On the link between oscillations and negative circuits in discrete genetic regulatory networks

Adrien Richard

INRIA Rhône-Alpes, 655 avenue de l'Europe, 38 334 Saint Ismier, France
adrien.richard@inria.fr

Abstract: The biologist René Thomas conjectured, twenty years ago, that the presence of a negative circuit in the interaction graph of a genetic regulatory network is a necessary condition for the presence of sustained oscillations in the dynamics of the network. Here, we state and prove this conjectured in a general discrete framework. The set of states of the network is assumed to be the Cartesian product X of n finite intervals of integers, and the dynamics of the networks is represented by an asynchronous state transition graph Γ on X, as in the Thomas' modeling. Then we derive from Γ an interactions graph G and we show that the presence, in Γ, of a strongly connected component A that we cannot leave and which is not reduced to a state, implies the presence of a negative circuit in G. This discrete version of the Thomas' conjecture was previously proved by Remy, Ruet and Thieffry under the strong hypothesis that A contains a unique cycle and under some assumptions on the form of Γ.

Keywords: Genetic regulatory network, negative circuit, oscillation, asynchronous state transition graph, Thomas' conjectures.

La structure des réseaux de régulation génétique est souvent représentée par un graphe d'interactions : les sommets correspondent aux gènes du réseau, et les arcs, qui sont soit positifs soit négatifs, indiquent les activations et les inhibitions présentes entre ces gènes. C'est généralement à partir de ce graphe que des modèles dynamiques (décrivant l'évolution temporelle du niveau d'expression des gènes) sont construits. Des paramètres précisant le fonctionnement des interactions sont alors introduits. La valeur des ces paramètres étant généralement inconnue et difficile à mesurer, il est intéressant de rechercher les propriétés dynamiques qui peuvent se déduire à la seule vue du graphe d'interactions (présentes quelle que soit la valeur des paramètres). Ayant cette perspective, le biologiste René Thomas a conjecturé, au début des années 80, que :

> *La présence d'un circuit négatif (contenant un nombre impaire d'inhibitions) dans le graphe d'interactions d'un réseau est une condition nécessaire pour la présence d'oscillations entretenues dans la dynamique de ce réseau* [1].

Il est à noter que ces oscillations sont capitales en biologie car elles permettent de modéliser les phénomènes d'homéostasie ; l'exemple type étant le cycle cellulaire.

Dans cette note, on énonce et on démontre la conjecture de Thomas dans un cadre discret général. On suppose que le niveau d'expression de chaque gène évolue dans un intervalle fini d'entiers. L'ensemble X des états possibles pour le réseau est alors le produit Cartésien de ces intervalles. On décrit ensuite la dynamique du réseau par un graphe de transitions d'états asynchrone Γ sur X, comme dans la *méthode logique généralisée* de Thomas [2]. Les oscillations entretenues sont alors décrites par les attracteurs cycliques de Γ, c'est à dire, les composantes fortement connexes de Γ qu'il est impossible de quitter et qui contiennent plus d'un état. On définit ensuite, à partir de Γ, un graphe d'interactions G nécessairement présent dans le graphe d'interactions du réseau. On montre enfin que :

Si Γ contient un attracteur cyclique alors G contient un circuit négatif (Théorème 1).

Cette version discrète de la conjecture de Thomas a récemment été démontrée par Remy, Ruet et Thieffry sous l'hypothèse forte que l'attracteur cyclique contient un unique cycle, et sous différentes hypothèses concernant Γ (dans [6] il est supposé que l'ensemble X des sommets de Γ est $\{0,1\}^n$, et dans [7] il est supposé que Γ peut être obtenu suivant la méthode de Thomas). Notons aussi que la conjecture de Thomas a été prouvée dans le cadre différentiel [4,5,3] ; cependant, les modèles discrets sont de nos jours de plus en plus utilisés du fait du manque de données précises sur le fonctionnement des régulations génétiques et du caractère non-linéaire de ces régulations.

1 Définitions

États. On s'intéresse donc à l'évolution d'un réseau regroupant n gènes en interactions, notés de 1 à n. On associe à chaque gène i un intervalle fini d'entiers X_i. Cet intervalle représente l'ensemble des *niveaux (d'expression)* possibles pour le gène i. L'ensemble des *états* possibles pour le réseau est alors le produit $X = X_1 \times \ldots \times X_n$ de ces intervalles.

Fonction signe. L'ensemble X des états du réseau étant fixé, on se donne ensuite une *fonction signe* sur X, c'est à dire, d'une application $d = (d_1, \ldots, d_n) : X \to \{-1, 0, 1\}^n$ telle que, pour tout $x \in X$, $x + d(x) \in X$; propriété qui se détaille en :

$$\forall x \in X,\ \forall i \in \{1, \ldots, n\}, \quad x_i + d_i(x) \in X_i.$$

Intuitivement, chaque "composante" d_i donne la direction dans laquelle le niveau de i évolue en fonction de l'état du réseau : si $d_i(x) = 1$ le niveau x_i de i est en "augmentation", si $d_i(x) = -1$ il est en "diminution", et si $d_i(x) = 0$ on considère qu'il est "stable". Posons :

$$\forall x \in X, \quad I_d(x) = \{i \in \{1, \ldots, n\} : d_i(x) \neq 0\}.$$

Suivant cette interprétation de d, $I_d(x)$ correspond à l'ensemble des gènes dont le niveau est en "augmentation" ou en "diminution" à l'état x.

Graphe de transitions asynchrone. Pour décrire, à partir de d, la dynamique du réseau, c'est à dire les changements d'états, ou *transitions*, qui peuvent avoir lieu au cours du temps, on prend ensuite en compte la remarque suivante : l'augmentation ou la diminution du niveau d'expression d'un gène passe par des phénomènes de synthèse et de dégradation qui nécessitent des délais très variables. L'évolution simultanée de plusieurs niveaux d'expression est donc irréaliste puisqu'elle implique des égalités improbables entre ces délais. C'est pourquoi, dans la méthode de Thomas [2], la dynamique est décrite de façon totalement *asynchrone* : "durant" une transition d'un état x à un état y, il y a au plus un gène $i \in I_d(x)$ dont le niveau x_i évolue, et ce niveau augmente (diminue) d'une unité si $d_i(x)$ est positif (négatif). Plus précisément, dans le modèle de Thomas, la dynamique est décrite par le *graphe de transitions asynchrone de* d, noté Γ_d : l'ensemble de ses sommets est X, et l'ensemble E_d de ses arcs est défini par :

$$E_d = \{(x, x + d_i(x)e_i) : x \in X,\ i \in I_d(x)\},$$

où, pour chaque gène i, e_i désigne le n-uplet dont toutes les composantes sont 0 sauf la $i$ème qui est égale à 1 (de sorte que $d_i(x)e_i = (0, \dots, d_i(x), \dots, 0)$). Notons que $|I_d(x)|$ est ainsi le nombre de successeurs de x dans Γ_d.

Attracteurs cycliques. On introduit maintenant les "structures" de Γ_d qui nous intéresserons dans la suite. Un *chemin de* Γ_d de longueur $r > 0$ est une séquence $(x^0, \dots, x^r)$ telle que $(x^p, x^{p+1}) \in E_d$ pour tout $0 \leq p < r$. On dit d'un tel chemin qu'il va de x^0 à x^r. Un *domaine clos de* Γ_d est un ensemble non vide $A \subseteq X$ tel que, pour tout $(x, y) \in E_d$, si $x \in A$ alors $y \in A$. Un domaine clos est donc, comme son nom l'indique, un ensemble d'états qu'il est impossible de quitter. Un *attracteur de* Γ_d est un plus petit domaine clos au sens de l'inclusion. Un attracteur est *cyclique* s'il est de cardinalité ≥ 2. Un *cycle stable* est un attracteur cyclique A tel que $|I_d(x)| = 1$ pour tout $x \in A$. Un *état stable* est un état x sans successeur dans Γ_d, ou, de façon équivalente, un état x tel que $\{x\}$ est un attracteur.

REMARQUE 1. Il est facile de voir qu'un attracteur cyclique A de Γ_d est une composante fortement connexe de Γ_d : pour tout $x, y \in A$ il existe un chemin de x à y (ne passant que par des états de A). On en déduit que pour tout $x \in A$, $I_d(x)$ est non vide (car x a au moins un successeur dans Γ_d). On en déduit également que si $x \in A$ et $i \in I_d(x)$, alors il existe $y \in A$ tel que $d_i(x) = -d_i(y)$. Pour s'en convaincre, supposons par exemple que $d_i(x) = 1$. Alors, $(x, x + e_i)$ est une transition de E_d durant laquelle le niveau de i augmente. Comme A est un attracteur cyclique, $x + e_i \in A$ et il existe un chemin de $x + e_i$ à x. Pour revenir en x, ce chemin, qui ne passe que par des états de A, empreinte nécessairement une transition de la forme $(y, y - e_i)$ durant laquelle le niveau de i diminue. On a alors $y \in A$ et $d_i(y) = -1$.

En somme, lorsque l'état du système est dans un attracteur cyclique, il parcours sans cesse les états de l'attracteur en empruntant des cycles. Les attracteurs cycliques permettent ainsi de modéliser les oscillations entretenues que l'on rencontre, par exemple, dans le cyclique cellulaire et, plus généralement, dans les phénomènes d'homéostasie.

graphe d'interactions. Dans ce paragraphe, on associe à la fonction signe d un graphe d'interactions. Il ne faut pas voir celui-ci comme *le* graphe d'interactions du réseau dont la dynamique est modélisée par Γ_d, mais comme un graphe d'interactions nécessairement présent dans le graphe d'interactions de ce réseau.

Formellement, un *graphe d'interactions* est un graphe orienté $G = (V, E)$ où chaque arc est étiqueté par un signe, c'est à dire, où $E \subseteq V \times \{-1, 1\} \times V$. Un *chemin de* G de longueur $r > 0$ est une séquence $(i_0, s_0, i_1, s_1, \dots, i_{r-1}, s_{r-1}, i_r)$ telle que $(i_p, s_p, i_{p+1}) \in E$ pour tout

$0 \leq p < r$. On dit d'un tel chemin qu'il va de i_0 à i_r, qu'il est un *circuit* si $i_0 = i_r$, et qu'il est *positif* (*négatif*) si son *signe* $\prod_{p=0}^{r-1} s_p$ est positif (négatif). Un graphe d'interactions $G' = (V', E')$ est un *sous-graphe* de G si $V' \subseteq V$ et $E' \subseteq E$.

Soit d une fonction signe sur X. Si $A \subseteq X$, on appelle *graphe d'interactions de d évalué sur A*, et on note $G_d(A)$, le graphe d'interactions dont l'ensemble des sommets est $\{1, \dots, n\}$, et qui contient un arc (i, s, j) s'il existe $x \in A$ tel que :

$$s = d_i(x)d_j(x + d_i(x)e_i) \quad \text{et} \quad d_j(x) \neq d_j(x + d_i(x)e_i).$$

Bien sûr, si $A' \subseteq A \subseteq X$ alors $G_d(A')$ est un sous-graphe de $G_d(A)$. Dans la suite, on notera simplement G_d le graphe d'interactions évalué sur l'ensemble X des états du réseau.

Pour comprendre en quoi G_d est nécessairement présent dans le graphe d'interactions du réseau modélisé par Γ_d, considérons un état $x \in X$ pour lequel il existe un gène i et un gène j tels que $s = d_i(x)d_j(x + d_i(x)e_i) \neq 0$ et $d_j(x) \neq d_j(x + d_i(x)e_i)$, et supposons que $d_i(x) = 1$, la situation étant symétrique dans le cas contraire. Il existe alors une transition de x à $x + e_i$, et si $d_j(x + e_i)$ est positif (négatif) alors $d_j(x)$ ne l'est pas : l'augmentation (la diminution) de j, impossible à l'état x, devient possible suite à l'augmentation de i, c'est à dire, à l'état $x + e_i$. Il est alors claire que i est un activateur (inhibiteur) de j, où encore, qu'il existe une interaction de i à j de signe $s = d_j(x + e_i)$ dans le graphe d'interactions du réseau.

2 Attracteurs cycliques et circuits négatifs

Dans cette section, on démontre que la présence d'un attracteur cyclique A dans Γ_d implique la présence d'un circuit négatif dans $G_d(A)$ (Théorème 1). Dans le cas très particulier où A correspond à un cycle stable, et sous différentes hypothèses concernant Γ_d, cette version discrète de la conjecture de Thomas a été démontrée par Remy, Ruet et Thieffry dans [6,7].

LEMME 1. *Soient d un fonction signe sur X et un chemin $(x^0, \dots, x^r)$ de Γ_d de longueur $r > 0$. S'il existe $j \in I_d(x^r)$ tel que $d_j(x^p) \neq d_j(x^r)$ pour tout $0 \leq p < r$, alors il existe $i \in I_d(x^0)$ tel que $G_d(\{x^0, \dots, x^r\})$ contient un chemin de i à j de signe $d_i(x^0)d_j(x^r)$.*

Preuve. On se place sous les conditions du lemme, et on raisonne par récurrence sur la longueur r du chemin. Afin d'alléger les notations, on pose $A = \{x^0, \dots, x^r\}$.

Cas $r = 1$: Comme $(x^0, x^1) \in E_d$, il existe $i \in I_d(x^0)$ tel que $x^1 = x^0 + d_i(x^0)e_i$. Or, sous les conditions du lemme, $j \in I_d(x^1)$ et $d_j(x^0) \neq d_j(x^1)$. On en déduit que $G_d(A)$ contient un chemin (de longueur 1) de i à j de signe $d_i(x^0)d_j(x^1)$.

Cas $r > 1$: Considérons le chemin (x^{r-1}, x^r) de Γ_d longueur 1. En utilisant le raisonnement précédant, on montre qu'il existe $k \in I_d(x^{r-1})$ tel que $G_d(A)$ contient un chemin de k à j de signe $d_k(x^{r-1})d_j(x^r)$. Si $d_k(x^0) = d_k(x^{r-1})$ il n'y a donc plus rien à démontrer. Sinon, il existe un plus petit indice $0 < p < r$ tel que $d_k(x^p) = d_k(x^{r-1})$. Du fait de la minimalité de cet indice, on a $d_k(x^q) \neq d_k(x^p)$ pour tout $0 \leq q < p$. Comme $0 < p < r$, on déduit de l'hypothèse d'induction qu'il existe $i \in I_d(x^0)$ tel que $G_d(A)$ contient un chemin de i à k de signe $d_i(x^0)d_k(x^p)$. Comme nous avons vu que $G_d(A)$ contient un chemin de k à j de signe $d_k(x^{r-1})d_j(x^r)$, on en déduit que $G_d(A)$ contient un chemin de i à j (passant par k) de signe $s = d_i(x^0)d_k(x^p)d_k(x^{r-1})d_j(x^r)$. Or $d_k(x^p) = d_k(x^{r-1})$ donc $s = d_i(x^0)d_j(x^r)$. □

LEMME 2. *Soient d une fonction signe sur X et un attracteur cyclique A de Γ_d. S'il existe $x \in A$ tel que $|I_d(x)| = 1$ alors $G_d(A)$ contient un circuit négatif.*

Preuve. On se place sous les conditions du lemme, et on note i l'unique élément de $I_d(x)$. Comme A est un attracteur cyclique, l'ensemble Y des états $y \in A$ tels que $d_i(y) = -d_i(x)$ est non vide (Remarque 1). Il existe donc un chemin de longueur minimal $(x^0, \dots, x^r)$ tel que $x = x^0$ et $x^r \in Y$, et celui-ci ne passe que par des états de A. Du fait de la minimalité de la longueur de ce chemin, on a $d_i(x^p) \neq d_i(x^r)$ pour tout $0 \leq p < r$. Donc, d'après le lemme 1, il existe $j \in I_d(x^0)$ tel que $G_d(\{x^0, \dots, x^r\})$ contient un chemin de j à i de signe $d_j(x^0)d_i(x^r)$. Or, comme i est l'unique élément de $I_d(x^0)$, on a $j = i$. On en déduit que $G_d(\{x^0, \dots, x^r\})$ contient un chemin de i à i, et donc un circuit, de signe $d_i(x^0)d_i(x^r)$. Comme $d_i(x^0) \neq d_i(x^r)$, ce circuit est négatif, et comme $\{x^0, \dots, x^r\} \subseteq A$, il est aussi présent dans $G_d(A)$. □

LEMME 3. *Soient d une fonction signe sur X et A un attracteur cyclique de Γ_d. Si, pour tout $x \in X$, $|I_d(x)| \geq 2$, alors il existe une fonction signe d' sur X telle que $E_{d'} \subset E_d$ et telle que $\Gamma_{d'}$ contient un attracteur cyclique A' tel que $G_{d'}(A')$ est un sous-graphe de $G_d(A)$.*

Preuve. On se place sous les conditions du lemme. Soient $z \in A$, $k \in I_d(z)$ et l'ensemble $D \subseteq A$ défini par : $D = \{x \in A : d_k(x) = d_k(z)\}$. Bien sûr, D est non vide puisque $z \in D$. Soit alors la fonction signe d' sur X définie par :

$$\forall x \in X,\ \forall i \in \{1, \dots, n\}, \qquad d'_i(x) = \begin{cases} 0 & \text{si } x \in D \text{ et } i = k, \\ d_i(x) & \text{sinon.} \end{cases}$$

Il est clair que :

$$\forall x \in X,\ \forall i \in \{1, \dots, n\}, \qquad d'_i(x) \neq 0 \Rightarrow d'_i(x) = d_i(x). \tag{1}$$

Il vient de là que $E_{d'} \subseteq E_d$, et comme pour tout $x \in D$ on a $k \in I_d(x) \setminus I_{d'}(x)$, l'inclusion est stricte. On déduit de $E_{d'} \subseteq E_d$ que A est un domaine clos de $\Gamma_{d'}$. Par conséquent $\Gamma_{d'}$ contient (au moins) un attracteur $A' \subseteq A$. Montrons que celui-ci est cyclique. Si x est l'unique élément de A', alors $I_{d'}(x)$ est vide. Comme $x \in A$, par hypothèse, $|I_d(x)| \geq 2$. On en déduit qu'il existe $i \in I_d(x) \setminus \{k\}$. Mais alors $d'_i(x) = d_i(x) \neq 0$, donc $i \in I_{d'}(x)$, une contradiction. Ainsi, A' contient au moins deux éléments : c'est un attracteur cyclique de $\Gamma_{d'}$. Montrons finalement que $G_{d'}(A')$ est un sous-graphe de $G_d(A)$. Supposons que (i, s, j) soit une interaction de $G_{d'}(A')$. Alors, il existe $x \in A'$ tel que :

$$s = d'_i(x)d'_j(x + d'_i(x)e_i) \neq 0 \quad \text{et} \quad d'_j(x) \neq d'_j(x + d'_i(x)e_i).$$

Donc $d'_i(x) \neq 0$ et $d'_j(x + d'_i(x)e_i) \neq 0$, d'où, d'après (1) :

$$d_i(x) = d'_i(x) \quad \text{et} \quad d_j(x + d_i(x)e_i) = d_j(x + d'_i(x)e_i) = d'_j(x + d'_i(x)e_i).$$

D'après ceci et le fait que $x \in A$, pour montrer que (i, s, j) est une interaction de $G_d(A)$, il suffit de montrer que $d_j(x) \neq d_j(x + d_i(x)e_i)$. Supposons par contradiction que :

$$d_j(x) = d_j(x + d_i(x)e_i).$$

Alors, $d_j(x) \neq d'_j(x)$ d'où $j = k$ et $x \in D$. Par conséquent $d_k(z) = d_k(x) = d_k(x + d_i(x)e_i)$. Comme $x \in A$ et $(x, x+d_i(x)e_i) \in E_d$ on a $x+d_i(x)e_i \in A$ et on en déduit que $x+d_i(x)e_i \in D$. Mais alors, d'après la définition de d', on a $d'_j(x + d'_i(x)e_i) = d'_k(x + d_i(x)e_i) = 0$, une contradiction. On en déduit que (i, s, j) est une interaction de $G_d(A)$, d'où le lemme. □

THÉORÈME 1. *Soit d une fonction signe sur X. Si A est un attracteur cyclique de Γ_d alors $G_d(A)$ contient un circuit négatif.*

Preuve. Soit U l'ensemble des couples (d, A) tels que d est une fonction signe sur X et tels que A est un attracteur cyclique de Γ_d. Soit $\prec$ la relation sur U définie par $(d', A') \prec (d, A)$ si et seulement si $E_{d'} \subset E_d$. La relation $\prec$ étant un ordre stricte bien fondée, on peut montrer, en raisonnant par induction, que tout élément $(d, A) \in U$ est tel que $G_d(A)$ admet un circuit négatif.

Soit (d, A) un élément minimal de $(U, \prec)$. Si pour tout $x \in A$ on a $|I_d(x)| \geq 2$ alors, d'après le lemme 3, il existe $(d', A') \in U$ tel que $(d', A') \prec (d, A)$ ce qui est faux puisque (d, A) est supposé minimal. On en déduit qu'il existe $x \in A$ tel que $|I_d(x)| = 1$. D'après le lemme 2, il existe alors un circuit négatif dans $G_d(A)$.

Soit (d, A) un élément non minimal de $(U, \prec)$. Supposons que pour tout $(d', A') \prec (d, A)$, $G_{d'}(A')$ contient un circuit négatif. Si, pour tout $x \in A$, on a $|I_d(x)| \geq 2$, alors, d'après le lemme 3, il existe $(d', A') \prec (d, A)$ tel que $G_{d'}(A')$ est un sous-graphe de $G_d(A)$. Par hypothèse d'induction, $G_{d'}(A')$ contient un circuit négatif, et comme $G_{d'}(A')$ est un sous-graphe de $G_d(A)$, ce circuit est aussi présent dans $G_d(A)$. Sinon, il existe $x \in A$ tel que $|I_d(x)| = 1$, et alors, d'après le lemme 2, $G_d(A)$ contient de nouveau un circuit négatif. □

On déduit de ce théorème et du fait que Γ_d contient forcement au moins un attracteur le :

COROLLAIRE 1. *Si G_d est sans circuit négatif, alors Γ_d contient un état stable.*

Références

[1] R. Thomas. On the relation between the logical structure of systems and their ability to generate multiple steady states and sustained oscillations. In *Series in Synergetics*, volume 9, pages 180-193. Springer, 1981.

[2] R. Thomas and M. Kaufman. Multistationarity, the basis of cell differentiation and memory. I. & II. *Chaos*, 11 :170-195, 2001.

[3] E.H. Snoussi. Necessary conditions for multistationarity and stable periodicity. *Journal of Biological Systems*, 6 :3-9, 1998.

[4] E. Plathe, T. Mestl and S.W. Omholt. Feedback loops, stability and multistationarity in dynamical systems. *Journal of Biological Systems*, 3 :569-577, 1995.

[5] J.L. Gouzé. Positive and negative circuits in dynamical systems. *Journal of Biological Systems*, 6 :11-15, 1998.

[6] E. Remy, P. Ruet and D. Thieffry. Graphics Requirement for Multistability and Attractive Cycles in a Boolean Dynamical Framework. Technical Report, IML, 2005.

[7] E. Remy, P. Ruet and D. Thieffry. Positive or negative regulatory circuit inference from multilevel dynamics. In *Positive Systems : Theory and Applications*, volume 341 of LNCIS, pages 263-270, Springer, 2006.