# Positive and negative feedback: striking a balance between necessary antagonists

Olivier Cinquin *, Jacques Demongeot •
TIMC-IMAG-CNRS UJF-INPG & IUF •, Faculty of Medicine
38706 LA TRONCHE
FRANCE

* To whom correspondence should be addressed, at Olivier.Cinquin@imag.fr

Journal of Theoretical Biology **216**(2), pp229-241 (2002)

**Abstract**

Most biological regulation systems comprise feedback circuits as crucial components. Negative feedback circuits have been well understood for a very long time; indeed, their understanding has been the basis for the engineering of cybernetic machines exhibiting stable behaviour. The importance of positive feedback circuits, considered as "vicious circles", has however been underestimated. In this article we give a demonstration based on degree theory for vector fields of the conjecture, made by René Thomas, that the presence of positive feedback circuits is a necessary condition for autonomous differential systems, covering a wide class of biologically relevant systems, to possess multiple steady states. We also show ways to derive constraints on the weights of positive and negative feedback circuits. These qualitative and quantitative results provide respectively structural constraints (*i.e.* related to the interaction graph) and numerical constraints (*i.e.* related to the magnitudes of the interactions) on systems exhibiting complex behaviours, and should make it easier to reverse-engineer the interaction networks animating those systems on the basis of partial, sometimes unreliable, experimental data. We illustrate these concepts on a model multistable switch, in the context of cellular differentiation, showing a requirement for sufficient cooperativity. Further developments are expected in the discovery and modelling of regulatory networks in general, and in the interpretation of bio-array hybridisation and proteomics experiments in particular.

Keywords: positive feedback, multistationarity, multistability, stability, regulation, interaction networks, switch, cellular differentiation

# 1 Introduction

The quantity of biological information available for analysis is expanding at a tremendous rate, especially information obtained from molecular genetics techniques, but the methods of analysis are lagging behind. One of the most interesting challenges from the theoretical and practical points of view is to unravel the networks of regulation, involving for example metabolism or cell differentiation. These networks are expected to exhibit stability (since loss or change of cell differentiation are usually observed under unfrequent conditions, respectively anaplasia and metaplasia) and multistationarity (as there are many different cell types in metazoans), and we thus investigate both aspects.

Our approach is to find conditions that are fulfilled by biological systems; our goal is to derive both new theoretical concepts, and practical constraints on interactions networks, to make more tractable the task of reverse-engineering such networks from partial, sometimes unreliable, experimental data. To become viable, regulation models are in need of *quantification*, as pointed out by Koshland (1998). Boolean networks do not meet this need entirely, and biological systems have not been proven to exhibit behaviours easily modelled in a boolean fashion (Kringstein *et al.*, 1998, have even shown in detail that eukaryotic transcription can give a graded response to concentrations of molecules); we thus consider continuous systems, which are much more difficult to deal with, but which provide much more accurate modelling of real systems.

There have been attempts to study properties of generic regulation systems of dimension 1 or 2; for example, a beautiful characterisation of multistability has been achieved by Cherry & Adler (2000). But higher-dimensional systems display much greater complexity, and probably are the rule rather than the exception in real-world systems (it is becoming clearer and clearer that signalling pathways are extremely intricate, see for example Jordan *et al.*, 2000). It seems difficult to scale up the results in 2 dimensions to higher dimensions, which is why we try not to make assumptions about the dimensions of the systems we study, even if it is at the cost of sometimes less powerful results.

In the first part of this article, we study steady states of a generic system, and derive qualitative and quantitative properties of the system according to its behaviour at these steady states. In the second part, we build on these results to tackle "Delbrück's conjecture"; Max Delbrück proposed in 1948 that cell differentiation could be established by a unique regulating system having distinct attractors (Delbrück, 1949). In 1980, R. Thomas made the conjecture that "the presence of a positive circuit in the logical structure

of an autonomous differential system is a necessary, although not sufficient, condition for multistationarity" (Thomas, 1981). The conjecture has been proven in the case of boolean networks (see Aracena *et al.*, 2000, and Demongeot *et al.*, 2000b), and partial demonstrations in the case of continuous systems have been given (Plahte *et al.*, 1995, as well as Snoussi, 1998, and Gouze, 1998, introduced by Demongeot, 1998), but their usefulness is deeply undermined by restrictive hypotheses. We propose another demonstration which introduces softer constraints. Finally, we study a multistable switch to illustrate our results.

# 2 Preliminaries

In the following, we consider an open domain $D \subset \Re^n$, a function $F \in C^\infty(D, \Re^n)$, a real interval $L$, and a function $x \in C^1(L, D)$ such that

$$\forall t \in L, \frac{\mathrm{d}x(t)}{\mathrm{d}t} = F(x(t)) \tag{1}$$

Vector $x$ describes the state of the system, and vector $F(x)$ gives the direction in which the system will evolve when it starts at point $x$ of its state space. We will use the abbreviation ss for stationary states of the system, *i.e.* the points where $F$ vanishes, and where the system maintains an equilibrium. We will suppose throughout this paper that no ss of $F$ is degenerate.

Let $J_F(x)$ be the Jacobian matrix of $F$ at $x \in D$ (*i.e.* the matrix of partial derivatives of the field components with respect to system variables), and $\mathrm{sgn}(J_F(x))$ the matrix obtained by taking the sign of each component of $J_F(x)$ (+1, 0, or −1). $\mathrm{sgn}(J_F(x))$ (respectively $J_F(x)$) can be considered to define an oriented, weighted *interaction graph* on the set $I = \{1, .., n\}$, in which an arc points from $i$ to $j$ if and only if $\frac{\partial F_j}{\partial x_i}(x) \neq 0$ (*i.e.* if variable $j$ depends on variable $i$), the associated weight being $\frac{\partial F_j}{\partial x_i}(x)$, *i.e.* the strength of the dependency of $F_j$ on $x_i$.

There is a circuit involving nodes $i_1, .., i_k$ if and only if $\prod_{j=2}^{k+1} \frac{\partial F_{i_j}}{\partial x_{i_{j-1}}} \neq 0$, where $i_{k+1} = i_1$. A same node is not allowed to appear more than once in the graph circuits considered in the following (*i.e.*, in the previous example, $\forall\ p, q \in \{1..k\}$ s.t. $p \neq q,\ i_p \neq i_q$). In the interaction graph, a node is part of a circuit if and only if the corresponding variable is part of a feedback circuit, *i.e.* if its own value affects the way it will change, either directly or through a chain of dependencies. Circuits comprising a single arc (from a vertex to itself) are allowed, and correspond to a diagonal term in the Jacobian, and a variable exerting a direct influence (positive or negative) on

itself. The sign of a circuit (positive or negative) is defined as the sign of the product of the weights of the associated arcs; *i.e.*, a circuit is negative if it has an odd number of negative interactions (in which case the corresponding variables are under negative feedback control), and positive otherwise (in which case the corresponding variables are under positive feedback control). Finally, the weight of a circuit is defined as the *product* of the weights of its arcs; the higher the weight, the higher the feedback intensity.

Note that, except in the trivial linear case (which doesn't give rise to interesting behaviours such as non-degenerate multistationarity), the Jacobian matrix is not constant within the domain $D$; it is thus possible, and likely, that the weights and the signs of the circuits are not constant, and even that the structure of the interaction graph itself is not constant. Note also that if there are two distinct mechanisms by which a variable depends upon another (these mechanisms possibly being antagonistic), it is the sum of the corresponding derivatives which will appear in the relevant Jacobian term. For example, if a protein whose concentration corresponds to $x_i$ exerts a positive but saturable effect on its own synthesis (positive autocatalysis), and undergoes exponential decay (negative autocatalysis), such that

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = -x_i + v\frac{x_i}{K + x_i},\ v > 0,\ K > 0,$$

then the overall autocatalysis for $x_i$, corresponding to the Jacobian term $J_F(x)_{i,i}$, will be positive for low values of $x_i$ (provided that $\frac{v}{K} > 1$), and negative for high values of $x_i$.

Finally, the following terminology will be useful in later sections:

- A matrix $M$ can be *decomposed into circuits* if there exists a subset of the arcs of its interaction graph such that considering this subset, every vertex is part of a single circuit; it can be easily shown that a matrix can be split into circuits if and only if there is a non-zero term in the development of its determinant
- $F$ is *inward-pointing* if and only if there exists a domain containing all steady states of the system, such that the frontier $\partial$ of the domain is diffeomorphic to the unit ball of $\Re^n$ (meaning intuitively that the frontier is smooth and its shape is sufficiently close to that of a ball), and such that $F$ points strictly toward the interior of the domain at points of $\partial$ (this means that if one chooses a smooth vector field $\chi$ normal to $\partial$ and respecting its canonical orientation, then $\forall x \in \partial, < \chi, F > (x) < 0$)
- We will call a ss $s$ *stable* if and only if it is locally asymptotically stable, *i.e.* if all eigenvalues of $J_F(s)$ have a strictly negative real part.

# 3 Previous results

Plathe *et al.* (1995) published the first attempt to show an existence result for positive circuits; however their demonstration does not take into consideration the case of a coefficient of $J_F$ vanishing without changing sign, and excludes without biological discussion the case of coefficients changing signs concomitantly. Snoussi derived independently a demonstration formalising that of Plathe *et al.*, without going into the detail of non-constant sign of the Jacobian. Gouze proposed another demonstration in which he also assumed that $\mathrm{sgn}(J_F)$ was constant throughout $D$, and that $D$ was convex, assumptions on which his demonstration crucially depended; under the supplementary assumptions that the vanishing set of $F$ comprised 2 or more isolated points of $D$, he showed that the interaction graph of $F$ had a positive circuit.

The hypothesis that $\mathrm{sgn}(J_F)$ is constant is a very restrictive one: it means that partial derivatives of $F$ may not cancel anywhere in $D$, *i.e.* that if variable $i$ exerts a positive (respectively negative) influence on variable $j$ for some state of the system, that influence remains positive (respectively negative) for all states of the system.

An example of a situation in which the sign of $J_F$ is not constant is that of the well-studied arabinose operon in *E. coli*: as a monomer, araC represses transcription of the arabinose operon but as a dimer, it activates transcription. The presence of arabinose induces dimerisation of araC, and thus induces transcription of the operon. If one considers total araC and arabinose concentrations as well as operon transcription as state variables, transcription increases as [araC] increases in the presence of arabinose, but decreases as [araC] increases in the absence of arabinose; the influence of [araC] on transcription thus changes from positive in the presence of arabinose to negative in the absence of arabinose.

More generally, it is possible to build systems for which the sign of the Jacobian matrix is not constant by considering a gene under transcriptional control of two binding sites for the same protein, one site having a high affinity for the regulating protein and activating transcription, and the other site having a lower affinity for the regulating protein and repressing transcription, as is the case for the TATA-Binding Protein of *Acanthamoeba* (Ogbourne & Antalis, 1998), or for the $\lambda$ repressor in *E. coli* (Ptashne, 1986). A similar situation can also be met with proteins having an activity-enhancing and an activity-repressing site for the same molecule, as is the case for example for phosphofructokinase with ATP.

It is possible to enhance Gouze's and Snoussi's theorems by only supposing that $\mathrm{sgn}(J_F)$ is constant in a convex neighbourhood of a segment joining

two stable ss, but it is hard to think of justifications for such a hypothesis. In the following, we thus attempt to relax the constant-sign hypothesis. In order to do so, we have to introduce a new condition, but for which we have a natural, biological justification.

# 4 Steady states, eigenvalues and circuits

Stability at stationary states is usually studied in terms of eigenvalues of the Jacobian matrix. This is the most appropriate approach from the mathematical point of view, but not from the biological one, from which one considers interactions between different components, and especially feedback circuits. The following lemma relates matrix eigenvalues to qualitative, circuit-related properties.

**Lemma 1.** *Let $M$ be a square, non-degenerate matrix of dimension $n$. If $\mathrm{sgn}(\det M) \neq (-1)^n$, then $M$ has a positive circuit. If $M$ is degenerate, then either $M$ has no decomposition in circuits or $M$ has a positive circuit.*

The proof of lemma 1 is technical but simple; it has been given by Plahte *et al.* (1995). In some cases, it is possible to derive a lower bound on the weight of the positive circuit for the sign of the determinant to be different from $(-1)^n$. We do not show the detail of the demonstration because the result one gets is heavily dependent on the particular structure of the matrix. However, we will illustrate such a lower bound on positive feedback in section 6.

The following corollary (whose proof can be found in the appendix) gives an application of lemma 1 which has an intuitive interpretation: if the system's motion in a certain direction is amplified (the direction being associated to a strictly positive eigenvalue), there must be some positive feedback in the matrix.

**Corollary 1.** *Let $M$ be a square, non-degenerate matrix of dimension $n$. If $M$ has at least one real, strictly positive eigenvalue, then $M$ has a positive circuit.*

Note however that the existence of an eigenvalue with a strictly positive real part is not a sufficient condition for the existence of a positive circuit, as illustrated by the matrix $(0, 1, 0; 0, 0, 1; -1, 0, 0)$, which has a negative circuit as single circuit, but two complex eigenvalues with strictly positive real parts. Thus, instability at a steady state is not necessarily provided by a positive feedback circuit.

Using the same mathematical approach as for the above lemma, and as was pointed out by Levins (1975), one can derive constraints on the weights of feedback circuits at a stable steady state:

**Lemma 2.** *Let $s \in D$ be a* stable *ss, $M = J_F(s)$ be the Jacobian matrix of $F$ at point $s$. We have*

$$\forall\ l = 1..n,\ (-1)^l \frac{\sum_{L \subset I, |L|=l} \det M_L}{\binom{n}{l}} > 0$$

One can see directly from the demonstration of lemma 1 that *positive* feedback circuits contribute *negative* terms to the sum in lemma 2. *Other things being equal*, positive feedback circuits are thus destabilising above a certain weight (in other cases, positive feedback can actually be stabilising and negative feedback destabilising, see Cinquin & Demongeot for details). This allows to derive constraints on the weight of positive feedback circuits at stable stationary states, but one must bear in mind that the weight of feedback circuits cannot usually be altered without displacing stationary states.

Thus, if we consider a ss of $F$ $s$, we have the following possibilities:

- $s$ is unstable, with a *real* eigenvalue being at least partly responsible for the unstability; in this case, there is a positive circuit in the Jacobian matrix
- $s$ is unstable only because of pairs of complex eigenvalues with positive real parts. In this case, the unstability stems from groups of 2 state variables which oscillate in an expanding fashion. In many regulation systems there is a key "switch" which integrates diverse information sources, and activates effectors when it reaches a critical threshold. The switch has to be upregulated when it reaches the threshold, so that there is a clear distinction between the "on" and "off" states. If the source of unstability is the switch-molecule, it is therefore unlikely that it is associated to a complex eigenvalue, because there would be an oscillation above and under the threshold.
- $s$ is stable; in this case, it is possible to derive constraints on weights of the circuits, following the approach detailed by Levins (1975).

The fact that there can be unstable ss without real, positive eigenvalues is interesting *per se*, but will not affect the extent of the existence results presented in the following sections.

# 5 Proof of Thomas's conjecture

## 5.1 Proof for inward-pointing fields

Steady states are the points of state space at which it is easiest to derive conditions about the Jacobian matrix of the system. Studies of the bifurcations of small-dimension systems often involve an unstable steady state. We follow the same approach for higher dimensions: having shown that certain types of unstable ss provide us with the existence of a positive circuit, we now proceed by showing that, under certain conditions, such an unstable ss can indeed be found, which will prove Thomas's conjecture under those conditions.

Theorem 1, which can be found in the appendix, makes use of degree theory and of the value of an integral involving $F$, which can be approximated by numerical methods. It is difficult to state beforehand whether our theorem will be applicable to a given vector field, but we think it should be able to cover a wide variety of systems. However our theorem does apply to all systems such that $F$ is inward-pointing, providing the following corollary:

**Corollary 2.** *If $F$ is inward-pointing, and if $F$ has at least 2 stable ss, then $J_F$ has a positive circuit.*

## 5.2 Making fields be inward-pointing

We will show in the following that, given biologically reasonable hypotheses, any field can be altered to be inward-pointing (and thus be subject to Corollary 2), in a way that doesn't add positive circuits.

We will consider a system fulfilling two requirements. Firstly, we will take the domain $D$ to contain the cone $\{x \in D \text{ st } \forall\ i \in I\ x_i > 0\}$, although this is mainly a matter of simplifying the demonstration.

Secondly, we will suppose that when a state variable vanishes there is a drive to replenish it, even if only by a very small amount. If we consider a set of genes under transcriptional control of one another, this corresponds biologically to the fact that repression is not "perfect", and that expression is "leaky". This is a well-known phenomenon in prokaryotes, and it has been suggested (Bird, 1995) that higher-order organisms developed new mechanisms to enhance repression because of their increased number of genes; however, it seems that even with mechanisms involving chromatin-structure and DNA methylation, there is still some expression-leakage (Chelly *et al.*, 1991).

More generally, specificity in biochemical processes is hardly ever perfect, and always entails a compromise between high precision and energetic price

(as illustrated by the proof-reading mechanism in DNA duplication, discussed by Savageau & Freter, 1979). Also, according to the thermodynamic laws describing the binding of molecules, it is impossible to reach a state in which a site (such as an operator site) is genuinely saturated.

When the concentration of a molecule approaches 0, the enzymatic mechanisms of its destruction or alteration should be proportional to its concentration, while its basal rate of production "leakage" is not dependent on its concentration. It would thus seem that when the concentration is low enough, there is a drive toward replenishment. If there is however a set of molecules dependent on one another as regards their production, it is possible that there is a steady state where all concentrations vanish; but such a steady state would not be stable, and could be eliminated from the study by restricting the domain of study to the quasi-cone $\{x \in D \text{ st } \forall\ i \in I\ x_i > \epsilon\}$, with $\epsilon$ suitably small.

These assumptions do not necessarily make $F$ an inward-pointing field, which would probably be too strict a requirement for our result to apply to many biological systems. However, as will be detailed in the appendix, $F$ can be "bent" into an inward-pointing field $G$, while preserving properties related to positive circuits. $G$ is no longer relevant from the biological point of view, but this is not important since it has been bent far enough from the origin to have at least two stable ss, and has no more positive circuits than $F$. Corollary 2 can be applied to $G$, showing that it has a positive circuit, and the same is thus true of $F$.

What's more, the way $G$ is built shows that the distance of the positive circuit found for $F$ to the origin is at most $\sqrt{n}$ times the distance of the second most distant stable ss of $F$ to the origin.

This demonstration is valid for example for all kinds of systems modelled by fields $F$ involving classical Michaelis-Menten, Hill, or Monod-Wyman-Changeux kinetics.

# 6 A model for a multistable switch

## 6.1 Introduction

Bistable switches have been studied extensively in the theoretical biology literature, and such switches have even been artificially built into biological organisms (see Gardner *et al.*, 2000, for a prokaryotic switch, and Becskei *et al.*, 2001, for a eukaryotic switch), displaying the properties that were expected from theoretical studies. However, to our knowledge, no study of $n$-stable switches has been proposed for $n > 2$ (a structure leading to

such systems was mentioned by Kling & Székely, 1968, in a different context). Such switches are much more resistant to mathematical analysis than bistable switches, but one would expect that they can appear in biological systems where a decision more complicated than a binary decision has to be made, especially in the case of cellular differentiation. Demongeot *et al.* (2000a) have for example proposed a 3-stable switch between the apex and cotyledonary bud growth variables, in order to explain storage and recall functions observed in *Bidens pilosa* L.

While it has long been clear that cellular differentiation is a dynamic phenomenon in developing organisms, it has only recently been widely recognised that differentiated cells are not always in a definitively fixed state, recorded for example by chromatin remodelling or DNA methylation. Indeed, there are many examples in which cells can be observed to undergo transdifferentiation or metaplasia in non-pathological situations (Slack & Tosh, 2001), regeneration in urodele amphibians is thought to involve dedifferentiation (Brockes, 1997), and in many cases stem cells have shown wider differentiation potential than expected (see for example Ferrari *et al.*, 1998, Bjornson *et al.*, 1999, and Clarke *et al.*, 2000). This points to autonomous differential systems, exhibiting switch-like behaviour, as a relevant type of mathematical model to study cellular differentiation. Biologically, such switches, while stable on their own, could be made to change states if they underwent the right perturbation, for example by external signalling factors.

One can think of cellular differentiation as following two different kinds of pathways; one extreme would be a hierarchical, sequential series of binary decisions, as illustrated in Figure 1a (such a pathway could correspond to the progressively more specialised progeny of stem cells; it would involve a low number of interactions between all components). Another extreme would be an "instantaneous" decision between all the possible outcomes (depending for example on external stimulations), as illustrated in Figure 1b (such a pathway structure makes it necessary that every single component has an inhibitory effect on all other components). These two proposed pathways are compatible with combinatorial choice of cell fate by different factors; we do not discuss this possibility because the way in which the cell-type is "read-out" does not influence the number of steady-states of the system. Biological studies haven't yet given a precise answer to how differentiation occurs, but it seems quite possible that real-world differentiation systems come half-way between these two extremes, and in particular that there are steps that involve "decisions" between at least 3 possible outcomes. It is thus very interesting to study $n$-stable switches with $n > 2$, since the study of hierarchical, sequential systems comprising such switches can easily be reduced to the study of $n$-stable switches, because the decision process can

be split up between the different decision steps. What's more, the study of such switches will give us the opportunity to illustrate remarks we made previously about the weight of positive and negative feedback circuits.

## 6.2 Results

In the following, we consider $n$ proteins, whose concentrations are denoted by $x_1$, .., $x_n$. Generalising the type of system proposed by Monod & Jacob (1961), and the type of equation that was used by Gardner *et al.* (2000), we suppose that each of these proteins undergoes exponential decay (be it natural or promoted by proteases functioning far from saturation), and that it inhibits the synthesis of other proteins in the switch, with cooperativity $c$ (see below for a discussion of $c$). Cooperativity and repressor site binding characteristics are taken to be the same for all proteins; repression is modelled by a simple, Michaelis-like function. Concentrations $x_1$, .., $x_n$ are normalised with respect to the Michaelis constant. $\sigma$ denotes the strength of unrepressed protein expression, relative to the exponential decay. The interaction graph of the system is of the type illustrated in Figure 1b. We get the following system of differential equations:

$$\begin{aligned}
\frac{\mathrm{d}x_1}{\mathrm{d}t} &= -x_1 + \frac{\sigma}{1+x_2^c+..+x_n^c} \\
&\dots \\
\frac{\mathrm{d}x_n}{\mathrm{d}t} &= -x_n + \frac{\sigma}{1+x_1^c+..+x_{n-1}^c}
\end{aligned}$$

Such a system could have evolved from duplication of genes involved in a bistable switch, and subsequent differentiation of repressor sequence and repressor-binding properties. It is trivial to show that the field defined by the system is inward-pointing on a domain of biological interest.

The differential system is difficult to study because there is no algebraic way to solve multivariate polynomial systems involving generic parameters (but it is possible to show the existence of $n$ ss which one would expect, see appendix). However, there are properties one can derive from general considerations, and simulations can be used to study systems with specific values of the parameters.

We now illustrate the fact that positive feedback, stemming from the interaction circuits between any single pair of proteins, must outweigh the negative feedback stemming from the exponential decay of the proteins.

In the general case, the system has a steady state where $x_1 = x_2 = .. = x_n$; in order to have a switch, such a steady state should clearly be unstable. As shown in the appendix, this requires both the cooperativity $c$ to have a lower bound (determined by the dimension of the system), and the strength of the

positive feedback circuits (which are proportional to $\sigma$ at any given point) to be strong enough:

- $$\sigma > (c-(n-1))^{-\frac{1}{c}} + (n-1)(c-(n-1))^{-\frac{c+1}{c}} \quad (2)$$

- $$c > n-1 \quad (3)$$

This theoretical study allows to derive constraints on the system if it is to exhibit switch-like behaviour, but doesn't show that it actually does exhibit such a behaviour. However, computer simulations (data not shown) indeed show that the proposed system, for values of $n$ up to 6, and such that $c$ and $\sigma$ match the conditions detailed above, has $n$ (and no more) stable steady states, each corresponding to a high concentration for one of the proteins, and a low concentration for all the other ones. We expect this result to hold for all values of $n$.

## 6.3 Discussion

In this specific example, because of the simple structure of the network, high values of the weight of the positive feedback circuits does not seem to affect the existence and stability of ss. In other cases, however, positive feedback can have adverse effects above a threshold value (see Cinquin & Demongeot); in those cases lemma 2 should allow to derive useful constraints.

Our demonstration of the existence of a positive feedback circuit doesn't give a precise location result, but intuitively it corresponds to the fact that there is a "destabilising force" that pushes the system to amplify small differences when protein concentrations are close to an unstable steady state; such unstable states necessarily exist if there are other stable steady states, and if the field is inward pointing. In the case of the model switch we propose, these unstable steady states would correspond to states in which some of the concentrations are equal (these states have to be unstable if we want switch-like behaviour).

Equation 2 from the previous subsection shows in particular that the condition $\sigma > 1$ must be met for the system to be switch-like (details not shown); since $\sigma$ is a measure of the unrepressed synthesis rate of each protein (*i.e.* in the absence of all other proteins) relative to the decay rate, this means that the unrepressed rate of synthesis (providing positive feedback circuits) must be stronger that the rate of decay (providing negative feedback in the

system). If the cooperativity $c$ is not an integer, as it tends to its lower bound $n-1$, the synthesis rate $\sigma$ must tend to infinity to achieve instability of the ss where all concentrations are equal. As the cooperativity $c$ tends to infinity, the lower bound on $\sigma$ tends to 1.

Equation 3 is a generalisation of a property of bistable switches studied by Gardner *et al.* (2000), which have been shown to require cooperativity strictly superior to 1 for one of the repressors (that requirement being a specific example of a more general property derived by Cherry & Adler, 2000). Higher dimensional systems of the same type need greater cooperativity to achieve switch-like behaviour.

Cooperativity can be achieved in multiple ways, for example by protein multimerisation, or by the presence of multiple binding sites. As the size of the system increases, and thus the minimum cooperativity, it seems that protein multimerisation quickly reaches a practical limit. This fact points to high-dimension multistable switches being more readily implemented in eukaryotes than in prokaryotes, since the eukaryotic mechanism of transcription can integrate the influence of many different *cis*-factors spread out over a lengthy sequence (Struhl, 1999), and seems to have a natural tendency to cooperativity (Carey, 1998); in eukaryotes, DNA-protein binding could also show cooperativity due to the presence of nucleosomes, and this cooperativity could thus be independent of specific protein-protein interactions (Polach & Widom, 1996). Sophisticated "*cis*-regulatory information processing" has been identified and modelled by Yuh *et al.* (2001), and a system in which different *cis*-regulatory elements were combined to drive additively the expression of a reporter has been engineered by Kirchhamer *et al.* (1996).

These remarks are of course consistent with the fact that sophisticated cellular memory, in the form of cellular differentiation, is a characteristic of eukaryotes.

# 7 Conclusion

We have shown that the existence of positive feedback is a necessary condition for multistationarity in a wide class of systems. This is an interesting property *per se*, but it is also a finding of practical importance. One can never be assured that the factors in a regulation network have all been identified, but it is also difficult to be certain that factors are missing. However, if one comes across a system exhibiting multistationarity, and the system has no positive feedback, then our result shows that the interaction graph is incomplete. For example, von Dassow *et al.* (2000) studied a system involved in cell differentiation, for which there was no experimental data asserting the

presence of a positive feedback circuit. The system displayed a very poor behaviour, which led the authors to assume two supplementary interactions in their model system. Most interestingly, these two assumed interactions each introduced a positive feedback circuit in the system, a fact which had not been explicitly pointed out by the authors.

Cellular differentiation is a process of major biological and medical interest. As illustrated by our multistable switch, mathematical modelling could prove to be useful in its investigation. Molecular information is becoming to be sufficiently available for the study of specific models; for example, dosed perturbation of the proteins involved in the transdifferentiation of pancreas to liver (Shen *et al.*, 2000) could allow to derive parameters for a possible switch.

# References


Aracena, J., Ben Lamine, S., Mermet, M.A., Cohen, O. & Demongeot, J. (2000) Mathematical modelling in genetic networks: relationships between the genetic expression and both chromosome breakage and positive circuits. In: *Bioinformatics and biomedical engineering* pp. 141-149. Washington: IEEE Computer Society.

Becskei, A., Séraphin, B., Serrano, L. (2001) Positive feedback in eukaryotic gene networks: cell differentiation by graded to binary response conversion. *EMBO J.* **20**, 2528-2535.

Berger, M. & Gostiaux B. (1992) Géométrie différentielle: variétés, courbes, et surfaces. Paris: PUF.

Bird, A.P. (1995) Gene number, noise reduction and biological complexity. *Trends Genet.* **11**, 94-100.

Bjornson, C.R., Rietze, R.L., Reynolds, B.A., Magli, M.C. & Vescovi, A.L. (1999) Turning brain into blood: a hematopoietic fate adopted by adult neural stem cells in vivo. *Science* **283**, 534-537.

Brockes, J.P. (1997) Amphibian limb regeneration: rebuilding a complex structure. *Science* **276**, 81-87.

Carey, M. (1998) The enhanceosome and transcriptional synergy. *Cell* **92**, 5-8.

Chelly, J., Hugnot, J.P., Concordet, J.P, Kaplan, J.C. & Kahn, A. (1991) Illegitimate (or ectopic) transcription proceeds through the usual promoters. *Biochem. Biophys. Res. Commun.* **178**, 553-7.

Cherry, J.L. & Adler, F.R. (2000) How to make a biological switch. *J. Theor. Biol.* **203**, 117-133.

Cinquin, O. & Demongeot, J. Positive and negative feedback: mending the ways of sloppy systems. To be published.

Clarke, D.L., Johansson, C.B., Wilbertz, J., Veress, B., Nilsson, E., Karlstrom, H., Lendahl, U. & Frisen, J. (2000) Generalized potential of adult neural stem cells. *Science* **288** 1660-1663.

von Dassow, G., Meir, E., Munro, E.M. & Odell, GM. (2000) The segment polarity network is a robust developmental module. *Nature* **406**, 188-92.

Delbrück, M. (1949) *Colloques internationaux CNRS.* In: Unités biologiques douées de continuité génétique. Paris: CNRS.

Demongeot, J., Thomas, R., & Thellier, M. (2000a) A mathematical model for storage and recall functions in plants. *C R Acad Sci III* **323**, 93-97.

Demongeot, J., Aracena, J., Ben Lamine, S., Mermet, M.A. & Cohen, O. (2000b) Hot spots in chromosomal breakage: from description to etiology. In: *Comparative genomics* (Sankoff, D. & Nadeau, J.H., eds) pp. 71-83. London: Kluwer.

Demongeot, J. (1998) Multistationarity and cell differentiation. *J. Biol. Sys.* **6**, 1-2.

Ferrari, G., Cusella-De Angelis, G., Coletta, M., Paolucci, E., Stornaiuolo, A., Cossu, G. & Mavilio, F. (1998) Muscle regeneration by bone marrow-derived myogenic progenitors. *Science* **279**, 1528-1530.

Gardner, T.S., Cantor, C.R., Collins, J.J. (2000) Construction of a genetic toggle switch in *Escherichia coli. Nature* **403**, 339-42.

Gouzé, J.L. (1998) Positive and negative circuits in dynamical systems. *J. Biol. Sys.* **6**, 11-15.

Jordan, J.D., Landau, E.M. & Iyengar, R. (2000) Signaling networks: the origins of cellular multitasking. *Cell* **103**, 193-200.

Kirchhamer, C.V., Bogarad, L.D. & Davidson, E.H. (1996) Developmental expression of synthetic cis-regulatory systems composed of spatial control elements from two different genes. *P.N.A.S.* **93**, 13849-13854.

Kling, U. & Székely, G. (1968) Simulation of rhythmic nervous activities. I: Function of networks with cyclic inhibitions. *Kybernetic* **5**, 89-103.

Koshland, D.E., Jr. (1998) The era of pathway quantification. *Science* **280**, 852-853.

Kringstein, A.M., Rossi, F.M., Hofman, A. & Blau, H.M. (1998) Graded transcriptional response to different concentrations of a single transactivator. *P.N.A.S.* **95**, 13670-13675.

Levins, R. (1975) Evolution in communities near equilibrium. In: *Ecology and evolution of communities* (Cody, M.L. & Diamond, J.M., eds) pp. 16-50. London: Belknap Press of Harvard University Press.

Monod, J. & Jacob, F. (1961) General conclusions: teleonomic mechanisms in cellular metabolism, growth, and differentiation. *Cold Spring Harbor Symp. Quant. Biol.* **26**, 389-401.

Ogbourne, S. & Antalis, T. M. (1998) Transcriptional control and the role of silencers in transcriptional regulation in eukaryotes. *Biochem. J.* **331**, 1-14.

Plahte, E., Mestl, T. & Omholt, W.S. (1995) Feedback circuits, stability and multistationarity in dynamical systems. *J. Biol. Sys.* **3**, 409-413.

Polach, K. J. & Widom, J. (1996) A model for the cooperative binding of eukaryotic regulatory proteins to nucleosomal target sites. *J. Mol. Biol.* **258**, 800-812.

Ptashne, M., (1986) A genetic switch, gene control and phage lamda. Cambridge, Massachusetts: Cell Press and Scientific Publications.

Savageau, M.A., Freter, R.R. (1979) On the evolution of accuracy and cost of proofreading tRNA aminoacylation. *P.N.A.S.* **76**, 4507-4510.

Shen, C.-N., Slack, J.M.W. & Tosh, D. (2000) Molecular basis of transdifferentiation of pancreas to liver. *Nat. Cell Biol.* **2**, 879-887.

Slack, J.M.W. & Tosh, D. (2001) Transdifferentiation and metaplasia - switching cell types. *Curr. Opin. Genet. Dev.* **11**, 581-586.

Snoussi, E.H. (1998) Necessary conditions for multistationarity and stable periodicity. *J. Biol. Sys.* **6**, 3-9.

Struhl, K. (1999) Fundamentally different logic of gene regulation in eukaryotes and prokaryotes. *Cell* **98**, 1-4.

Thomas, R. (1981) On the relation between the logical structure of systems and their ability to generate multiple steady states or sustained oscillations. *Springer Ser. Synergetics* **9**, 180-193.

Yuh, C.H., Bolouri, H., Davidson, E.H. (2001) *Cis*-regulatory logic in the *endo16* gene: switching from a specification to a differentiation mode of control. *Development* **128**, 617-629.

# Appendix

## Corollary 1 (section 4)

*Proof.* Let $P$ be the characteristic polynomial of $M$; by hypothesis, $P$ has a root $\alpha \in \Re_+^*$, which means that $\det(M-\alpha I) = 0$. Since $M$ is not degenerate, it has a decomposition in circuits. Let us suppose that $M$ has no positive circuit. $M-\alpha I$ possesses the same decompositions in circuits as $M$: diagonal terms of $M$ were negative and we substracted a positive number, so if they were not 0 in $M$ they are not 0 in $M-\alpha I$; other terms are untouched. If we

apply lemma 1 to matrix $M-\alpha I$, which is degenerate but has a decomposition in circuits, we find that $M-\alpha I$ has a positive circuit, and so does $M$. We thus arrive to a contradiction which concludes our proof. □

## Theorem 1 (section 5.1)

Theorem 1 makes use of an integral noted $I_F$, which is called the index of F on the unit sphere. $I_F$ is in fact an integer; in two dimensions, it is a measure of the number of times $F$ winds around the origin. $I_F$ is by no means trivial to compute, unless $F$ has specific, qualitative properties.

**Theorem 1.** *Following the widespread notations used in Berger & Gostiaux (1992), let us write* $I_F = \int_T (\frac{F}{\|F\|})^* \sigma$, *where* $T$ *is a hypersurface diffeomorphic to the unit sphere, such that the inside of* $T$ *is included in the domain of definition* $D$ *of* $F$ *and contains all ss of* $F$, *and where* $\sigma$ *is a volume form compatible with the canonical orientation of* $T$. *If* $F$ *has at least* $1+(-1)^n I_F$ stable *ss, then* $J_F$ *has a positive circuit.*

*Proof.* In the following, we denote ss of $F$ by $\{x_1, .., x_p\}$. We have the following result from degree theory (Berger & Gostiaux (1992), p 292):

$$\sum_{i=1}^{p} \text{index}_{x_i} F = I_F$$

For any non degenerate ss $x_i$, we have

$$\text{index}_{x_i} F = \text{sgn}(\det J_F(x_i))$$

We will denote the set of ss by $R_1 \cup R_2$, where

- $\forall\ r \in R_1,\ \text{sgn}(\det J_F(r)) = (-1)^n$
- $\forall\ r \in R_2,\ \text{sgn}(\det J_F(r)) = (-1)^{n+1}$

Stable ss are in $R_1$, because complex eigenvalues come in pairs, and if there are only complex eigenvalues then the dimension $n$ must be even. We have

$$\mid R_2 \mid (-1)^{n+1} + \mid R_1 \mid (-1)^n = I_F$$

$$\mid R_2 \mid = (-1)^{n+1} I_F + \mid R_1 \mid \tag{4}$$

Thus, if $\mid R_1 \mid > (-1)^n\ I_F$ (which derives directly from the last hypothesis in the theorem), one derives $\mid R_2 \mid > 0$, and $F$ has a ss $s$ verifying $\text{sgn}(\det J_F(s)) = (-1)^{n+1}$; according to lemma 1, $J_F(s)$ has a positive circuit. This proves the theorem. Note that the existence of a positive circuit

is an open condition : a positive circuit can disappear only if one of its elements vanishes; by taking the finite intersection of neighbourhoods in which each element doesn't vanish, one gets a neighbourhood of the ss in which a positive circuit is always present. □

This type of result can also be derived on compact manifolds, replacing the computation of the integral $I_F$ by the computation of the Euler characteristic, but it is not clear whether there are examples of dynamical systems on compact manifolds which are relevant to biological systems.

The application of the previous theorem involves knowing $I_F$. Whenever possible, it is probably easiest to derive $I_F$ from qualitative observations (for example, when $F$ is inward-pointing, the determination is direct, see below). For more difficult cases, we propose to compute an approximate value; since it is known from degree theory that $I_F$ is an integer, we can actually turn this approximate value into the exact value, provided we have an upper bound on the integration error, which is rather easy to find. Theorem 1 is interesting in that it replaces an extremely complicated computational problem, solving a multi-dimensional system of non-linear equations (to find ss of $F$) and computing eigenvalues of the Jacobian matrices (to find the existence of a real, positive eigenvalue), into a more straightforward problem of numerical analysis: the computation of an approximate integral. However, computing an approximate integral also becomes very difficult when the dimension rises, because of the exponential dependence of the complexity on the dimension of the problem.

If $F$ is inward-pointing, then $I_F = (-1)^n$, and Corollary 2 is thus an immediate consequence of Theorem 1: if $F$ is inward-pointing, and if $F$ has at least 2 stable ss, then $J_F$ has a positive circuit.

## Bending of $F$ into $G$ (section 5.2)

Let $C_s$ be the distance to the origin of the second most distant stable ss of $F$ (it is possible for more than 1 such ss to be situated at distance $C_s$ from the origin), let $C$ be a real number strictly superior to $C_s$, and $\epsilon > 0$. We will add a "pull-back" term to $F$ at points whose norm is greater than $C$. Since the function $x \mapsto < x, F(x) >$ is continuous, there is a constant $M \in \Re_+^*$ such that:

$$\forall\ x \in D \text{ st } \parallel x \parallel \leq \sqrt{n}(C+\epsilon),\ < x, F(x) > < M$$

In the following, we will use a smooth function $\Xi$ such that:

- $\forall\ x \in \Re \text{ st } x \leq C,\ \Xi(x) = 0$
- $\forall\ x \in \Re \text{ st } x \geq C + \epsilon,\ \Xi(x) = 1$

- $\forall\ x \in \Re\ \Xi'(x) \geq 0$

We will consider the new vector field $G$ defined by

$$\forall x \in D,\ \forall\ i \in I,\ G_i(x) = F_i(x) - \alpha \Xi(x_i) x_i,$$

where $\alpha \in \Re_+^*$ will be defined later on. Function $\Xi$ is defined in such a way that $G$ shares at least two stable ss with $F$.

Because of our assumption about expression leakage, we have:

$$\forall\ i \in I,\ \forall\ x \in D \text{ st } x_i = 0,\ F_i(x) > 0$$

This is also verified by $G$:

$$\forall\ i \in I,\ \forall\ x \in D \text{ st } x_i = 0,\ G_i(x) > 0$$

We have:

$$\forall\ x \in D,\ < x, G(x) > = < x, F(x) > - \alpha \sum_{i \in I} \Xi(x_i) x_i^2$$

Let us study the behaviour of the "pull-back" term we added to $F$:

$$\sum_{i \in I} x_i^2 = \sum_{i \in J_{C+\epsilon}(x)} x_i^2 + \sum_{i \in I \setminus J_{C+\epsilon}(x)} x_i^2,$$

where $J_c(x) = \{i \in I \text{ st } x_i < c\}$.

It is obvious that:

$$\sum_{i \in J_{C+\epsilon}(x)} x_i^2 < n(C+\epsilon)^2,$$

and by definition of $\Xi$ we have:

$$\forall\ x \in D,\ \sum_{i \in I} \Xi(x_i) x_i^2 \geq \sum_{i \in I \setminus J_{C+\epsilon}(x)} x_i^2$$

We thus conclude that

$$\forall\ t \text{ st } t > \sqrt{n}(C+\epsilon),\ \text{Min}_{\|x\|=t} \sum_{i \in I} \Xi(x_i) x_i^2 > 0$$

Therefore, we can choose $\alpha \in \Re_+^*$ such that:

$$\forall\ x \in D \text{ st } \|\ x\ \| = \sqrt{n}(C+\epsilon),\ \alpha \sum_{i \in I} \Xi(x_i) x_i^2 > M$$

Finally, we derive:

$$\forall\ x \in D \text{ st } \parallel x \parallel = \sqrt{n}(C+\epsilon), < x, G(x) >< 0$$

$G$ is thus inward-pointing on the sphere of centre 0 and radius $\sqrt{n}(C+\epsilon)$ intersected with $D$. $G$ has at least two stable ss inside this portion of sphere, because $G$ coincides with $F$ within the sphere of centre 0 and radius $C$ intersected with $D$. The area of space which we will consider is the intersection $\Delta$ of this portion of sphere with the cone $\{x \in D \text{ st } \forall\ i \in I,\ x_i > 0\}$. Because of the hypotheses we made, $G$ is inward-pointing on $\Delta$; the problem with $\Delta$ is that it is not smooth, but it is possible to "round the angles" into a surface $S$ close enough to $\Delta$ for $G$ to be inward-pointing on $S$ (the demonstration will be detailed elsewhere).

Finally we can apply corollary 2 to surface $S$: $G$ has a positive circuit at some point of $S$; since we added no positive circuits (the terms we added yield supplementary negative diagonal terms in the Jacobian matrix), $F$ has the same positive circuit. By letting $\epsilon$ be arbitrarily small, and $C$ arbitrarily close to $C_s$, we see that this positive circuit is at most at distance $\sqrt{n}C_s$ from the origin.

## Multistable switch (section 6)

### Existence of equilibria

Let us consider a ss of the system in which the variable $x_i$ has a given value, and all other variables share the same value $\epsilon$. We have

$$x_i = \frac{\sigma}{1+(n-1)\epsilon^c}$$
$$\epsilon = \frac{\sigma}{1+x_i^c+(n-2)\epsilon^c}$$

Rewriting the last equation as

$$\epsilon(1 + (\frac{\sigma}{1+(n-1)\epsilon^c})^c + (n-2)\epsilon^c) = \sigma,$$

and considering limits when $\epsilon$ tends to 0 and to infinity, one sees that the system has at least one solution.

### Bounds on $c$ and $\sigma$

Let us note $x_e$ verifying $F(x_e, .., x_e) = 0$. The set of such equilibria $x_e$ is defined by

$$x_e + (n-1)x_e^{c+1} = \sigma \tag{5}$$

$s$ being a strictly positive number, and considering limits when $x_e$ tends to 0 and to infinity, one sees that this equation has at least one solution. There is thus at least one undesirable equilibrium; this equilibrium should be unstable. We have

$$J_F(x_e,..,x_e) = \begin{pmatrix} -1 & a & \dots & a \\ a & -1 & \dots & a \\ \vdots & \vdots & \vdots & \vdots \\ a & a & \dots & -1 \end{pmatrix},$$

where

$$a = \frac{-\sigma c x_e^{c-1}}{(1+(n-1)x_e^c)^2}$$

The eigenvalues of $J_F(x_e,..,x_e)$ are $-(a+1)$ and $-1-(n-1)a$. Thus, a necessary and sufficient condition for the ss to be unstable is $-a > 1$, *i.e.*

$$\frac{\sigma c x_e^{c-1}}{(1+(n-1)x_e^c)^2} > 1$$

Using equation 5, we derive

$$c x_e^{c+1} > \sigma$$

Replacing $\sigma$ by its value derived from equation 5, we have

$$c x_e^{c+1} > x_e + (n-1)x_e^{c+1},$$
$$x_e^c(c-(n-1)) > 1 \ (x_e \neq 0),$$

from which we derive

$$c > n-1$$

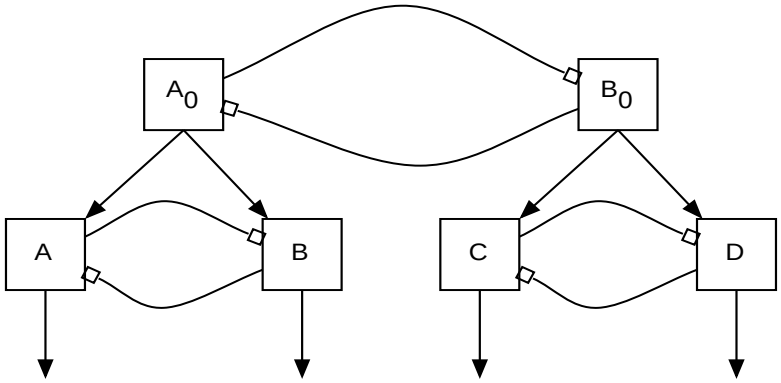
and

$$x_e^c > \frac{1}{c-(n-1)}$$

Combining this last equation with equation 5, we find

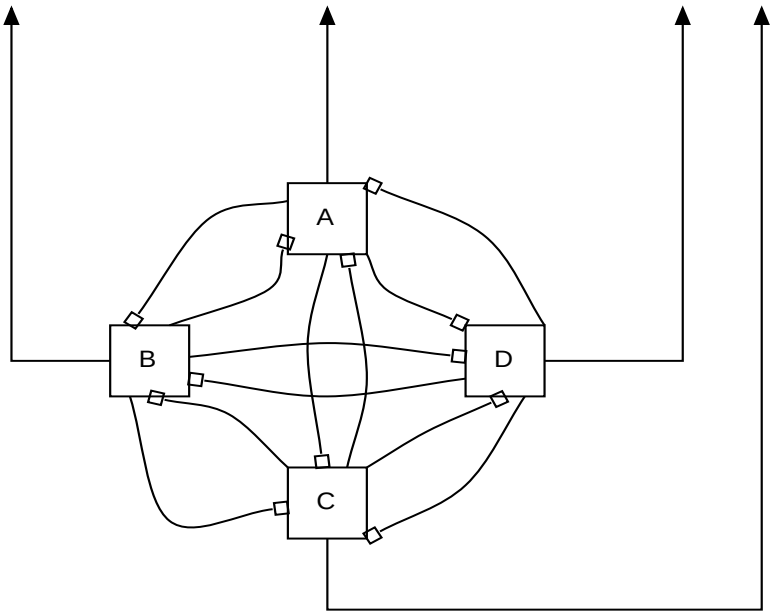
$$\sigma > (c-(n-1))^{-\frac{1}{c}} + (n-1)(c-(n-1))^{-\frac{c+1}{c}}$$

a) Hierarchic decision model




Genes specific for cell-type

b) Simultaneous decision model



Legend for Figure 1: Two basic models for cell differentiation. Empty squares correspond to repression, and arrows to activation. 1a: hierarchic decision model: genes $A_0$ and $B_0$, $A$ and $B$, and $C$ and $D$ form mutually repressive circuits, providing positive feedback. Differentiation would involve setting a first bistable switch, formed by $A_0$ and $B_0$, and then a second one, formed by either $A$ and $B$, or $C$ and $D$. 1b: simultaneous decision model: genes $A$, $B$, $C$ and $D$ form a 4-stable switch, and require mutual inhibition. Differentiation would involve setting the switch to the right state in one step.