Modular bond-graph modelling and analysis of biomolecular systems

Peter J. Gawthrop1,2,3✉, Edmund J. Crampin1,3,4,5,6

1Systems Biology Laboratory, Melbourne School of Engineering, University of Melbourne, Victoria 3010, Australia
2Department of Electrical and Electronic Engineering, Melbourne School of Engineering, University of Melbourne, Victoria 3010, Australia
3Centre for Systems Genomics, University of Melbourne, Victoria 3010, Australia
4School of Mathematics and Statistics, University of Melbourne, Victoria 3010, Australia
5School of Medicine, University of Melbourne, Victoria 3010, Australia
6ARC Centre of Excellence in Convergent Bio-Nano Science, Melbourne School of Engineering, University of Melbourne, Victoria 3010, Australia
✉ E-mail: peter.gawthrop@unimelb.edu.au

Abstract: Bond graphs can be used to build thermodynamically-compliant hierarchical models of biomolecular systems. As bond graphs have been widely used to model, analyse and synthesise engineering systems, this study suggests that they can play the same rôle in the modelling, analysis and synthesis of biomolecular systems. The particular structure of bond graphs arising from biomolecular systems is established and used to elucidate the relation between thermodynamically closed and open systems. Block diagram representations of the dynamics implied by these bond graphs are used to reveal implicit feedback structures and are linearised to allow the application of control-theoretical methods. Two concepts of modularity are examined: computational modularity where physical correctness is retained and behavioural modularity where module behaviour (such as ultrasensitivity) is retained. As well as providing computational modularity, bond graphs provide a natural formulation of behavioural modularity and reveal the sources of retroactivity. A bond graph approach to reducing retroactivity, and thus inter-module interaction, is shown to require a power supply such as that provided by the ATP ⇌ ADP + Pi reaction. The mitogen-activated protein kinase cascade (Raf–MEK–ERK pathway) is used as an illustrative example.

1 Introduction

In their review paper, The rôle of control and system theory in systems biology, Wellstead et al. [1] suggest that ‘systems biology is an area where systematic methods for model development and analysis, such as bond graphs, could make useful new contributions as they have done in the physical world’. The purpose of this paper is to show that bond graphs not only provide a systematic method for model development and analysis of biomolecular systems, but also provide a bridge allowing application of control engineering methodology, in particular feedback concepts, to systems biology.

Bond graphs were introduced by Paynter [2] and their engineering application is described in number of text books [3–6] and a tutorial for control engineers [7]. Bond graphs were first used to model chemical reaction networks by Oster et al. [8] and a detailed account is given by Cellier [10], Thoma and Mocellin [11] and Greifeneder and Cellier [12]. More recently, the bond graph approach has been used to analyse biochemical cycles by Gawthrop and Crampin [13] and has been shown to provide a modular approach to building hierarchical biomolecular system models which are robustly thermodynamically compliant [14]; combining thermodynamically compliant modules gives a thermodynamically compliant system. In this paper, we will call this concept, computational modularity.

Computational modularity is a necessary condition for building physically correct computational models of biomolecular systems. However, computational modularity does not imply that module properties (such as ultrasensitivity) are retained when a module is incorporated into a larger system. In the context of engineering, modules often have buffer amplifiers at the interface so that they have unidirectional connections and may thus be represented and analysed on a block diagram or signal flow graph where the properties of each module are retained. This will be called behavioural modularity in this paper. However, biological networks do not usually have this unidirectional property, but rather display retroactivity [15–20]; retroactivity modifies the properties of the interacting modules. As will be shown, the property of retroactivity is naturally captured by bond graphs. In particular, a bond graph approach to reducing retroactivity, and thus inter-module interaction, is discussed and shown to require a power supply such as that provided by the ATP ⇌ ADP + Pi reaction.

Early attempts at modelling the mitogen-activated protein kinase (MAPK) cascade [21, 22], used modules which displayed behavioural modularity. However, because they use the Michaelis–Menten approximation, the modules do not have the property of computational modularity and thus the results were based on a non-physical model. This was noted in later work which examined the neglected interactions: in particular, Ortega et al. [23] show that ‘product dependence and bifunctionality compromise the ultrasensitivity of signal transduction cascades’ and the ‘effects of sequestration on signal transduction cascades’ are considered by Bluthgen et al. [24]. In this paper, the MAPK cascade is used as an illustrative example which illustrates how a computationally modular approach based on bond graphs avoids the errors associated with assuming irreversible Michaelis–Menten kinetics. Moreover, the bond graph approach to reducing retroactivity is used to make the modules approximately modular in the behavioural sense. This emphasises the necessity for a power supply to support signalling networks in biology as well as in engineering.
The bond graph approach gives the set of non-linear ordinary differential equations describing the biomolecular system being modelled. Linearisation of non-linear systems is a standard technique in control engineering; as discussed by Goodwin et al. [25], ‘The incentive to try to approximate a non-linear system by a linear model is that the science and art of linear control is vastly more complete and simpler than they are for the nonlinear case’. Nevertheless, it is important to realise that conclusions drawn from linearisation can only be verified using the full non-linear equations. In the context of bond graphs, linearisation (and the associated concept of sensitivity) has been treated by the authors [26–28]. This paper builds on this work to explicitly derive the bond graph corresponding to the linearised non-linear system and thus provide a method to analyse behavioural modularity.

Section 2 briefly shows how biomolecular systems can be modelled using bond graphs. Section 3 shows how thermodynamically closed systems can be converted to thermodynamically open systems using the twin notions of chemostats and flowstats. Linearisation is required to understand module behaviour, and this is developed in Section 4. Section 5 looks at modularity, retroactivity and feedback and Section 6 illustrates the main results using the MAPK cascade example. Section 7 concludes the paper and suggests future research directions.

2 Bond graph modelling of biomolecular systems

As discussed by Maxwell [29], the use of ‘mathematical or formal analogy’ enables us to avail ‘ourselves of the mathematical labours of those who had already solved problems essentially the same.’ The bond graph approach provides a systematic approach to the use of analogy in the modelling of systems across different physical domains; in the context of this paper, this allows engineering concepts to be carried across to biomolecular systems. A number of text books about bond graphs [3–6] and a tutorial for control engineers [7] are available. Briefly, bond graphs focus on a pair of variables generically termed effort e and flow f whose product is power $p = ef$. In the electrical domain, effort is identified with voltage $V$ (V) and flow with current $i$ (C·s$^{-1}$) and in the mechanical domain effort is identified with force $F$ (N) and flow with velocity $v$ (m·s$^{-1}$). Thus voltage and force are effort analogues and current and velocity are flow analogues. Although the effort (and the flow) variables have different units in each domain, their product (power) has the same units (W or J·s$^{-1}$); power is the common currency of disparate physical domains. The pair e f is represented on the bond graph by the harpoon symbol: — which can be optionally annotated with specific effort and flow variables, for example $\frac{v}{i}$. Sign convention is handled by the harpoon direction: thus if e and f are positive, the flow f is in the harpoon direction.

As well as analogous variables, bond graphs deal in analogous components. Thus the bond graph C component models both the ideal electrical capacitor (with capacitance $c_e$) and the ideal mechanical spring (with stiffness $K$). In both cases, the C component physically accumulates flow to give the integrated flow $q$ corresponding to electrical charge or mechanical displacement. In the linear case, this gives an effort proportional to $q$. To summarise

$$\dot{Q} = i, \quad V = \frac{Q}{c_q} \quad \text{(electrical)}$$

$$\dot{q} = f, \quad e = Kq \quad \text{(generic)}$$

Similarly, electrical resistors and mechanical dampers are represented by bond graph R components where

$$V = ri \quad \text{(electrical)}$$

$$F = rv \quad \text{(mechanical)}$$

$$e = rf \quad \text{(generic)}$$

where r represents the (linear) electrical resistance and mechanical damping factor.

Bonds are connected by 0 and 1 junctions, which again conserve energy; the 0 junction gives the same effort on each impinging bond and the 1 junction gives the same flow on each impinging bond.

Fig. 1 shows a simple electrical circuit connecting two capacitors with capacitance $c_1$ and $c_2$ by a resistor with resistance $r_1$. In the linear case, this gives an ordinary differential equation which represents the two equations

$$f_1 = p_1, \quad i_1 = e_1$$

Note that energy is conserved as

$$p_2 = e_2f_2 = \frac{e_1}{r_1}p_1 = e_1f_1 = p_1$$

2.1 Biomolecular bond graph components

It is assumed that biochemical reactions occur under conditions of constant pressure (isobaric) and constant temperature (isothermal). Under these conditions, the chemical potential $\mu_A$ of substance $A$ is given [30] in terms of its mole fraction $\chi_A$ as

$$\mu_A = \mu_A^* + RT\ln\chi_A(J\cdot mol^{-1})$$

where the standard chemical potential $\mu_A^*$ is the value of $\mu_A$ when $A$
The key to modelling chemical reactions by bond graphs is to determine the appropriate effort and flow variables. As discussed by Oster et al. [8, 9], the appropriate effort variable is the chemical potential \( \mu \) and the appropriate flow variable is molar flow rate \( v \).

In the context of chemical reactions, the bond graph \( C \) component of (1) is defined by (7) as

\[
\dot{x}_A = v_A (\text{mol} \cdot \text{s}^{-1}) \quad \mu_A = RT \ln K_A x_A (\text{J} \cdot \text{mol}^{-1}) \quad (\text{chemical})
\]

where \( x_A \) is the molar amount of \( A \) and \( K_A \) is the equilibrium constant for the reaction of \( A \) defined by (8) as

\[
K_A = \frac{1}{n_{\text{total}}} \exp \left( \frac{\mu_A^*}{RT} \right) = \frac{1}{n_{\text{total}}} \exp \left( \frac{\mu_A}{RT} \right) \quad (\text{mol}^{-1})
\]

where \( n_{\text{total}} \) is the total number of moles in the mixture. Alternatively, (9) can be written more simply in terms of the normalised chemical potential \( \tilde{\mu}_A \) of (8) as

\[
\dot{x}_A = v_A (\text{mol} \cdot \text{s}^{-1}) \quad \tilde{\mu}_A = \ln K_A x_A \quad (\text{11})
\]

We follow Oster et al. [9] in describing chemical reactions in terms of the Marcelin–de Donder formulae as discussed by Van RysSELBERGE [32] and Gawthrop and Crampin [13]. In particular, given the \( i \)th reaction \( \sum_{j} v_i^j A_j + \sum_{k} v_i^k B_k = \sum_{r} v_i^r C_r + \sum_{s} v_i^s D_s \) of (12) as

\[
v_i^j A_j + v_i^k B_k = v_i^r C_r + v_i^s D_s \quad (\text{12})
\]

where the stoichiometric coefficients \( v_i \) are either zero or positive integers, the forward affinity \( A_i^f \) and the reverse affinity \( A_i^r \) are defined as

\[
A_i^f = v_i^j \mu_A + v_i^k \mu_B + v_i^r \mu_C + v_i^s \mu_D \quad \text{and} \quad A_i^r = v_i^j \mu_A + v_i^k \mu_B + v_i^r \mu_C + v_i^s \mu_D \quad (\text{13})
\]

The units of affinity are the same as those of chemical potential: \( \text{J} \cdot \text{mol}^{-1} \). Again, normalised affinities are useful

\[
\tilde{A}_i^f = \frac{A_i^f}{RT}, \quad \tilde{A}_i^r = \frac{A_i^r}{RT} \quad (\text{15})
\]

The \( i \)th reaction flow \( v_i \) is then given by

\[
v_i = \kappa_i (v_i^0 - v_i^0) \quad \text{where} \quad v_i^0 = e^\kappa_i (\text{mol} \cdot \text{s}^{-1}) \quad (\text{16})
\]

Note that the arguments of the exponential terms are dimensionless as are \( v_i^0 \) and \( v_i^0 \). The units of the reaction rate constant \( \kappa_i \) are those of molar flow rate: \( \text{mol} \cdot \text{s}^{-1} \).

The nth reaction flow \( v_n \) depends on the forward and reverse affinities \( A_i^f \) and \( A_i^r \), but cannot be written as the difference between the affinities. Unlike the electrical \( R \) component (see Fig. 1), it cannot be written as a one port component with the flow dependent on the difference between the efforts. However, as discussed by Gawthrop and Crampin [12], a two port resistive component, the \( R \) component, can be used to model the reaction (16).

The fact that the capacitive \( C \) and resistive \( Re \) components are intrinsically non-linear is one factor distinguishing biochemical systems from the electrical and mechanical systems of (1).

The \( TF \) component is used in this context to account for any non-unity and non-zero stoichiometric coefficients \( v \) in (12) [8, 9, 13]. Moreover, as will be discussed in the following section, the \( TF \) component can be used to abstract the entire network of bonds, 0 and 1 junctions connecting the \( C \) and \( Re \) components.

2.2 Examples

Consider the simple reaction \( A = B \). In this case, \( \tilde{A}^f = \tilde{\mu}_A \) and \( \tilde{A}^r = \tilde{\mu}_B \). With reference to Fig. 1c, substance \( A \), \( B \), and the reaction by \( \text{Re}_r \). The equations of the \( C \) components correspond to (9) and that of the \( \text{Re}_r \) component to (16). The equations are

\[
\tilde{A}^f = \tilde{\mu}_A = \ln K_A x_A, \quad \tilde{A}^r = \tilde{\mu}_B = \ln K_B x_B \quad (\text{17})
\]

\[
v = \kappa (K_A x_A - K_B x_B) = k^+ x_A - k^- x_B \quad (\text{18})
\]

\[
K_{\text{eq}} = \frac{k^+}{k^-} = \frac{K_A}{K_B} \quad (\text{20})
\]

and is thus a function of the thermodynamic constants \( K_A \) and \( K_B \), but not the reaction rate-constant \( \kappa \). A general formula relating all the equilibrium constants in a biomolecular network to the rate constants is given by Gawthrop et al. [14, §3].

The enzyme catalysed reaction

\[
A + E \xrightarrow{1 \, \text{C}} 2 = B + E \quad (\text{21})
\]

where \( A \) is the reactant, \( B \) is the product, \( C \) is the intermediate complex and \( E \) is the enzyme, is ubiquitous in biochemical systems. Reaction (21) was first modelled using bond graphs by Oster et al. [9, Fig. 5.9].

Fig. 2a shows the bond graph corresponding to (21). The components \( \text{Re}_r_1 \) and \( \text{Re}_r_2 \) represent the reactions \( 1 \) and \( 2 \) and the four \( C \) components \( C.A, C.B, C.C \) and \( C.E \) represent the four species \( A, B, C \) and \( E \). The left-hand 1 junction ensures that the flow out of \( C.A \) and \( C.E \) is the reaction flow \( v_1 \) and the right-hand 1 junction ensures that the flow into \( C.B \) and \( C.E \) is the reaction flow \( v_2 \). The net flow into \( C.E \) is thus \( v_1 - v_2 \)

The additional reaction \( \text{Re}_r_0 \) has been added in Fig. 2b together with the zero-potential source \( \phi \); this can be used to model enzyme degradation. \( C.A \) and \( C.B \) are used in Section 3 as an example of a chemostat and \( \text{Re}_r_0 \) as an example of a flowstat. The enzyme catalysed reaction is analysed further in Section 5.2.

3. Closed systems and open systems: chemostats and flowstats

Specific bond graphs (such as Figs. 1c and 2a) model specific sets of chemical reactions. It is convenient to generalise such bond graphs to
allow generic statements to be made and generic equations to be written. The molar amounts of the \( n_X \) species \( x_A, x_B, \ldots \), the corresponding chemical potentials \( \mu_A, \mu_B, \ldots \) and the corresponding thermodynamic constants \( K_A, K_B, \ldots \) are collected into column vectors

\[
X = \begin{pmatrix} x_A \\ x_B \\ \vdots \end{pmatrix}, \quad \mu = \begin{pmatrix} \mu_A \\ \mu_B \\ \vdots \end{pmatrix}, \quad K = \begin{pmatrix} K_A \\ K_B \\ \vdots \end{pmatrix}
\]  

(22)

Similarly, the \( n_V \) reaction flows \( v_1, v_2, \ldots \), affinities (forward and reverse) \( A_1, A_2, \ldots \) and the corresponding reaction constants \( \kappa_1, \kappa_2, \ldots \) are collected into column vectors

\[
V = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \end{pmatrix}, \quad A = \begin{pmatrix} A_1 \\ A_2 \\ \vdots \end{pmatrix}, \quad \kappa = \begin{pmatrix} \kappa_1 \\ \kappa_2 \\ \vdots \end{pmatrix}
\]  

(23)

As discussed by Karnopp et al. [6], the \( C \) components can be subsumed into a single \( C \)-field, the \( R \) components (as two-port \( R \) components) subsumed into an \( R \)-field and the connecting bonds, \( 0 \) and \( 1 \) junctions subsumed into a junction structure. Moreover, as this junction structure transmits, but does not store or dissipate energy, it can be modelled as the two multiport transformers \( TF:N_f \) and \( TF:N_r \) as shown in Fig. 3a. These two multiport

---

**Fig. 2** Example: Enzyme catalysed reaction. (a) Bond graph of the enzyme catalysed reaction \( A + E \xrightarrow{1} C \xrightarrow{2} B + E \). \( A \) is the reactant, \( B \) the product, \( C \) the intermediate complex and \( E \) the enzyme. (A and B are used as chemostats given in Section 3). (b) An enzyme degradation reaction \( Re:R_0 \) is added. (\( Re:R_0 \) is used as a flowstat given in Section 3)

a Bond graph  
b Bond graph with enzyme degradation

---

**Fig. 3** Closed and open systems: (a) General closed system. Bond symbols \( S \) correspond to vectors of bonds; \( C, Re \) and \( O \) correspond to arrays of \( C, Re \) and \( O \) components; the two \( TF \) components represent the intervening junction structure comprising bonds, \( 0 \) and \( 1 \) junctions and \( TF \) components. \( N_f \) and \( N_r \) are the forward and reverse stoichiometric matrices. (b) Corresponding block diagram. (c) Addition of the chemostat and flowstat flows \( V^* \) to the closed system of Fig. 3a gives an open system. (d) Corresponding block diagram

a Closed system bond graph  
b Closed system block-diagram  
c Open system bond graph  
d Open system block diagram
transformers are defined to transform flows as

\[ \dot{X}^I = N^I V, \quad \dot{X}^F = N^F V \] (24)

As they do not store or dissipate energy, it follows that the affinities are given by

\[ A^I = N^I T \mu, \quad A^F = N^F T \mu \] (25)

As discussed by Gawthrop and Crampin [13], and with reference to Fig. 3a, the system states \( X \) correspond to the molar amounts of each species stored in each \( C \) component and are given in terms of the reaction flows \( V \) as

\[ \dot{X} = X^I - X^F = NV \quad \text{where} \quad N = N^I - N^F \] (26)

\( N \) is the stoichiometric matrix \([33]\); \( N^I \) and \( N^F \) are referred to as the forward and backward stoichiometric matrices.

From (9), the composite chemical potential \( \mu \) is given by the non-linear equation [Following [34], we use the convenient notation \( \text{Exp}\ X \) to denote the vector whose \( i \)th element is the natural logarithm of the \( i \)th element of \( X \)]

\[ \frac{\mu}{RT} = \mu = \text{Ln} K X \quad \text{where} \quad K = \text{diag} K \] (27)

and from (16), the composite reaction flow \( V \) is given by the non-linear equations

\[ V^+_0 = \text{Exp} \left( \frac{A^I}{RT} \right) = \text{Exp} A^T, \quad V^-_0 = \text{Exp} \left( \frac{A^F}{RT} \right) = \text{Exp} A^T \] (28)

\[ V = \kappa (V^+_0 - V^-_0) \quad \text{where} \quad \kappa = \text{diag} \kappa \] (29)

Defining the composite stoichiometric and composite reaction constant matrices \( N^\kappa \) and \( K^\kappa \) as

\[ N^\kappa = \left( \begin{array}{c} N^I \\ N^F \end{array} \right) \quad \text{and} \quad K^\kappa = \left( \begin{array}{c} \kappa \\ -\kappa \end{array} \right) \] (30)

Equations (25), (28) and (29) can be rewritten in a more compact form as

\[ \text{Exp} A^\kappa = N^\kappa T \mu, \quad V_0 = \text{Exp} A^\kappa, \quad V = K^\kappa V_0 \] (31)

which can be combined to give a compact expression for the flows \( V \) in terms of the state \( X \)

\[ V = K^\kappa \text{Exp} (N^\kappa T \text{Ln} K X) \] (32)

### 3.1 Block diagrams

Block diagrams are the conventional way of describing systems in the context of control design [25]. However, as discussed by Gawthrop and Bevan [7], bond graphs are superior to block diagrams in the context of system modeling. Nevertheless, block diagrams have advantages when analyzing the system dynamics arising from the bond graph model; in particular, block diagrams expose the underlying feedback structure of the equations arising from the bond graph model. Fig. 3b shows the block diagram corresponding to the closed system bond graph of Fig. 3a; it is a diagrammatic way of writing down (26), (27) and (31). Each arrow corresponds to a vector of signals corresponding to: the \( n_c \) species concentrations \( X \) and normalised chemical potentials \( \mu \), the \( n_r \) reaction flows \( V \) and the \( 2n_r \) normalised forward and reverse affinities \( A^I, \dot{f} \) which represent the integrations of \( X \) to give \( X \) implied by (26). \( \text{Ln} \) and \( \text{Exp} \) represent the non-linear functions in (27) and (29), respectively.

### 3.2 Examples

For example, in the case of the simple reaction \( A = B \) of Fig. 1c

\[ X = \left( \begin{array}{c} x_A \\ x_B \end{array} \right), \quad V = \left( \begin{array}{c} v_1 \\ v_2 \end{array} \right), \quad N^I = \left( \begin{array}{c} 1 \\ 0 \\ 0 \end{array} \right), \quad N^F = \left( \begin{array}{c} 0 \\ 1 \end{array} \right) \] (33)

As \( N^I \) is a unit matrix, the ODE is

\[ -\dot{x}_A = \dot{x}_B = v_1 = \kappa (K_A x_A - K_B x_B) \] (34)

In the case of the enzyme-catalysed reaction \( A + E = C \) of Fig. 2b

\[ X = \left( \begin{array}{c} x_A \\ x_B \\ x_E \end{array} \right), \quad V = \left( \begin{array}{c} v_1 \\ v_2 \end{array} \right), \quad N^I = \left( \begin{array}{c} 1 \\ 0 \\ 0 \end{array} \right), \quad N^F = \left( \begin{array}{c} 0 \\ 1 \\ 1 \\ 1 \end{array} \right) \] (35)

Substituting into (29) gives

\[ v_1 = \kappa_1 (K_A x_A x_E - K_C x_E), \quad v_2 = \kappa_2 (K_C x_C - K_B x_E x_B) \] (36)

and substituting into (26) gives

\[ -\dot{x}_A = -v_1, \quad -\dot{x}_B = v_2, \quad -\dot{x}_C = -v_2 = (v_1 - v_2) \] (37)

### 3.3 Chemostats

As discussed by Polettini and Esposito [35], the notion of a chemostat is useful in creating an open system from a closed system; a similar approach is used by Qian and Beard [36] who use the phrase ‘concentration clamping’. The chemostat has three interpretations:

(i) One or more species is fixed to give a constant concentration [14]; this implies that an appropriate external flow is applied to balance the internal flow of the species.

(ii) An ideal feedback controller is applied to species to be fixed with setpoint as the fixed concentration and control signal an external flow.

(iii) As a \( C \) component with a fixed state.

Define \( T^c \) as the set containing the indices of the species corresponding to the chemostats. Then the \( n_c \times n_c \) diagonal matrices \( T^c \) and \( T^{cs} \) are defined as

\[ T^c = \left\{ \begin{array}{ll} 1 & \text{if } i \in T^c \\ 0 & \text{if } i \notin T^c \end{array} \right\}, \quad T^{cs} = \left\{ \begin{array}{ll} 1 & \text{if } i \in T^{cs} \\ 0 & \text{if } i \notin T^{cs} \end{array} \right\} \] (38)

It follows that \( I_C = T^c + T^{cs} \) where \( I_C \) is the \( n_C \times n_C \) unit matrix. The stoichiometric matrix \( N \) can then be expressed as the sum of two matrices: the chemostatic stoichiometric matrix \( N^{cs} \) and the
chemodynamic stoichiometric matrix $N^{cd}$ as

$$N = N^{cs} + N^{cd} \quad (39)$$

where

$$N^{cs} = I^{cs} N \quad \text{and} \quad N^{cd} = I^{cd} N \quad (40)$$

Note that $N^{cd}$ is the same as $N$ except that the rows corresponding to the chemostat variables are set to zero. The stoichiometric properties of $N^{cd}$, rather than $N$, determine system properties when chemostats are present. When chemostats are used, the state (26) is replaced by

$$\dot{X} = N^{cd} V = NV - V^{cs} \quad \text{where} \quad V^{cs} = N^{cs} V \quad (41)$$

and thus the fixed states are held constant by the external flow $V = -V^{cs} = -N^{cs} V$ acting at the $C$ components. Thus the closed-system bond graph of Fig. 3a is replaced by the open-system bond graph of Fig. 3c where the external flows $V$ have been added.

### 3.4 Flowstats

In addition to ‘concentration clamping’ (identified with chemostats in Section 3.3), Qian and Beard [36] also use ‘boundary flux injection’ to convert closed to open systems. Here we ‘fix’ flows though $Re$ components to create flowstats. Although Polletini and Esposito [35] ‘focus on chemostats for thermodynamic modelling’ and note that chemostats can be used to create fixed currents, it is argued that flowstats provide a useful complement to chemostats.

In a similar way to Section 3.3, define $I^{fs}$ as the set containing the indices of the reactions corresponding to the flowstats. Then the $n_{t} \times n_{t}$ diagonal matrices $I^{fs}$ and $I^{fd}$ are defined as

$$I^{fs} = \begin{cases} 1 \quad \text{if} \quad i \in I^{fs} \\ 0 \quad \text{if} \quad i \notin I^{fs} \end{cases}, \quad I^{fd} = \begin{cases} 1 \quad \text{if} \quad i \in I^{fd} \\ 0 \quad \text{if} \quad i \notin I^{fd} \end{cases} \quad (42)$$

It follows that $I_{f} = I^{fs} + I^{fd}$ where $I_{f}$ is the $n_{t} \times n_{t}$ unit matrix. Thus the flows $V$ are replaced by $V^{fs}$ where

$$V^{fs} = I^{fs} V + I^{fd} V^{fs} \quad (43)$$

Assuming that chemostats are also present, (41) is replaced by

$$\dot{X} = N^{cd} V^{fs} = N^{cd} (I^{fs} V + I^{fd} V^{fs}) \quad (44)$$

If $V^{fs} \neq 0$, the stoichiometric properties of $N^{cd}$ (i.e. determined by the chemostats) determine system properties. However, if $V^{fs} = 0$, then the stoichiometric properties of

$$N^{d} = N^{cd} I^{fd} = I^{cs} N I^{fd} \quad (45)$$

(i.e. both chemostats and flowstats) determine system properties. Note that $N^{d}$ is the same as $N^{cd}$ except that the columns corresponding to the flowstat variables are set to zero.

Fig. 3d shows the block diagram corresponding to the open-system bond graph of Fig. 3c. It differs from Fig. 3a in that $N$ of (27) is replaced by $N^{d}$ of (41) to reflect the fact that the chemostat states are not affected by $V$ and thus correspond to the zero rows of $N^{cs}$. Moreover, the matrices $I^{fs}$ and $I^{fd}$, and the flows $V^{fs}$ are added to reflect the effect of the flowstats.

### 3.5 Reduced-order equations

As discussed by the authors [37, 38], the presence of conserved moieties leads to potential numerical difficulties with the solution of (26). As chemostats introduce further conserved moieties it is important to resolve this issue. The following outline uses the notation and approach of Gawthrop and Crampin [13, §3(c)].

Defining $G^{cd}$ as the left null-space matrix of $N^{cd}$ it follows that

$$G^{cd} \dot{X} = G^{cd} N^{cd} V = 0 \quad (46)$$

Hence, each of the $n_{G}$ rows of $G^{cd}$ defines an algebraic relationship between the states contained in $X$. Thus the number of independent states $n_{s}$ is given in terms of the total number of states $n_{X}$ by

$$n_{s} = n_{X} - n_{G} \quad (47)$$

The derivative of the independent states $x$ is given in terms of the derivative of state $X$ by the $n_{s} \times n_{X}$ transformation matrix $L_{X}^{cd}$

$$\dot{x} = L_{X}^{cd} \dot{X} \quad (48)$$

Similarly

$$\dot{X} = L_{X}^{cd} x \quad (49)$$

where $L_{X}^{cd}$ is an $n_{x} \times n_{x}$ matrix. Integrating (49)

$$X = L_{X}^{cd} x + L_{X}^{cd} x_{0} = L_{X}^{cd} x + G_{X} x_{0} \quad (50)$$

where

$$G_{X} = I_{n_{X} \times n_{x}} - L_{X}^{cd} L_{X}^{cd} \quad (51)$$

and $x_{0}$ and $X_{0}$ are the values of $x$ and $X$ at time $t = 0$, respectively.

Fig. 4a corresponds to the open-system bond graph and block diagram of Fig. 3, but the reduced-order (48) and (50) have been incorporated. The block $L_{X}^{cd}$ contracts the state dimension from $n_{X}$ to $n_{s}$ and the block $L_{X}^{cd}$ expands it again. The initial condition term $G_{X} x(0)$ becomes an exogenous signal analogous to the setpoint term of feedback control; note that this includes the states of all of the chemostats.

### 3.6 Examples

The simple reaction $A \xrightarrow{1} B$ of Fig. 1c has a single conserved moiety represented by

$$x_{A} + x_{B} = x_{AB} \quad (52)$$

where $x_{AB}$ is a constant. One possibility is

$$x = (x_{A}), \quad L_{X} = \begin{pmatrix} 1 & 0 \end{pmatrix}, \quad G_{X} = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} \quad (53)$$

The enzyme-catalysed reaction $A + E \xrightarrow{1} C \xrightarrow{2} B + E$ of Fig. 2b has a number of possible representations depending on which $C$ components are chemostats and which $Re$ components are flowstats. Two of these are examined here.

First, consider the case where both $C A$ and $C B$ are chemostats and $Re r0$ is a flowstat with zero flow. The relevant stoichiometric matrix is thus $N^{d}$ of (45) that determines system properties and

$$N^{d} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{pmatrix}, \quad G^{d} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \quad (54)$$

$G^{d}$ has three rows corresponding to the three conserved moieties $x_{A}$, $x_{B}$ and $x_{C} + x_{E}$. The first correspond to the two chemostats, the third to the well-known conserved moiety for enzyme-catalysed reactions.
the total enzyme amount is conserved. There is only one independent state which is chosen as $x_C$. With this choice

$$L_{xX} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ \end{pmatrix}, \quad L_{x\xi} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ \end{pmatrix},$$

$$G_X = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ \end{pmatrix} \text{ (55)}$$

Second, consider the case where both $C:A$ and $C:B$ are chemostats and $Re:Theta$ is a flowstat with non-zero flow. The relevant stoichiometric matrix is thus $N^{cd}$ of (39) that determines system properties and

$$N^{cd} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & -1 & -1 \\ \end{pmatrix}, \quad G^{cd} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \end{pmatrix} \text{ (56)}$$

The effect of the variable flowstat is to remove the third conserved moiety leaving only the chemostat states $x_d$ and $x_b$. There are now two independent states $x_C$ and $x_E$. This gives

$$L_{xX} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \end{pmatrix}, \quad L_{x\xi} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ \end{pmatrix} \text{ (57)}$$

$$G_X = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \end{pmatrix} \text{ (58)}$$

4 **Linearisation**

As discussed in Section 1, linearisation of non-linear systems is a standard technique in control engineering. In Section 5, linearisation is used to analyse the properties of modules.

Assuming that the system reaches a steady-state $\bar{X}$, i.e. $\dot{X} = 0$ when $X = \bar{X}$, the system can be linearised about that steady state by introducing perturbation variables $\tilde{X}$ so that $X = \bar{X} + \tilde{X}$. These can be defined for each relevant variable; e.g.

$$x = \bar{x} + \tilde{x}, \quad \mu = \bar{\mu} + \tilde{\mu}, \quad A^{f} = \bar{A}^{f} + \tilde{A}^{f}, \quad v = \bar{v} + \tilde{v} \text{ (59)}$$

If the perturbation is small, each variable can be approximated using a first-order Taylor series; thus, e.g. $\tilde{\mu} \approx (\partial \mu / \partial x) \tilde{x}$.

4.1 **Component linearisation: C and Re**

The non-linear $C$ component is defined by (11). In particular, for substance $A$

$$\tilde{\mu}_d = \ln K_d x_d, \quad \tilde{x}_d = v_d \text{ (60)}$$

Using the perturbation approach, it follows that the linearised $C$ component is defined by the equations

$$\tilde{\mu}_d = \tilde{K}_d \tilde{x}_d, \quad \tilde{K}_d = \frac{\partial \tilde{\mu}_d}{\partial \tilde{x}_d} = \frac{1}{x_d}, \quad \tilde{x}_d = \tilde{v}_d \text{ (61)}$$

The non-linear $Re$ component representing the $i$th reaction (12) is defined by (16), in particular

$$\nu_i = \kappa_i (v_i^+ - v_i^-), \quad v_i^+ = \frac{\kappa_i}{T}, \quad \nu_i^- = \frac{\kappa_i}{T} \text{ (62)}$$

Hence the linearised $Re$ component is defined by the following
equations
\[ \ddot{\mathbf{v}}_i = \mathbf{K}_i \dot{\mathbf{A}}_i^T - \mathbf{K}_j \dot{A}_j^T, \]
\[ \mathbf{K}_i^T = \frac{\partial \mathbf{v}_i}{\partial \dot{A}_i} = \mathbf{K}_j \mathbf{v}_0, \quad \mathbf{K}_j^T = -\frac{\partial \mathbf{v}_j}{\partial \dot{A}_j^T} = \mathbf{K}_j \mathbf{v}_0 \] \hspace{1cm} (62)

4.2 Linearised system equations

Section 4.1 shows how the bond graph components are linearised; essentially the non-linear exp and In functions are replaced by linear gains dependent on the steady-state flows and steady-state states, respectively. The \( n_a \) constants \( \tilde{K}_A, \tilde{K}_B, \ldots \) of the linearised \( C \) components, and the \( n_c \) constants \( \bar{K}_1^T, \bar{K}_2^T, \ldots \) and \( \hat{K}_1^T, \hat{K}_2^T, \ldots \) are collected into column vectors
\[ \mathbf{K}_A = \begin{pmatrix} \tilde{K}_A \\ \hat{K}_A^T \end{pmatrix}, \quad \mathbf{K}_B = \begin{pmatrix} \tilde{K}_B \\ \hat{K}_B^T \end{pmatrix}, \quad \mathbf{K}_C = \begin{pmatrix} \tilde{K}_C \\ \hat{K}_C^T \end{pmatrix} \] \hspace{1cm} (63)

Fig. 4b shows the block diagram corresponding to the linearisation of the reduced-order system depicted in Fig. 4a where
\[ \mathbf{K} = \text{diag} \mathbf{K}, \quad \bar{\mathbf{K}} = \text{diag} \mathbf{K}_C, \quad \hat{\mathbf{K}} = \text{diag} \mathbf{K}_C \] \hspace{1cm} (64)

The block diagram of the linearised version of the full system of Fig. 3d gives the following linear state-space equations
\[ \dot{\mathbf{V}} = \mathbf{C} \tilde{\mathbf{X}}, \quad \dot{\mathbf{V}}^\text{cd} = \mathbf{I}^\text{cd} \dot{\mathbf{V}} + \mathbf{I}^\text{cd} \mathbf{V}^\text{fs}, \quad \dot{\mathbf{X}} = \mathbf{A} \mathbf{X} + \mathbf{B} \mathbf{V}^\text{fs} \] \hspace{1cm} (65)

where
\[ \mathbf{A} = \mathbf{N}^{\text{cd}} \mathbf{L}^\text{cd} \mathbf{C} = \mathbf{N}^\text{cd} \mathbf{C}, \quad \mathbf{B} = \mathbf{N}^{\text{cd}} \mathbf{I}^\text{fs}, \quad \mathbf{C} = \mathbf{K}_C^T \mathbf{N}^{\text{fs}} \mathbf{X} \] \hspace{1cm} (66)

The block diagram of the reduced system shown in Fig. 4b gives the following linear state-space equations
\[ \dot{\mathbf{V}} = \mathbf{c} \tilde{\mathbf{X}} + \mathbf{d} \mathbf{X}_0, \quad \dot{\mathbf{X}} = \mathbf{a} \tilde{\mathbf{X}} + \mathbf{b} \mathbf{V}^\text{fs} + \mathbf{b} \mathbf{X}_0 \] \hspace{1cm} (67)

where
\[ \mathbf{a} = \mathbf{L}^{\text{cd}} \mathbf{A} \mathbf{L}^{\text{fs}}, \quad \mathbf{b} = \mathbf{L}^{\text{cd}} \mathbf{b}_t, \quad \mathbf{b}_t = L^{\text{cd}} A G^{\text{cd}} \mathbf{c} = \mathbf{C} \mathbf{L}^{\text{cd}} \mathbf{X} + \mathbf{d} = \mathbf{C} \mathbf{L}^{\text{cd}} \mathbf{X}^\text{fs} \] \hspace{1cm} (68)

Equations (67) can be written more compactly as
\[ \dot{\mathbf{V}} = \mathbf{c} \tilde{\mathbf{X}} + \mathbf{d} \dot{\mathbf{X}} \] \hspace{1cm} (69)
\[ \dot{\mathbf{X}} = \mathbf{a} \tilde{\mathbf{X}} + \mathbf{b} \dot{\mathbf{X}} \] \hspace{1cm} (70)

where
\[ \mathbf{U} = \begin{pmatrix} \dot{\mathbf{X}}_0 \\ \mathbf{V}^\text{fs} \end{pmatrix} \] \hspace{1cm} (71)
\[ \mathbf{b} = (\mathbf{b}_t, \mathbf{b}_c) \] \hspace{1cm} (72)

and
\[ \mathbf{d} = (\mathbf{d}, 0_{n_x \times n_y}) \] \hspace{1cm} (73)

where \( 0_{n_x \times n_y} \) is the zero matrix with indicated dimensions.

As the state-space systems (65) and (67) are linear, they can also be represented as transfer functions in the Laplace variable \( s \). In particular, the reduced-order system (67) has the transfer function \( G(s) \) given by
\[ G(s) = c \left( A_{n_x \times n_x} - a \right)^{-1} b + d \] \hspace{1cm} (74)

where \( I_{n_x \times n_y} \) is the unit matrix with indicated dimensions.

4.3 Examples

The simple reaction \( A \rightleftharpoons B \) of Fig. 1c has a flow given by (19). As both state derivatives are proportional to \( \mathbf{v} \), it follows that the steady-state is defined by: \( \mathbf{v} = \kappa (K_B x_A - K_A x_B) = 0 \). As noted in (52), \( x_A + x_B = x_{AB} \) where \( x_{AB} \) is a constant. It follows that the steady-state values of \( x_A \) and \( x_B \) are
\[ \tilde{x}_A = \frac{K_B}{K_A + K_B} x_{AB}, \quad \tilde{x}_B = \frac{K_A}{K_A + K_B} x_{AB} \] \hspace{1cm} (75)

From (18), \( \tilde{v}_A^\text{fs} = K_A x_A \) and \( \tilde{v}_0^\text{fs} = K_B x_B \). Hence
\[ \tilde{v}_A^\text{fs} = \tilde{v}_0^\text{fs} = \frac{K_A K_B}{K_A + K_B} x_{AB} \] \hspace{1cm} (76)

Using formulae (60) and (62), it follows that the coefficients of the linearised \( C \) components are
\[ \mathbf{K}_C = \begin{pmatrix} 1 \\ K_A + K_B \end{pmatrix}, \quad \mathbf{K}_{C} = \begin{pmatrix} 0 \\ K_A + K_B \end{pmatrix} \] \hspace{1cm} (77)

and that the coefficients of the linearised \( Re \) component are
\[ \mathbf{K}_R = \mathbf{K}_C \] \hspace{1cm} (78)

Hence the linearised equations for the flow are
\[ \dot{\mathbf{V}} = \mathbf{c} \tilde{\mathbf{X}} + \mathbf{d} \dot{\mathbf{X}} \] \hspace{1cm} (79)

As expected, the linearisation of a linear equation is the same as the linear equation.

5 Modularity, retroactivity and feedback

Modularity provides one approach to understanding the complex systems associated with biochemical systems [39–44]. However, as discussed by Kaltenbach and Stelling [45] there are many possible concepts of modularity. These include structure deduced from the stoichiometric matrix [46, 47]; modular construction of in silico models [48]; and modular structure designed to minimise the retroactivity between modules [19, 20, 45].

This paper focuses on two overlapping, but conceptually different concepts of modularity:

- **Computation modularity**: modules retain physically correct results when connected together to form a system.
- **Behavioural modularity**: modules retain their behaviour (such as ultrasensitivity) when connected together.

Gawthrop et al. [14] have shown that bond graphs provide an effective foundation for modular construction of computer models of biochemical systems. This paper focuses on the second interpretation of modularity and shows that bond graphs provide a natural interpretation of inter-module retroactivity [15–17, 19, 20, 49, 50]. Retroactivity has been illustrated experimentally in the context of ‘signalling properties of a covalent modification cycle’
Using the superscripts &IET Syst. Biol. Series connection: linearised block diagram

Series connection: bond graph

Modularity, retroactivity and feedback

As discussed in Section 1, feedback is another concept crucial to the understanding of complex systems. Kholodenko [22], Brightman and Fell [56], Asthagiri and Lauffenburger [57], Kolch et al. [58], Hornberg et al. [59] and Sauro and Ingalls [60] investigate the feedback in the context of MAPK cascades. As will be shown in this paper, retroactivity and feedback are closely related concepts. As will be seen, feedback arises in a number of ways including:

Intrinsic feedback: due to the interaction of reactions and species within and between modules
Conserved moieties: implicitly generate feedback loops
Feedback inhibition: explicitly uses negative feedback.

As discussed in Section 4, linearisation of a non-linear system allows a wide range of control engineering techniques to be applied. In this section, linearisation is used to investigate transfer functions and frequency-domain methods.

Fig. 5a shows the series interconnection of two bond graph modules labelled A and B. In this example, each module has two ports labelled 1 and 2 and the modules are interconnected to form a composite module AB with two ports. To create a block diagram from a bond graph, the concept of causality is required. This concept is discussed in detail in [3–6], but here it suffices to know that causality determines which variable on a bond impinging on a system is the input, and which the output. For example, in this case the causality is such that flow v is the input (and effort μ the output) on port 1 and that effort μ is the input (and flow v the output) on port 2.

As discussed by Gawthrop et al. [14], the bond graph approach can be used to build arbitrarily complex systems out of such modules. However, to delve more deeply into the power of the bond graph approach and to understand how modules interact, it is instructive to look at the block diagram equivalents following linearisation as discussed in Section 4. With the assumed causality, each module can be represented by four transfer functions and (80) can be represented for each of the interconnected modules as Fig. 5b. Connecting port 2 of A to port 1 of B shown in Fig. 5a is equivalent to connecting the corresponding signals shown in Fig. 5b:

\[ v_2' = v_1' \quad \text{and} \quad \mu_2' = \mu_1' \tag{81} \]

This connection induces a feedback loop involving \( G_{12}^A \) and \( G_{21}^B \) thus the properties of the composite system are dependent on the loop gain \( L_1 \) of this feedback loop.

In particular, using (80) for \( A \) and \( B \) and substituting (81) gives the transfer function \( G_{AB} \) for the composite module as:

\[ G_{11}^{AB} = \frac{G_{11}^A G_{11}^B}{1 + L_1} \tag{82} \]
\[ G_{12}^{AB} = G_{12}^B + \frac{G_{12}^A G_{22}^B}{1 + L_1} \tag{83} \]
\[ G_{21}^{AB} = G_{21}^A + \frac{G_{21}^B G_{11}^B}{1 + L_1} \tag{84} \]
\[ G_{22}^{AB} = \frac{G_{22}^A G_{22}^B}{1 + L_1} \tag{85} \]
\[ L_1 = -G_{12}^B G_{21}^B \tag{86} \]

While this will be called the interaction loop-gain. In linear systems, feedback shifts system poles and therefore changes the behaviour of the interacting systems. In particular, each of the transfer functions \( G_{ij}^B \) of (82)–(85) is modified by the interaction loop-gain. Thus the feedback loop comprising \( G_{12}^B \) and \( G_{21}^B \) is the source of behaviour alteration when two modules are connected. It follows that approximate behavioural modularity is achieved by making the interaction loop-gain as small as possible. Indeed, in the special case that \( G_{12}^B = G_{21}^B = 0 \) and so \( L_1 = 0 \) then:

\[ G_{11}^{AB} = G_{11}^A G_{11}^B, \quad G_{12}^{AB} = G_{21}^B, \quad G_{21}^{AB} = G_{21}^B \tag{87} \]

5.1 Example module: simple reaction

Fig. 6a shows a simple reaction system comprising a species represented by the C component \( C \) and the reaction component \( Rer \). This closed system is converted to an open system by appending a flowstat \( Rer0 \) with flow \( v_0 \) and a chemostat \( C0S \) with state \( x_0 \). The system is linear and the reaction flow (though \( Rer \)) is given by:

\[ v = \kappa (K_C x_C - K_{x_C}) \tag{88} \]

and the rate of change of \( x_C \) is:

\[ \dot{x}_C = v_0 - v \tag{89} \]

Equations (88) and (89) can be visualised using the block diagram of Fig. 6b, which clearly shows the implicit feedback structure with loop gain:

\[ L = \frac{\kappa K_C}{s} \tag{90} \]

It follows from the block diagram of Fig. 6b that:

\[ \frac{v}{x_C} = \frac{1}{1 + L} \left( \frac{L}{\kappa K_C} \right) \tag{91} \]
\[ \frac{v_0}{x_C} = \frac{1}{s + \kappa K_C} \left( \frac{-s \kappa K_C}{1 + \kappa K_C} \right) \frac{v_0}{x_C} \]
In the particular case, that \( \kappa = K_r = K_s = 1 \)

\[
G_{11} = G_{22} = G_{21} = \frac{1}{s+1}, \quad G_{12} = \frac{-s}{s+1} \quad (92)
\]

If two identical copies of this module are placed in series as shown in Fig. 6d

\[
L_I = \frac{s}{(s+1)^2}, \quad \frac{1}{1+L_I} = \frac{(s+1)^2}{s^2 + 3s + 1} \quad (93)
\]

and the resulting overall transfer function is

\[
G_{11}^{AB} = G_{22} = \frac{1}{s^2 + 3s + 1}, \quad G_{21}^{AB} = \frac{-s(s+2)}{s^2 + 3s + 1}, \quad G_{12}^{AB} = \frac{s+2}{s^2 + 3s + 1} \quad (94)
\]

The isolated modules each have a single pole at \( s = -1 \); the series modules have a pole at \( s = -0.38 \) and at \( s = -2.62 \). This shift in pole location is due to non-zero interaction loop-gain \( L_I \) (86).

Such reaction systems are often incorrectly modelled using an irreversible reaction where the flow is independent of \( \mu_2 \). This would imply that \( G_{12} = G_{22} = L = 0 \) and thus the overall transfer function would be

\[
G_{11}^{AB} = \frac{1}{(s+1)^2}, \quad s^2 + 2s + 1 \quad (95)
\]

This thermodynamically incorrect system has zero retroactivity. As will be shown in the sequel, approximate irreversibility, and thus approximate zero retroactivity, can be achieved, but at the metabolic cost of using a power supply such as that provided by the ATP \( \rightarrow \) ADP + Pi reaction.

5.2 Example module: enzyme-catalysed reaction

As an example, the enzyme-catalysed reaction of Fig. 2b is considered as a two-port module (as illustrated in Fig. 5). In particular, the flowstat corresponding to \( \text{Recr0} \) is replaced by port 1 and the chemostat corresponding to \( \text{C:B} \) is replaced by port 2.

Thus (80) becomes

\[
\left( \begin{array}{c}
\tilde{v}_2 \\
\mu_k
\end{array} \right) = \left( \begin{array}{cc}
G_{11} & G_{12} \\
G_{21} & G_{22}
\end{array} \right) \left( \begin{array}{c}
\tilde{v}_0 \\
\mu_B
\end{array} \right) \quad (96)
\]

The system parameters were \( K_R = K_C = K_E = 1, \kappa_1 = 10 \) and \( \kappa_2 = 1 \). Three alternative values were used for \( K_R: 2, 10 \) and 100. Using an initial state \( X_0 = (100 \ 1 \ 0 \ 1) \), the steady states were found for each value of \( K_R \) and the system was linearised using the method of Section 4. The transfer functions for the three cases were found to be

\[
G_2 = \left( \begin{array}{cc}
\frac{-s+10}{s^2 + 32} & -0.34s - 10.31 \\
\frac{-s+90}{s^2 + 32} & -1
\end{array} \right) \quad (97)
\]

\[
G_{10} = \left( \begin{array}{cc}
\frac{s+90}{s^2 + 112} & -0.10s - 10.80 \\
\frac{s+990}{s^2 + 112} & -1
\end{array} \right) \quad (98)
\]

Although these transfer functions are simple enough to analyse directly, in more complex cases it is useful to look at the transfer function frequency responses obtained by replacing the Laplace variable \( s \) by \( j\omega \) where \( j = \sqrt{-1} \) and \( \omega \) is a frequency in rad \( \cdot \) s\(^{-1}\). Fig. 7 gives the frequency response magnitude of the three transfer functions: \( G_{11} \) relating \( \tilde{v}_0 \) to \( \tilde{v}_2 \), \( G_{22} \) relating to \( \tilde{\mu}_B \) to \( \tilde{\mu}_E \) and the loop-interaction \( L_I = -G_{11}G_{21} \) for each of the three cases.

The forward transfer function \( G_{11} \) approaches 1/s as \( K_R \) increases, the transfer functions \( G_{22} \) and \( L_I \) decrease as \( K_R \) increases. Thus larger values of \( K_R \) give approximate behavioural modularity. However, this comes at an energetic cost measured by the external flow associated with the chemostat \( \text{C:A} \).

5.3 Example module: phosphorylation/dephosphorylation

A bond graph model of the thermodynamically correct formulation of the phosphorylation/dephosphorylation cycle of Beard and Qian [61] was presented by Gawthrop and Crampin [13]. Fig. 9a shows a modular version where the two ports are given by flowstat \( \text{Recr0} \)
the chemostat. The three components representing ATP, ADP and Pi (ATP, ADP and Pi) are also chemostats and provide the power source for the module.

As shown in Section 5.2, this module can be analysed by plotting the frequency response of the three transfer functions. The parameters (which are illustrative and do not correspond to a specific biological instance) are

\[
X = \begin{pmatrix}
  x_{E1} \\
  x_{C1} \\
  x_{E2} \\
  x_{C2} \\
  x_{ADP} \\
  x_p \\
  x_M \\
  x_{MP}
\end{pmatrix}, \\
X_0 = \begin{pmatrix}
  0 \\
  0 \\
  0.001 \\
  110, 100 \\
  1 \\
  0.01 \\
  10 \\
  0
\end{pmatrix},
\]

\[
K = \begin{pmatrix}
  100 \\
  1 \\
  100 \\
  1 \\
  0.1 \\
  0.001 \\
  0.001 \\
  1
\end{pmatrix}, \\
\kappa = \begin{pmatrix}
  10 \\
  1000 \\
  10 \\
  1000 \\
  10 \\
  1000 \\
  10 \\
  1000
\end{pmatrix}
\]

(100)

The (fixed) amount of ATP was set at three alternative values: \( x_{ATP} = 110, 100 \). As shown in Section 5.2, larger values give reduced interaction at the expense of more power needed to drive the module.

5.4 Example module: feedback inhibition

The idea that a product can inhibit an enzyme and thus give negative feedback is a well-established concept in biology [62–65]. This section focuses on one possible mechanism, competitive inhibition [66, §1.4.3]. The basic idea is that the product P binds to the enzyme E to form a complex C (thus partially sequestering E) via the reaction

\[
P + E \rightarrow C
\]

(101)

Together with an additional flow of enzyme modelled by \( \text{Recr0} \), this reaction is modelled by the bond graph of Fig. 10a. This can be represented as a two-port module if \( C.P, \text{Recr0} \) and associated junctions are replaced by ports. This module will be used in the sequel to apply feedback inhibition to the MAPK cascade.

6 MAPK cascades

The MAPK cascade is a well-studied signalling pathway with ultrasensitive components [21, 22, 58]. However, the use of the Michaelis–Menten approximation to enzyme-catalysed reactions can be misleading in this context. In particular, as discussed by Voit [67, §9.5], ‘It is tempting to set up the two phosphorylation steps with Michaelis–Menten rate functions, but such a strategy is not the best option, because (i) the enzyme concentration is not constant, (ii) the enzyme concentration is not necessarily smaller than the substrate concentration and (iii) the two reaction steps are competing for the same enzyme’.

This section shows that the bond graph property of computational modularity can be used to build a computational model of the MAPK cascade which is thermodynamically correct and thus avoids the pitfalls associated with inappropriate use of the Michaelis–Menten approximation. Moreover, having seen in Section 5.3 that the bond graph module corresponding to phosphorylation/dephosphorylation can be designed to give approximate behavioural modularity, the MAPK cascade can be built with approximate behavioural modularity.

Fig. 9b shows the bond graph of the MAPK cascade based on the phosphorylation/dephosphorylation module \( PD \) of Fig. 8 and the double phosphorylation/dephosphorylation module \( DPD \) the \( \text{Recr0} \) component is used as a flowstat generating a flow \( v_0 \) as discussed in Section 5.2. The non-linear system of ODEs corresponding to Fig. 9b was simulated for 100 time units with an input \( v_0 \) given by

\[
v_0 = \begin{cases}
  10^{-6} & \text{if } 20 \geq t \geq 30 \\
  -10^{-6} & \text{if } 50 \geq t \geq 60 \\
  0 & \text{otherwise}
\end{cases}
\]

(102)

This gives a maximum value of the total enzyme of \( e_{\text{max}} = 10^{-5} \). The system parameters are those used in Section 5.3.

Fig. 9c shows the corresponding time courses for the total amount of enzyme \( e_{\text{tot}} \) and the amounts of MKKKP, MKKPP and MKPP. A logarithmic scale is used to account for the large range of values. Note that the gain between \( e_{\text{tot}} \) and the concentration \( x_{MKPP} \) is of the order of \( 10^5 \).

The steady-state value of \( x_{MKPP} \) was computed for a range of values of \( e_{\text{tot}} \) and the incremental values (\( \Delta x_{MKPP}/\Delta e_{\text{tot}} \)) were computed numerically for three values of ATP. The system behaves as designed to give approximately linear response. The high gain due to the ultrasensitivity of the phosphorylation/dephosphorylation modules vanishes between ATP amounts of 2 and 5.

In his seminal paper Black [68] points out that ‘by building an amplifier whose gain is deliberately made, say 40 dB higher than necessary … and then feeding the output back on the input in such a way as to throw away the excess gain, it has been found possible to effect extraordinary improvement in constancy of amplification and freedom from non-linearity.’ In this context, Fig. 10b is the same as shown in Fig. 9b except that the feedback

\[
IET Syst. Biol., 2016, Vol. 10, Iss. 5, pp. 187–201
© The Institution of Engineering and Technology 2016
197
inhibition module of Section 5.4 is incorporated into the bond graph and the system is re-simulated with $K_f = 4$ and $k_{fi} = 1$.

The steady-state value of $x_{MKPP}$ was computed for a range of values of $e_{tot}$ and the incremental values ($dx_{MKPP}/de_{tot}$) were computed numerically both with and without feedback and plotted in Fig. 10d. The gain of the system is reduced by a factor of about 20, but the system is now more linear: the gain is approximately constant over a wider range of $e_{tot}$ than was the case without feedback.

7 Conclusions

Building on its inherent computational modularity, it has been shown that the bond graph approach can be used to explain and adjust behavioural modularity. The MAPK cascade was used as an example to illustrate this point. It would be interesting to repeat the MAPK examples with parameter values taken from the literature [69, 70]. This may provide insight into the evolutionary trade-off between energy consumption and signalling performance [71–73].
Control-theoretic concepts based on linearisation were shown to provide a quantitative analysis of behavioural modularity. However, non-linear systems can be approximated in other ways apart from linearisation. In the context of metabolic network modelling, Heijnen [74] discusses and compares a number of approximations including: logarithmic-linear, power law generalised mass action, S-systems [63, 67] and linear logarithmic [75, 76]. It would be interesting to see whether such approximations provide an alternative to linearisation in analysing behavioural modularity.

It has been suggested that metabolism and its dysfunctions may relate to certain diseases including Parkinson’s disease [77, 78], heart disease [79], cancer [80, 81] and chronic fatigue [82]. It is envisaged the energy-based approach used in this paper will help to understand such energy-related diseases.

The example in this paper examines a signalling network as an analogy to an electronic amplifier. Gene regulatory networks have been analysed and synthesised as amplifiers [83–85]. Future work will examine the bond graph-based analysis and synthesis of gene regulatory networks.

8 Acknowledgments

Peter Gawthrop thanked the Melbourne School of Engineering for its support via a Professorial Fellowship. This research was in part conducted and funded by the Australian Research Council Centre of Excellence in Convergent Bio-Nano Science and Technology (project number CE140100036), and by the Virtual Physiological Rat Centre for the Study of Physiology and Genomics, funded through NIH grant P50-GM094503.

9 References

1. Wellstead, P., Bullinger, E., Kalamatianos, D., et al.: ‘The role of control and system theory in systems biology’, Annu. Rev. Control, 2008, 32, (1), pp. 33–47, ISSN 1367-5788. doi: 10.1016/j.arcontrol.2008.02.001
2. Paynter, H.M.: ’Analysis and design of engineering systems’ (MIT Press, Cambridge, Mass, 1961)
3. Wellstead, P.E.: ’Introduction to physical system modelling’ (Academic Press, 1979)
4. Gawthrop, P.J., Smith, L.P.S.: ’Metamodelling: bond graphs and dynamic systems’ (Prentice-Hall, Hemel Hempstead, Herts, England, 1996), ISBN 0-13-489624-9
5. Mukherjee, A., Karmaker, R., Samantaray, A.K.: ’Bond graph in modeling, simulation and fault indentification’ (I.K. International, New Delhi, 2006)
6. Karnopp, D.C., Margolis, D.L., Rosenberg, R.C.: ’System dynamics: modeling, simulation, and control of mechatronic systems’ (John Wiley & Sons, 2012, 5th edn.), ISBN 978-0470889084
7. Gawthrop, P.J., Bevan, G.P.: ’Bond-graph modeling: a tutorial introduction for control engineers’, IEEE Control Syst. Mag., 2007, 27, (2), pp. 24–45, doi: 10.1109/MCS.2007.338279
8. Oster, G., Perelson, A., Katchalsky, A.: ’Network thermodynamics’, Nature, 1971, 230, pp. 393–399, doi: 10.1038/230393a0
9. Oster, G.F., Perelson, A.S., Katchalsky, A.: ’Network thermodynamics: dynamic modelling of biophysical systems’, Q Rev. Biophys., 1973, 6, (01), pp. 1–134, doi: 10.1017/S0035563500008081
10. Cellier, F.E.: ’Continuous system modelling’ (Springer-Verlag, 1991)
11 Thoma, J.U., Moccilin, G.: ‘Simulation with entropic thermodynamics: understanding complex systems with bionet graphs’ (Springer, 2005).

12 Greifeneder, J., Cellier, F.E.: ‘Modeling chemical reactions using bond graphs’ Proc. ICBGM21, 10th SCS Int. Conf. on Bond Graph Modeling and Simulation, Germany, 2006.

13 Gawthrop, P.J., Crampin, E.J.: ‘Energy-based analysis of biochemical cycles using bond graphs’, Proc. R. Soc. A Math. Phys. Sci. Eng., 2014, 470, (2171), pp. 1–25, doi: 10.1098/rspa.2014.0459.

14 Gawthrop, P.J., Crampin, E.J.: ‘Hierarchical bond graph modelling of biochemical networks’, Proc. R. Soc. A Math. Phys. Sci. Eng., 2015, 471, (2184), pp. 1–23, doi: 10.1098/rspa.2015.0642. Available at arXiv:1503.01814.

15 Saco-Rodriguez, J., Kremling, A., Conzelmann, H., et al.: ‘Modular analysis of signal transduction networks’, IEEE Control Syst. 2004, 24, (4), pp. 35–52, doi: 10.1109/MCS.2004.1316652.

16 Saco-Rodriguez, J., Kremling, A., Gilles, E.D.: ‘Dissecting the puzzle of life: modularization of signal transduction networks’, Comput. Chem. Eng., 2005, 29, (3), pp. 619–629, doi: 10.1016/j.compchemeng.2004.08.035.

17 Vecchio, D.D., Ninfa, A.J., Sontag, E.D., et al.: ‘Retroactivity controls the temporal dynamics of gene transcription’, ACS Synth. Biol., 2013, 2, (8), pp. 431–441, doi: 10.1021/bi300090x.

18 Vecchio, D.D., Sontag, E.D.: ‘Engineering principles in bio-molecular systems: From retroactivity to modularity’, Eur. J. Control, 2009, 15, (4–5), pp. 389–397, doi: 10.3166/ejc.15.389-397.

19 Barton, J., Sontag, E.D., et al.: ‘Control of modular signaling cascade: The energy costs of biological insulators’, arXiv preprint arXiv:1201.3809, 2012.

20 Brightman, F.A., Fell, D.A.: ‘Differential feedback regulation of the MAPK cascade underlies the stability of product dependent signaling’, FEBS Lett., 2000, 473, (1–2), pp. 85–96, doi: 10.1016/S0014-5793(00)02037-8.

21 Asthagiri, A.R., Lauffenburger, D.A.: ‘A computational study of feedback effects on signal dynamics in a mitogen-activated protein kinase (MAPK) pathway model’, Bioinformatics, 2002, 18, (2), pp. 227–239, doi: 10.1093/bioinformatics/18.3.227.

22 Kolch, W., Calder, M., Gilbert, D.: ‘When kinases meet mathematics: the systems biology of MAPK signalling’, FEBS Lett., 2005, 579, (8), pp. 1891–1895, doi: 10.1016/j.febslet.2005.02.002.

23 Hornberg, J., Binder, B., Bruggeman, F.J., et al.: ‘Control of MAPK signalling from complexity to what really matters’, Oncogene, 2005, 24, (33), pp. 5533–5542, doi: 10.1038/sj.onc.1208187.

24 Savageau, M.A.: ‘Biochemical systems analysis: A study of function and design in molecular biology’ (Addison-Wesley, Reading, Mass, 2009, 40th anniversary issue edition).

25 Fell, D.: ‘Understanding the control of metabolism, volume 2 of Frontiers in metabolism’ (Portland Press, London, 1997).

26 Cornish-Bowden, A.: ‘Fundamentals of enzyme kinetics’ (Wiley-Blackwell, Oxford, 2005).

27 Keener, J.P., Sneyd, J.: ‘Mathematical physiology I: cellular physiology’ (Springer, 2009, 2nd edn.), vol. 1, 21.

28 Voit, E.O.: ‘A first course in systems biology’ (Garland Science, New York and London, 2010).

29 Black, H.S.: ‘Stabilized feedback amplifiers’, Bell Syst. Tech. J., 1943, 13, (1), pp. 1–18, doi: 10.1002/j.1538-7305.1943.tb00622.x.

30 Nguyen, L.K., Matalanas, D., Croucher, D.D., et al.: ‘Signalling by protein phosphatases and drug development: a systems-centred view’, FEBS J., 2013, 280, (2), pp. 751–765, doi: 10.1111/1742-4658.2012.08322.x.

31 Ahmed, S., Grant, K.G., Edwards, L.E., et al.: ‘Data-driven modeling reconciles kinetics of ERK phosphorylation, localization, and activity states’, Mol. Syst. Biol., 2014, 10, (1), p. 718.

32 Hasenstaub, A., Otte, S., Callaway, E., et al.: ‘Metabolic cost as a unifying principle governing neuronal biochemistry’, Proc. Natl. Acad. Sci. USA, 2010, 107, (27), pp. 12329–12334, doi: 10.1073/pnas.0914886107.

33 de Atauri, P., Orrell, D., Ramsey, S., et al.: ‘Evolution of design principles in biochemical networks’, IEEE Proc. Syst. Biol., 2004, 1, (1), pp. 28–40, doi: 10.1109/IPSBN.2004.405133.

34 Albergante, L., Julian Blow, J., Newman, T.J.: ‘Buffered qualitative stability explains the robustness and evolvability of transcriptional networks’, eLife, 2014, 3, pp. 1–28, doi: 10.7554/eLife.02863.

35 Heijnen, J.J.: ‘Approximative kinetic formats used in metabolic network modeling’, Biotechnol. Bioeng., 2005, 91, (5), pp. 534–545, doi: 10.1002/bit.20558.

36 Wang, F.-S., Ko, C.-L., Voit, E.O.: ‘Kinetic modeling using S-systems and lin-log approaches’, Bioinformatics, 2016, 32, (3), pp. 438–447, doi: 10.1093/bioinformatics/btw652.

37 Cloutier, M., Middleton, R., Wellstead, P.: ‘Feedback motif for the pathogenesis of Parkinson’s disease’, IET Syst. Biol., 2012, 6, (3), pp. 86–93, doi: 10.1049/iet-syb.2011.0076.
