A Mathematical Framework for Describing and Analysing Gene Regulatory Networks

THOMAS MESTL,† ERIK PLAHTED† AND STIG W. OMHOLT‡§

† Department of Mathematical Sciences, Agricultural University of Norway, P.O. Box 5035, N-1432 Aas, Norway and ‡ Department of Animal Science, Agricultural University of Norway, P.O. Box 5025, N-1432 Aas, Norway

(Received on 28 March 1994, Accepted in revised form on 3 May 1995)

This paper presents a mathematical framework for describing and analysing gene regulatory networks by autonomous differential equations. It represents an improvement on existing frameworks in that it may handle a wider range of gene regulatory mechanisms. Gene regulatory networks are frequently threshold-dominated, i.e. genes are activated only when the concentration of certain gene products lie between definite thresholds. Here, the concept of regulatory domain is introduced to describe these regions in the phase space. To each regulatory domain is associated an indicator function whose value is 1 inside and 0 outside the domain. The indicator functions thus reflect the logical structure of the network. The sharp borders between the regulatory domains may be smoothed by replacing the logical step functions by continuous sigmoids or so-called logoid functions. A logoid function coincides with the step function outside a narrow interval around the threshold, and rises continuously from 0 to 1 inside it. Using logoids, the task of finding steady states is considerably simplified. A list of regions in phase space comprising all steady states lying close to a threshold is obtained by examining a certain type of matrix called the Logoid-Jacobian. In addition, this matrix leads to the conditions necessary for stability of the steady states. External signals may be conveniently incorporated in the form of Boolean variables. Thus the framework is well suited for studying gene regulatory networks both in single cells and multicellular systems.

1. Introduction

A predominant feature of biological (and other higher level) systems is the presence of nonlinear cause-and-effect relationships between the system variables. In many instances, one may talk about threshold relationships, i.e. below (or above) a certain level variable \( x \) has no influence on the behaviour of variable \( y \), while above (or below) this level the effect of \( x \) on \( y \) saturates rapidly to a constant level.

For example, in gene regulatory systems the setup of stable transcription complexes, mRNA stability, protein degradation and protein phosphorylation seem to a great extent to be based on threshold regulation in that several factors may have to be present above or below certain concentration levels before a regulatory mechanism becomes operative (Yagil & Yagil, 1971; Yagil, 1975; Latchman, 1990; Lawrence, 1992).

Despite the existence of numerical methods, the wish to develop analytical methods for investigation of complex gene regulatory systems may be motivated by the hope of obtaining a deeper insight into such systems through the mutually reinforcing development of analytical and numerical methods.

Descriptively, a mathematical framework of gene regulatory systems should be able to handle the basic features of threshold regulation of production as well as degradation of gene products. In addition, it should account for the fact that the elements in gene regulatory networks seldom depend on just one factor, and that the components frequently combine...
into regulatory complexes in a fundamentally non-additive way (Davidson, 1990; Latchman, 1990).

Mathematically, the framework also ought to permit the development of analytical methods not generally available for the localization of all stationary states, attractor types and basins of attraction.

During the last two decades analytical approaches for the study of gene regulatory systems dominated by threshold phenomena have received considerable attention (Kauffman, 1977; Rigney, 1977; Banks, 1978; Yamamoto, 1985; Bignone, 1988; Okamoto et al., 1989). The existence of mathematical descriptions of such systems can apparently be classified in three categories. In models based on differential equations, threshold phenomena have been described by sigmoid functions in a wide range of applications (Goodwin, 1965; Monod et al., 1965; Yagil & Yagil, 1971; Othmer, 1976; Banks & Mahaffy, 1978; Dabrushin et al., 1990).

However, the use of sigmoid functions makes it almost impossible to find steady states by analytical means, and, in general, numerical methods have to be applied.

A second approach has been to use discontinuous step functions as an approximation to sigmoid functions. Although this avoids the overall nonlinearity of sigmoids and leads to a simplification of the differential equations, it introduces unpleasant discontinuities which often make an analysis difficult (Glass, 1975a, 1977a, b; Snoussi, 1989; Snoussi & Thomas, 1993).

The third category is characterized by systems of Boolean or multivalued logical equations of varying complexity (Kauffman, 1969; Kauffman & Goodwin, 1990; Thomas & D’Ari, 1990; Prokudina et al., 1991; Tchuraev, 1991; Thomas, 1991). Replacing the differential equations by logical equations, a much simpler description of the system is obtained which, however, has a lower level of resolution. Since the logical variables only tell whether the system is below or above thresholds, trajectories can no longer be followed, and the location of many steady states can only be given qualitatively. In addition, a mathematical stability analysis at thresholds seems difficult.

Combining the strong features of the three approaches into a unified formalism, the framework presented here satisfies several of the descriptive and analytical requirements mentioned above.

2. Description of the Mathematical Framework

A regulatory factor frequently acts on other variables or itself at different concentration levels—hence several thresholds may be associated to it. All the thresholds for all variables taken together divide phase space into rectangular, open, but not necessarily finite boxes. When going to the limit of infinitely steep sigmoids, the threshold functions become step functions \( X_\theta = X(x, \theta) \), where \( X_\theta = 0 \) when \( x < \theta \), \( X_\theta = 1 \) when \( x > \theta \), and \( \theta \) is threshold number \( j \) of \( x \), the amount of concentration of the product of gene \( i \). The alternative case where the values above and below the threshold are interchanged is described by \( \tilde{X}_\theta = 1 - X_\theta \). In the analysis of actual models, other, continuous threshold functions may be introduced. However, the step function suffices in many situations, and is in any case well suited at an introductory level by the fact that it unravels the main regulatory structure of a system very efficiently.

With this choice of threshold function, \( x \) is presumed to be governed by the locally valid linear differential equation

\[
\frac{dx_i}{dt} = \gamma_i x_i, \quad i = 1, \ldots, n.
\]

Within each box, \( x \) is the constant production rate and \( \gamma \) the constant, relative degradation rate of the product of gene \( i \). In this way, the production and relative degradation rates become piecewise constant functions in phase space, i.e. \( x_i = x_i(\mathcal{X}) \) and \( \gamma_i = \gamma_i(\mathcal{X}) \), where \( \mathcal{X} \) is the set of all threshold functions. The open region in phase space given by the union of boxes with identical production or degradation rates for \( x \) is called a regulatory production or degradation domain for \( x_i \), respectively.

Each regulatory domain is represented by a Boolean function, which we call its indicator function \( I \), having the value 1 inside the domain and 0 otherwise. The indicator functions specifying production or degradation domains number \( r \) and \( s \) of \( x \) will be denoted \( P^r \) or \( D^s \), respectively. The intersection of the production domain \( P^r \) and the degradation domain \( D^s \) is the combined regulatory domain \( I^r \) of variable \( x \). It is also presumed that irrespective of any regulation there will always be a basal minimum degradation rate \( \gamma_0 > 0 \) of \( x \).

By these definitions the production and relative degradation rates \( z_i(\mathcal{X}) \) and \( \gamma_i(\mathcal{X}) \) can be given as a sum of disjoint terms representing the various domains:

\[
z_i(\mathcal{X}) = z_{i_1} P_1^r + z_{i_2} P_2^r + \cdots + z_{i_n} P_n^r, \quad (2)
\]

\[
\gamma_i(\mathcal{X}) = \gamma_{i_1} D_1^s + \gamma_{i_2} D_2^s + \cdots + \gamma_{i_n} D_n^s, \quad (3)
\]

where \( n \) and \( u \) are the number of production and degradation domains for \( x \), respectively. A regulatory system may then be described by the equations

\[
\frac{dx_i}{dt} = z_i(\mathcal{X}) - (\gamma_0 + \gamma_i(\mathcal{X})) x_i, \quad i = 1, \ldots, n. \quad (4)
\]

For example, suppose that the following differential equation, with step functions as threshold functions,
describes the rate of change of gene product $x_1$ in a system:

$$\frac{dx_1}{dt} = (a_{11} X(x_2, \theta_{21}) + a_{12} X(x_2, \theta_{22})) - (\gamma_{10} + \gamma_{11} X(x_1, \theta_{11}) X(x_2, \theta_{22})) x_1. \quad (5)$$

The gene product $x_1$ is produced at rates $a_{11}$ and $a_{12}$ in the production domains specified by $X(x_2, \theta_{21}) = 1$ and $X(x_2, \theta_{22}) = 1$, respectively. The relative degradation rate of $x_1$ depends on itself (autoregulation) and on the amount of factor $x_2$ in a non-additive way, reflected by the product of threshold functions. Inside the domain given by $X(x_1, \theta_{11}) X(x_2, \theta_{22}) = 1$, $x_1$ degrades at the relative rate $\gamma_{10} + \gamma_{11}$, while outside this domain it degrades only at the basal degradation rate $\gamma_{10}$. Using indicator functions eqn (5) can be rewritten as

$$\frac{dx_1}{dt} = (a_{11} I_1^{11} + a_{12} I_1^{12}) - (\gamma_{10} + \gamma_{11} I_1^{11}) x_1. \quad (6)$$

The regulatory domains of $x_1$ are depicted in Fig. 1, showing that the production (and degradation domains) are mutually disjoint. In the combined regulatory domain $I_1^{11}$ we have $a_1(x) = a_{11}$ and $\gamma_1(x) = \gamma_{10} + \gamma_{11}$, so eqn (6) can be rewritten as

$$\frac{dx_1}{dt} = [a_{11} - (\gamma_{10} + \gamma_{11}) I_1^{11}] x_1\quad + [a_{12} - (\gamma_{10} + \gamma_{11}) I_1^{12}] x_1\quad + [a_{11} - \gamma_{10} I_1^{10}] x_1\quad + [a_{12} - \gamma_{10} I_1^{12}] x_1. \quad (7)$$

Each pair of brackets encloses a linear rate function which is only active inside the corresponding regulatory domain. As will be seen below, expressing a regulatory system in the form given by eqn (7) simplifies the analysis of the so-called regular steady points of the system.

If $\theta_{21} < \theta_{22}$, the indicator functions $I_1^{11}$ associated with these combined regulatory domains can be found from Fig. 2 to be

$$I_1^{11} = X(x_1, \theta_{11}) X(x_2, \theta_{21}),$$

$$I_1^{10} = X(x_1, \theta_{11}) X(x_2, \theta_{22}),$$

$$I_1^{12} = X(x_1, \theta_{11}) X(x_2, \theta_{22}) X(x_2, \theta_{21}),$$

$$I_1^{10} = X(x_1, \theta_{11}) X(x_2, \theta_{22}) + X(x_1, \theta_{11}) X(x_2, \theta_{21}).$$

The step functions were only used to identify the regulatory domains and to allow a compact description. A conversion into a continuous system function is done simply by substituting the step functions by sigmoid functions of choice.

External signals of importance for the dynamics of a network may be conveniently incorporated in our framework through the indicator functions $I$ in the form of additional Boolean variables. As an example, consider the response of a cell exposed to a heat-shock. After a heat shock, the synthesis of most enzymes are shut down, while protective heat-shock proteins are being produced. In this case, the gene regulatory system may be divided into two different parts, one active at lower temperatures, the other at higher ones, denoted by $(x_1 - \gamma_i x)$ and $(x_2 - \gamma_j x)$, respectively. Let $I_0 = 1, I_1 = 0$ for $T < T^* \quad$ and $I_0 = 0, I_1 = 1$ for
$T > T^\circ$ be indicator functions where $T^\circ$ represents the temperature at which the heat shock is triggered. The system equation incorporating the external temperature signal can then be written as

$$\frac{dx}{dt} = (x_1 - \gamma_1 x_1)T^\circ + (x_2 - \gamma_2 x_2)I^\circ.$$  

When modelling gene regulatory processes in a single cell in a multicellular context, one may view the various signals from neighbouring cells as external inputs analogous to the heat-shock example above. However, in several situations, e.g. when studying determination and differentiation processes in cell clusters, it will be an advantage to model gene regulatory processes of several cells from the same cluster simultaneously. With reasonable assumptions on the inter-cellular interactions, the above single-cell framework could be generalized to cover multicellular systems by including a description of the overall regulatory structure and the dynamics of each cell. This can be achieved by a moderate and straightforward elaboration of the notational complexity of our framework.

3. Localizing Steady Points in Models based on the Framework

In biology, threshold relationships are usually continuous and therefore commonly represented by sigmoids. However, as shown in the previous section the use of step functions instead of sigmoid functions opens for a simpler description and helps to unravel the logical structure of the system. In the step function limit the set of steady points can be divided into two subsets, called regular steady points (RSPs) and singular steady points (SSPs) (Snoussi & Thomas, 1993). Mathematically, this distinction is very convenient. It may also have a biological interpretation (see Discussion).

RSPs are lying inside a combined regulatory domain and are always stable nodes. SSPs are always located at the boundary of a domain, i.e. the point of discontinuity for a step function, and can be nodes, saddle points or focuses. However, at the boundary of a domain the system is not defined, and the exact location and the nature of a SSP is usually not easily determined. This problem due to discontinuity is avoided by replacing the step function with a sigmoid function and let the sigmoid approach the step function in the limit. This alternating between sigmoid and step functions allows an almost exhaustive steady-state analysis.

3.1. Localization of Regular Steady Points (RSP)

In principle, RSPs are easy to handle. One method is to convert the differential equations into logical equations and analyse the resulting state table (Thomas & D’Ari, 1990). However, this gives only the regulatory domains containing RSPs but not their exact positions.

Another way is to obtain the RSPs directly from the differential equations. To see this, recall that inside a combined regulatory domain (CRD) the system equations (5) are strictly linear. The steady-state condition $\frac{dx}{dt} = 0$ yields $x^\circ = x_1/\gamma_1$. The point $x^*$ is called the focal point (Glass & Pasternack, 1978) of the CRD, because the system evolves towards this point. If $x^*$ does not lie in the CRD, the system will sooner or later leave it, and enter a new CRD that may have a different focal point. Thus, the focal point $x^*$ of a CRD is a RSP if and only if it lies inside the CRD, and in case it exists the RSP is a stable node. It follows that all RSPs can be found by locating the focal points of all CRDs.

3.2. Localization of Singular Steady Points (SSP)

A SSP is characterized by the fact that at least one of its components lies on a threshold, i.e. on a boundary of a domain. The existence of a SSP is normally not due to a very special choice of parameter values placing a focal point exactly on the boundary of a domain, but due to the nonlinear interactions present in the system. It is reasonable to assume that actual models describing regulatory systems are rather insensitive to slightly perturbed threshold values, and in the following we will assume that focal points do not lie in threshold planes.

Even though the transition from one set of parameters to another may be relatively abrupt in a real biological system, the step function represents a crude idealization. In order to make a better fit with reality and to avoid some of the analytical difficulties associated with sigmoids (and step functions), Plahte et al. (1994) suggested that what have been termed logoids should be preferred as threshold functions when analysing the steady states of threshold dominated systems. A logoid $Z_i = Z(x, \theta_\circ)$ has the value 0 or 1 as long as $x_i$ is between a certain distance from the threshold $\theta_\circ$, i.e. outside a so called $\delta$-interval. However, passing through this interval, the logoid increases continuously and monotonously from 0 to 1.

Thus $Z_i = Z(x, \theta_\circ) = 0$ for $x_i < \theta_\circ - \delta/2$, $Z_i$ increases continuously and monotonously from 0 to 1 for $x_i \in (\theta_\circ - \delta/2, \theta_\circ + \delta/2)$, and $Z_i = 1$ for $x_i \geq \theta_\circ + \delta/2$. Here $\delta > 0$ is a constant so that no $\delta$-intervals overlap. We also define $Z_i = 1 - Z_i$ and assume $dZ(x, \theta_\circ)/dx = c/\delta$ for $x_i = \theta_\circ$, where $c \geq 1$. When $\delta$ tends to zero, the logoid approaches the step function.

Substituting the step functions in the indicator functions by corresponding logoids, we obtain a
system of continuous differential equations. Instead of discontinuous change between sets of parameters at the boundaries, we now have a continuous transition. Due to the Boolean character of logoids, the system equations are still linear inside the "regular" part of a domain, i.e. RSPs can be found as before.

Thus logoids preserve the advantage of piecewise linearity within domains and restrict nonlinearity to small regions around thresholds or threshold intersections. The region in phase space where \( k \) "primary" variables lie in a \( \delta \)-interval, and all the logoids of the remaining "secondary" variables have fixed values, is called a \( \Delta \)-region of dimension \( k \) (Plahte et al., 1994). Steady points arising due to the non-linearity of the logoids can only occur in \( \Delta \)-regions. In order to find all SSPs one must in principle investigate all \( \Delta \)-regions. The rule below simplifies this investigation considerably (cf. the Appendix for formal details):

**Rule 1.** Let \( L \) be the matrix obtained by differentiation of the \( n \) rate functions with respect to the logoids \( Z_k \) in a \( n \)-dimensional \( \Delta \)-region obtained by selecting a particular threshold for each of the \( n \) variables. To each principal submatrix \( L_k \) of \( L \) there corresponds a \( k \)-dimensional \( \Delta \)-region whose primary variables correspond to the rows and columns in \( L_k \). If a row or a column of \( L_k \) is identically zero, there is no SSP in that \( k \)-dimensional \( \Delta \)-region in the limit \( \delta \to 0 \).

A proof of this rule is very simple. Assume row \( i \) in \( L_k \) is identically zero. Then \( z_i(\mathbf{x}) \) and \( \gamma_i(\mathbf{x}) \) do not vary in the \( \Delta \)-region considered. In the \( x_i \) direction the system evolves towards the focal point \( \mathbf{x}\ast \) as described in Section (3.1). Since by assumption no focal point is lying on a threshold, the system will not become steady in this \( k \)-dimensional \( \Delta \)-region in the limit \( \delta \to 0 \).

Suppose now column \( i \) in \( L_k \) is identically zero. Then the system equations describing the evolution inside the \( k \)-dimensional \( \Delta \)-region contains no threshold function for \( x_i \). This means the system does not recognize the crossing of \( \theta_i \), from which it follows that only for special parameter values it may happen that \( x_i \) becomes steady at \( \theta_i \) (in which case a bifurcation would occur). But by the assumption that our mathematical model is insensitive to slightly perturbed threshold values, there is no SSP in these \( \Delta \)-regions either.

The matrix \( L \) is called the **Logoid-Jacobian**. Each \( n \)-dimensional \( \Delta \)-region has its own Logoid-Jacobian matrix \( L_k \). If each variable has only one threshold, there is only one \( n \)-dimensional \( \Delta \)-region. Otherwise the number of \( n \times n \) Logoid-Jacobians is given by the product of the number of thresholds for each variable. In this case many \( k \)-dimensional \( \Delta \)-regions, where \( k < n \), will be encountered several times during investigation of all \( L_k \).

By the above rule, numerous regions, which never contain a SSP, irrespective of the parameter values, can be eliminated from further analysis. In addition, \( \Delta \)-regions which may contain a SSP only for a very special choice of parameters are sorted out. Thus we are left with a list of \( \Delta \)-regions which may, but need not, contain SSPs in the limit \( \delta \to 0 \). In these regions further analysis has to be performed.

Remembering that the logoid of a secondary variable has the value 0 or 1, while the logoid of a primary variable has a value between 0 and 1, and that in the limit \( \delta = 0 \) the primary variables will obtain their threshold values for that region, the determination whether there is a SSP in such a \( \Delta \)-region is straightforward:

**Rule 2.** Consider a \( \Delta \)-region of dimension \( k \) remaining after the application of Rule 1. Let \( x_i \) represent a primary variable, \( Z_i \) its logoid, \( \theta_i \) its threshold, and \( x_i \) a secondary variable. Let \( Z \) represent the \( k \)-dimensional vector with \( Z_i \) as components. In the limit \( \delta \to 0 \), \( x_i \to \theta_i \), and the steady state conditions for \( x_i \) are \( z_i(Z)/(g_\theta(Z)\theta_i) = 0 \). Assume there is a solution \( \mathbf{Z}\ast \).

Then the solution of the secondary steady state equations is \( x_i = z_i(Z\ast)/(g_\theta(Z\ast)) \). A SSP exists in this \( \Delta \)-region for small \( \delta > 0 \) if

\[(a) \text{ Det}(L_k) \neq 0, \quad (b) \text{ } 0 < Z_i\ast < 1, \quad (c) \text{ } x_i \text{ lies within the range of the } \Delta \text{-region.}\]

It should be noted that Rule 2 restricts the region in parameter space where a SSP exists. If a solution of the steady-state conditions exists, the rule gives the exact position of a SSP in the limit \( \delta \to 0 \). It is also worth mentioning that given the exact position of a SSP in the case \( \delta = 0 \), an approximate position for \( \delta > 0 \) can be calculated.

Having obtained all the SSPs, the next step is to decide whether they are stable or not. The eigenvalues of the Logoid-Jacobians provide a zero order approximation in \( \delta \) of the eigenvalues of the ordinary Jacobian. This leads to a much simpler analysis.

### 3.3. An Example

It is beyond the scope of this paper to give a detailed analysis of a specific model, though the usefulness of our framework will ultimately be determined through concrete model building. The main aim of the example below is to demonstrate the strength of our approach in the description and analysis of threshold-dominated systems compared to other approaches. Suppose we want to describe and investigate a regulatory network...
consisting of three genes expressing the proteins $x_1$, $x_2$ and $x_3$ with an interaction structure as depicted in Fig. 3, their production or relative degradation rates only changing when certain thresholds are crossed.

Assume that protein $x_1$ is produced at two different rates, $x_{1o}$ in production domain $P^1$ and $x_{12}$ in $P^2$, and that it always disintegrates with a basal relative rate of $\gamma_{10}$. The second protein $x_2$ is only produced in $P^2$ with rate $x_{21}$ and disintegrates with a basal relative rate of $\gamma_{20}$. In contrast, $x_3$ has a constant non-regulated production rate $x_{31}$, but its degradation is subject to regulation, with relative rate $\gamma_{30}+\gamma_{31}$ inside degradation domain $P^1$, and with relative rate $\gamma_{30}$ outside this domain. The system equation can then be written as:

$$\begin{align*}
\frac{dx_1}{dt} &= x_{11} P^1_1 + x_{12} P^2_1 - \gamma_{10} x_1, \\
\frac{dx_2}{dt} &= x_{21} P^2_1 - \gamma_{20} x_2, \\
\frac{dx_3}{dt} &= x_{31} - [\gamma_{30} + \gamma_{31}] x_3.
\end{align*}$$

Assume the indicator functions of regulatory domains are given by $P^1 = \bar{X}_{11}, P^2 = X_{11} X_{12} X_{21}, P^1 = X_{11}$ and $P^2 = \bar{X}_{11} \bar{X}_{21}$. Let the parameter values be $\theta_{11} = 1$, $\theta_{21} = 1$, $\theta_{12} = 1$, $\gamma_{11} = 1$, $\gamma_{12} = 2$, $x_{12} = 2$, $x_{32} = 4$, $x_{31} = 5$ and $\gamma_{30} = 2$, $\gamma_{31} = 4$. The system equation is then

$$\begin{align*}
\frac{dx_1}{dt} &= 2Z_{11} + 4Z_{12} Z_{21} Z_{31} - x_1, \\
\frac{dx_2}{dt} &= 4Z_{11} - 2x_2, \\
\frac{dx_3}{dt} &= 5 - [2 + 4Z_{11} Z_{21}] x_3,
\end{align*}$$

where $Z_\cdot$ could be a step function, a sigmoid or a logoid.

Assume $Z_\cdot$ is a logoid. To find the RSPs each box has to be investigated. Out of the eight boxes in this example only one, the box $x_1 > x_{11}, x_2 > x_{21}, x_3 > x_{31}$, is found to contain a RSP in the point $(x_{12}/\gamma_{12}, x_{21}/\gamma_{20}, x_{31}/\gamma_{30}) = 4, 2, 2.5$.

In this three-dimensional system with only one threshold for each variable there are 19 $\Delta$-regions (one three-dimensional, six two-dimensional and twelve one-dimensional) which have to be examined for SSPs.

The lower dimensional $\Delta$-regions are represented by lower dimensional principal submatrices $L_k$ in $L$. In general, the number of $k$-dimensional $\Delta$-regions corresponding to a $k$-dimensional submatrix $L_k$ is $2^{n-k}$. Since a nonzero $L_k$ is a necessary condition for the existence of a SSP, it is quite easy to read directly out from $L_k$ which $\Delta$-regions can be omitted from further analysis.

To localize SSPs we first calculate the Logoid-Jacobian,

$$\begin{equation}
L = \begin{bmatrix}
4Z_{21} Z_{31} & 4Z_{11} Z_{31} & -2 + 4Z_{11} Z_{21} \\
4 & 0 & 0 \\
4Z_{21} & 4Z_{11} & 0
\end{bmatrix},
\end{equation}$$

which has only one one-dimensional principal submatrix, $L_x(\theta_{11}, Z_{21}, X_{12}) = [4Z_{11} Z_{21}]$, which is nonzero only in the one-dimensional $\Delta$-region $(\theta_{11}, x_2 > x_{21}, x_3 > x_{31})$. Further, there are two two-dimensional and one three-dimensional principal submatrices which are nonzero only in the $\Delta$-regions $(\theta_{11}, \theta_{21}, x_1 > x_{12}), (\theta_{11}, x_2 < x_{21}, \theta_{32})$ and $(\theta_{11}, \theta_{21}, \theta_{31})$ respectively. Using Rule 1, we only need to investigate the solution of the steady state conditions $dx_i/dt = 0$ in each of these regions, and thereafter apply Rule 2.

For example, inside the $\Delta$-region $\Delta(\theta_{11}, x_{21}, \theta_{31})$ (i.e. intersection of three thresholds), the steady state conditions are

$$\begin{align*}
0 &= 2Z_{21} + 4Z_{11} Z_{21} Z_{31} - 1, \\
0 &= 4Z_{11} - 2, \\
0 &= 3 - 4Z_{11} Z_{21},
\end{align*}$$

which gives $Z_{11} = 1/2, Z_{21} = -1/2$ and $Z_{31} = 1/3$. Thus an SSP does not exist in this $\Delta$-region as $Z_{21}$ is not in the interval $(0, 1)$. Based on eqns (9) the remaining regions can be investigated in the same way, but in these cases values of the secondary variables will be obtained too. For example, in the $\Delta$-region $(\theta_{11}, x_{21}, x_3 > x_{31})$ we get

$$\begin{align*}
0 &= 4Z_{11} Z_{21} - 1, \\
0 &= 4Z_{11} - 2, \\
0 &= 5 - (2 + 4Z_{11} Z_{21}) x_3,
\end{align*}$$

which gives $Z_{11} = 1/2, Z_{21} = 1/2$ and $x_3 = 5/3 > \theta_{31}$, which is in agreement with the value $Z_{31} = 1$ in this region. Thus a SSP exists in this region in the limit $\delta \to 0$ and for small $\delta$.

By performing the same analysis of the remaining regions, it is found that also the $\Delta$-regions $(\theta_{11}, x_2 < x_{21}, \theta_{31})$ with $x_2 = 0.5$ and $(\theta_{11}, x_2 > x_{21}, x_3 > x_{31})$ with $x_2 = 4$ and $x_3 = 2.5$, contain SSPs. The eigenvalues of the Logoid-Jacobian in these points show that only the
SSP \((\theta_{11}, 0.5, \theta_{31})\) in the \(\Delta\)-region \((\theta_{11}, x_1 < \theta_{31}, \theta_{31} )\) is stable in the limit \(\delta \to 0\) and for small \(\delta\).

If the step functions in the example system are replaced by sigmoid functions, analytical solutions of the steady points are no longer attainable. However, using a perturbation method, approximate values for the steady points may be found for steep sigmoid functions (cf. Plahte et al. (1995) for a summary and an application).

Using the correspondence between differential equations and logical equations developed by Thomas & D’Ari (1990), the system can be described by a set of logical equations with the same logical structure as eqn (9), and analysed by the logical methods developed in their book. However, from this approach it is difficult to extract quantitative results.

An additional advantage of our framework is that due to the analytical solutions, an analysis can be performed even if parameter values are unknown. Indeed, one of the results of the analysis is restrictions on the parameter values necessary to ensure the existence of the steady points.

5. Discussion

The simplicity of the presented framework is based on the assumption that the interactions between variables can be represented by a sum of products of steep threshold functions only. In models where, in addition to threshold functions, other functional dependencies also occur, only the threshold part can be handled by our approach. This, however, will often lead to a simplification of the remaining system. As stated above, the presented framework is aimed at describing regulatory networks at a medium level of resolution where it is valid to treat a cell as a point with no extension, and where there are no time delays due to diffusion or transport processes. When addressing problems where such phenomena have to be considered, one needs mathematical descriptions at a higher level of resolution which incorporates additional phenomena like reaction-diffusion mechanisms or the occurrence of time delays. Thomas & D’Ari (1990) present a simple example demonstrating how diffusion may be treated by using a step function approach. However, we would like to stress that the present framework is appropriate for a lot of problems, and it can handle an immense regulatory complexity involving regulation of for example transcription, mRNA processing, mRNA transport, mRNA longevity, translation, protein longevity, protein phosphorylation and dephosphorylation. Furthermore, it allows formulation of models where the parameters can be interpreted biologically, as well as measured experimentally.

There may be a biological interpretation of RSPs and SSPs. A RSP for a gene product \(x\) is defined as the equilibrium concentration given by its rate of production and relative decay. Thus, there is no negative autoregulation or regulatory feedback structure between \(x\) and other gene products involved to ensure that \(x\) stays at the actual concentration level. However, other gene products may be active for ensuring its activation, continued production and deactivation. A positive autoregulation may also be used to ensure continued production. For example, a gene may be activated by one gene product, kept activated by its own product (staying at its RSP), and be deactivated by a third gene product. This is a simple, robust and energetically very efficient mechanism, and the fact that positive autoregulation is a frequent characteristic of gene regulation, seems to indicate that RSP are real entities.

On the other hand, a regulation leading to a RSP is not an efficient mechanism when there is need for rapid changes in concentration levels, and when many factors have to be kept at various concentration levels due to reasons of timing and tuning during the unfolding of for example a developmental process. Then positive as well as negative feedback elements have to be present, and this will automatically generate SSPs.

A system with only SSPs may be interpreted as one where all the stable concentration levels of the actual gene products (given by the stable SSPs) are determined by active positive and negative feedback regulation. Thus there is at least some reason to believe that RSPs as well as SSPs describe real regulatory phenomena in gene networks. Needless to say, this conjecture needs much more experimental evidence before it can be stated with some certainty.

An alternative mathematical framework to ours is the one developed by Thomas and co-workers (Thomas, 1978, 1991; Snoussi, 1989; Thomas & D’Ari, 1990; Snoussi & Thomas, 1993). Much of their recent efforts are based on the mathematical structure

\[
\frac{dx_i}{dt} = k_0 + \sum k_{ij} S_{ij}(x_j, \theta_j) - \gamma x_i, \tag{11}
\]

where \(S_{ij}(x_j, \theta_j)\) are step functions, \(x_j \) equals + or - to describe positive or negative regulation, \(k_{ij} \geq 0\) are production rates, \(\gamma_j > 0\) are the relative decay rates and \(\theta_j\) are the thresholds of \(x_j\). The sum runs from \(j = 1\) to \(n\), \(n\) being the number of variables. From eqn (11) and the logical equations for the corresponding logical variables, a number of results on the qualitative behaviour of such systems have been obtained.

However, eqn (11) does not include the phenomenon of non-additivity in a natural way. Although it may describe non-additive phenomena, the production
rates $k_i$ lose their biological meaning, and we consider this to be an unpleasant feature. Furthermore, in eqn (11) the degradation rates $\gamma_j$ are not under regulatory control. It would, however, be straightforward to apply the same regulatory structure to the decay terms as the production terms. But we do not know how much of their methodology for analysing systems based on eqn (11) would still be valid.

Concerning the analytical tools developed by Thomas and coworkers for studying the logical equations based on eqn (11), their recent result on logical identification of all steady states deserves attention (Snoussi & Thomas, 1993). Here they introduce the concept of characteristic state of a feedback loop. It is defined as the state for which each variable of the loop is located at the threshold value above which it is active in the loop considered. They show that among the singular states, only those which are loop-characteristic can be steady. This provides a considerable simplification of the analysis (Snoussi & Thomas, 1993).

How is their way of simplifying this analysis related to ours? In fact, the two methods turn out to be very similar. Snoussi and Thomas define a singular state to be a state where at least one variable is at its threshold value. Let us assume a state where $k$ out of $n$ variables are at a threshold value. Then this singular state is analogous to a $k$-dimensional $\Delta$-region. The number of $\Delta$-regions is identical to the number of singular states of a system.

It is nice to observe that there is a close relationship between the $\Delta$-regions satisfying Rule 1 introduced above and the loop-characteristic states. The Logoid-Jacobian $L_\Delta$ is made up of only the logoids that vary in the actual $\Delta$-region, the other logoids disappearing from $L_\Delta$ because they are constant. Thus $L_\Delta$ only comprises operative logoids. When a row of column in $L_\Delta$ is identically zero, there is no $k$-dimensional feedback loop in the corresponding $\Delta$-region, because if row $i$ is zero, $x_i$ is not influenced by any variable including itself, and if column $j$ is zero, $x_j$ does not influence any variable including itself. That means that all feedback loops being active in a given $\Delta$-region are contained in the corresponding Logoid-Jacobian. However, even if there is no $k$-dimensional feedback loop, all rows and columns in $L_\Delta$ may be non-zero. Thus we may be left with a greater number of $\Delta$-regions or singular states to investigate than by the method of Snoussi & Thomas (1993). If there exist SSPs in these additional $\Delta$-regions, all of them have to be bifurcation points since $\text{Det}(L_\Delta) = 0$. As a consequence, a separate bifurcation analysis has to be performed in order to determine the dynamics around the SSP.

Hence, the statement by Snoussi and Thomas that only singular states that are loop-characteristic can be steady, appears to be not completely true. But this is not a major problem, and we conclude that Rule 1 and the method of detecting loop-characteristic states are very similar. Furthermore, Rule 2 gives simple sufficient conditions for a singular state to be steady.

In many practical situations it is very convenient, if not necessary, to make use of a computer to perform the analysis even if analytic methods for analysing complex gene regulatory systems are available. Thomas and coworkers have computerized their procedures (Thieffrey et al., 1993). Fortunately, this can also be done with our procedures.

We conclude that our approach represents, to some degree, an improvement on the approach advocated by Snoussi & Thomas (1993), as

(i) our framework has a somewhat higher descriptive capability, i.e. allows for non-additivity and varying degradation rates,

(ii) we do not need to perform the analysis in the logical realm and have developed comparable analytical methods for analysing steady states,

(iii) our methods can be computerized as easily as theirs.

However, the approach developed by Thomas and coworkers may be appropriate in many cases, and that the logical method per se as a means to obtain preliminary qualitative information about complex regulatory systems is a very valuable tool.

The analytical part of this paper has focused on methods of localizing steady points. They are not restricted to gene regulatory networks but may also be applicable to other kind of systems, such as food web models (Plahte et al., 1995). Steady points are important characteristics of dynamical systems, but limit cycles and the character of attractor basins are also important aspects. Ideally, a mathematical framework of gene regulatory networks should have a mathematical structure such that these aspects can also be handled by analytical methods. For the moment, we are not able to say to what extent our framework allows analytical methods to be developed. However, we have recently obtained some encouraging results concerning methods for localization of limit cycles and periodic trajectories (Mestl et al., 1995).

We would like to thank one of the referees for very constructive comments. This work was partly supported by the Norwegian Research Council by a grant to Thomas Mestl.
REFERENCES


APPENDIX

Derivation of the Logoid-Jacobian

In the following, a $k$-dimensional $\Delta$-region in an $n$-dimensional system ($k < n$) is considered. Without loss of generality it can be assumed that variables $x_1, \ldots, x_k$ are the primary variables, i.e. $x = \theta$, whereas the remaining are secondary variables, i.e. $x < \theta$ or $x > \theta$. Let $Z$ denote a set of logoid threshold functions. The differential equation (5) is written in compact vector and matrix notation as

$$\frac{dx}{dt} = z(Z) - [B + G(Z)]x = F(x),$$

where $B$ and $G(Z)$ are diagonal matrices with $\gamma_a$ and $\gamma_r(Z)$ as elements, respectively. Assume that in the limit $\delta \to 0$ eqn (A.1) has an SSP inside the $k$-dimensional $\Delta$-region. The Taylor representation of eqn (A.1) in this $\Delta$-region is

$$\frac{dx}{dt} = F(\theta) + F'(\theta)(x - \theta) + R(x),$$

where $F'(\theta)$ denotes the Jacobian. Higher order terms are denoted by $R(x)$. The Jacobian is given by

$$F'(\theta) = \frac{\partial [a(Z) - (B + G(Z))x]}{\partial Z} \frac{\partial Z}{\partial x} - [B + G(\theta)].$$

From the definition of a logoid we have $\frac{\partial Z(\theta, \theta)}{\partial Z} = \delta$, so

$$F'(\theta) = \frac{\partial}{\partial Z} [a(Z) - G(Z)\theta] - [B + G(\theta)].$$

The logoids of the secondary variables have constant values $0$ or $1$, hence their derivatives are always zero and do not contribute to the Jacobian. That is, the first term in (A.3) is a matrix of rank less or equal to $k$, and is due to the derivatives of the $k$ logoids of the primary
variables only. This is what we call the Logoid-Jacobian \( L_k \). The second term in (A.3) is a diagonal matrix of rank \( n \) with negative diagonal elements only. If the steepness of the threshold functions is low, i.e. \( \delta \gg 0 \), the Logoid-Jacobian \( L_k \) tends to zero and the Jacobian is approximately \( J \approx -[B + G(\theta)] \), i.e. always stable. However, if the steepness of the logoids increases, \( \delta \to 0 \), the stability of the Jacobian will now entirely depend on the stability of the Logoid-Jacobian \( L_k \). That is, if in the limit \( \delta \to 0 \) the Logoid-Jacobian is stable, then the corresponding Jacobian is also stable at the actual SSP.