucose, it enhances the relevant gene-expression more than proportionately. Metabolic regulation then runs behind to the extent that it acts homeostatically.

### 3.3. Distribution of types of regulation

The above being mere examples of different scenarios of regulation, we did a more complete analysis for glycolytic enzymes in three organisms. The following types of regulation were found more than twice: largely metabolic enzymes in three organisms. The following types of regulation, we did a more complete analysis for glycolytic enzymes. Metabolic regulation then runs behind to the extent that it acts homeostatically.

First, even at steady state, message levels did not correlate with enzyme activities in the three species used for this study [5,6]. Secondly, when one does not study the system at steady state, the enzyme concentrations may not yet have relaxed to their new steady-state values, whereas the mRNA concentrations have. This more trivial explanation is not relevant to steady-state chemostat cultures. Another explanation for the absence of correlation between mRNA levels and fluxes has been regulation at the posttranslational-modification level. Because we compared enzyme activities to fluxes, this cannot explain our findings either. We conclude that the both transcriptome and proteome may be vastly incomplete monitors of regulation of cell function. The transcriptome [7,8] must also be taken in consideration in functional genomics.

That perhaps being a negative note, there are positive notes as well. Firstly, this paper shows a new and quantitative way to evaluate the importance of metabolic regulation by demonstrating that it equals 1 minus hierarchical regulation, provided that either type of regulation is expressed in terms of regulation coefficients. Secondly, it reveals that the extent to which regulation is hierarchical versus metabolic can vary with growth conditions of the organism. Consequently, it may enable one to find conditions where regulation is largely hierarchical and transcriptome/proteome analysis should be telling. When regulation is shown to be completely metabolic, there is also no need to search for regulatory mechanisms at the transcriptional or translational level. In addition, knowing that a particular enzyme is regulated largely metabolically may make it worthwhile to undertake the more tedious metabolic analyses, because the latter can then be focused on fewer enzymes.

**Acknowledgements:** We thank David Fell, Les Grivell, Jannie Hofmeyr and Fred Opperdoes for critical reading of an earlier version of the manuscript. This study was supported by the Leiden Institute of Chemistry and the Netherlands Association of Biotechnology Research Schools.

### References