

MODÉLISATION DE RÉSEAUX DE RÉGULATION GÉNÉTIQUE INTERCELLULAIRES

Anne Crumière

Résumé. — Les biologistes représentent souvent les interactions génétiques par des graphes orientés, appelés graphes de régulation génétique. Les sommets désignent les gènes du système et les arêtes, les effets de régulation d'un gène sur un autre. Elles sont munies d'un signe positif dans le cas d'une activation et négatif pour une inhibition. Ce papier traite des relations entre la structure de ces graphes et leurs propriétés dynamiques.

1. Introduction

Les biologistes représentent souvent les réseaux de régulation génétique par des graphes. Ces graphes, appelés graphes d'interaction, sont des graphes orientés et signés, notés $G = (V, E)$, où l'ensemble des nœuds, $V = \{1, \dots, n\}$, représente les n gènes du système et l'ensemble des arêtes E représente les régulations : $(i, j) \in E$ si le gène i est un régulateur du gène j . Les arêtes sont signées, positivement ($+1$) dans le cas d'une activation, c'est à dire quand la protéine codée par le gène i favorise l'expression du gène j , et négativement dans le cas d'une inhibition (-1), c'est à dire quand la protéine codée par le gène i ralentit ou stoppe l'expression du gène j .

Un tel graphe d'interaction est une représentation d'un réseau permettant d'expliquer au niveau de la régulation génétique, un processus biologique donné. Les niveaux d'expression des n gènes à un instant donné représentent un *état* du système.

Dans ce papier, on s'intéresse aux propriétés de stabilité de ce système ; dans le cadre des systèmes dynamiques, cela se traduit par les notions d'attracteurs, d'états stationnaires, de cycles attractifs... L'étude de ces propriétés demande soit dans un cadre continu la résolution d'un système d'équations différentielles assez grand, soit dans le cadre discret, l'étude d'un graphe de grande taille représentant la dynamique. Dans les deux cas, on est confronté à des problèmes de très grande complexité.

D’où l’idée de revenir aux graphes d’interaction qui correspondent à ces dynamiques, et qui ont l’avantage d’être de taille plus petite et donc plus facile à analyser. De tels liens entre la dynamique et les graphes d’interaction existent : les *règles de Thomas* en sont un bel exemple. En effet, dans les années 80, le biologiste René Thomas énonça la règle suivante [5] : *une condition nécessaire pour l’existence de plusieurs états stationnaires dans la dynamique est la présence d’un circuit positif dans le graphe d’interaction (le signe d’un circuit étant le produit des signes des arêtes).*

Durant cette décennie, de nombreux chercheurs ont formalisé et démontré cette règle dans des cadres différents [2, 3, 4], mais toujours dans le cas où les gènes sont répartis dans une même cellule. Ce travail étend dans un cadre discret ces règles à un réseau génétique intercellulaire.

2. Formalisme

2.1. Dynamique du réseau de régulation. — On s’intéresse à l’évolution d’un système composé de cellules, chaque cellule contient la même collection de gènes choisis dans un ensemble fini I . Pour un gène $i \in I$, l’intervalle $\mathcal{A}_i = [0, k_i]$ désigne les niveaux d’expression possibles du gène i . Un *état* d’une cellule est un élément du produit cartésien $\mathcal{A} = \prod_{i \in I} \mathcal{A}_i$.

Généralement, un système biologique est constitué de plusieurs cellules. On peut supposer que les cellules sont réparties régulièrement et disposées suivant un réseau \mathbb{M} , i.e. un sous-groupe discret de \mathbb{R}^d muni de l’opération $+$. Chaque cellule est dans un état $\alpha \in \mathcal{A}$. Un *état du système* est ainsi une suite d’éléments de \mathcal{A} indexée par \mathbb{M} , i.e. un élément de $\mathcal{A}^{\mathbb{M}}$. Pour tous $s \in \mathcal{A}^{\mathbb{M}}$ et $\mathbb{U} \subset \mathbb{M}$, on note $s_{\mathbb{U}}$ la *restriction* de s à \mathbb{U} . Lorsque l’état du système est donné par $s \in \mathcal{A}^{\mathbb{M}}$, l’état de la cellule $x \in \mathbb{M}$ est noté $s(x)$ et pour un gène $i \in I$, le niveau d’expression du gène i dans la cellule x est noté $s(x, i)$.

Le niveau d’expression d’un gène dans une cellule varie au cours du temps en fonction des niveaux d’expression des gènes dans cette cellule et dans les cellules voisines. Chaque cellule communique avec ses voisines de manière uniforme dans l’espace. Pour modéliser ce phénomène, on considère l’ensemble fini $\mathbb{V} \subset \mathbb{M}$ appelé *voisinage* et une *fonction locale* $f : \mathcal{A}^{\mathbb{V}} \rightarrow \mathcal{A}$. On suppose que l’élément nul de \mathbb{M} appartient à \mathbb{V} .

La *dynamique globale* du système peut alors être donnée par l’*automate cellulaire* $F : \mathcal{A}^{\mathbb{M}} \rightarrow \mathcal{A}^{\mathbb{M}}$ défini par $F(s)(x) = f((s(x + v))_{v \in \mathbb{V}})$ pour tous $s \in \mathcal{A}^{\mathbb{M}}$ et $x \in \mathbb{M}$. On désigne par $F(s)(x, i)$ la valeur vers laquelle le niveau d’expression du gène i dans la cellule x tend quand le système est à l’état s . Tout naturellement, un *état stationnaire* de F est un état $s \in \mathcal{A}^{\mathbb{M}}$ tel que $F(s) = s$, i.e. un état dans lequel le niveau d’expression de chaque gène n’évolue pas.

2.2. Graphe d'interaction. — Généralement, on observe les variations du niveau d'expression d'un gène lorsque les autres gènes interagissent avec celui-ci. Ces variations sont mises en évidence par le calcul de la Jacobienne discrète et visualisées sous la forme d'un graphe d'interactions.

Notation 2.1. — • Soit $t \in \mathbb{Z}$, on définit $\text{sg}(t) = 0$ si $t = 0$, $\text{sg}(t) = +1$ si $t > 0$ et $\text{sg}(t) = -1$ si $t < 0$.

• Soient $(s, s') \in (\mathcal{A}^{\mathbb{M}})^2$ et $(x, i) \in \mathbb{M} \times I$, on définit $s^{(x,i) \triangleleft s'}$ par : pour tout $(y, j) \in \mathbb{M} \times I$,

$$s^{(x,i) \triangleleft s'}(y, j) = \begin{cases} s(y, j) & \text{si } (x, i) \neq (y, j), \\ s(y, j) + \text{sg}(s'(y, j) - s(y, j)) & \text{sinon.} \end{cases}$$

Définition 2.2. — Soient $s, s' \in \mathcal{A}^{\mathbb{M}}$ et $((x, i), (y, j)) \in (\mathbb{M} \times I)^2$, on définit la **Jacobienne discrète** de F en $s \in \mathcal{A}^{\mathbb{M}}$ suivant la direction $s' \in \mathcal{A}^{\mathbb{M}}$ comme étant la matrice dont les coefficients sont :

$$\partial_{(x,i),(y,j)} F(s, s') = \text{sg}(s'(x, i) - s(x, i)) \text{sg}(F(s^{(x,i) \triangleleft s'})(y, j) - F(s)(y, j)).$$

Pour visualiser ou plutôt représenter les différentes actions d'un gène sur un autre dans une même cellule ou dans une cellule voisine dans une région $\mathbb{U} \subset \mathbb{M}$, on définit le **graphe d'interaction** de $\partial F(s_{\mathbb{U}}, s'_{\mathbb{U}})$, noté $G(\partial F(s_{\mathbb{U}}, s'_{\mathbb{U}}))$. Le graphe d'interaction est un graphe orienté signé, *i.e.*, avec un signe $+1$ ou -1 , attaché à chaque arête et défini par :

- les sommets sont les gènes de chaque cellule contenue dans \mathbb{U} , c'est à dire $\mathbb{U} \times I$,
- il y a une arête allant du gène i dans la cellule x vers le gène j dans la cellule y si $\partial_{(x,i),(y,j)} F(s_{\mathbb{U}}, s'_{\mathbb{U}}) \neq 0$ et $\text{sg}(F(s)(y, j) - s(y, j)) \neq \text{sg}(F(s^{(x,i) \triangleleft s'})(y, j) - s(y, j))$. Le signe de l'arête est alors celui de $\partial_{(x,i),(y,j)} F(s_{\mathbb{U}}, s'_{\mathbb{U}})$, il est positif (resp. négatif) si le niveau d'expression du gène i est en "augmentation" (resp. "diminution").

3. Conditions nécessaires de multistabilité

Définition 3.1. — On rappelle que \mathbb{V} est le voisinage de la fonction globale F . Soient $\mathbb{U}', \mathbb{U}''$ deux sous-ensembles de \mathbb{M} , on définit le sous-ensemble $\mathbb{U}' + \mathbb{U}'' = \{u' + u'' : u' \in \mathbb{U}' \text{ et } u'' \in \mathbb{U}''\}$. Pour $\mathbb{U} \subset \mathbb{M}$ fini, on note $\partial\mathbb{U} = (\mathbb{U} + \mathbb{V}) \setminus \mathbb{U}$ le **bord** de \mathbb{U} .

Notre formalisme pour décrire les interactions génétiques dans un cadre intercellulaire est maintenant mis en place. Nous pouvons énoncer l'adaptation de la règle de Thomas au système intercellulaire avec la même collection de gènes dans chaque cellule :

Théorème 3.2. — [1] Soient $F : \mathcal{A}^{\mathbb{M}} \rightarrow \mathcal{A}^{\mathbb{M}}$ une fonction globale et $\mathbb{U} \subset \mathbb{M}$. Si $r, t \in \mathcal{A}^{\mathbb{M}}$ sont deux états stationnaires qui vérifient $(r)_{\partial\mathbb{U}} = (t)_{\partial\mathbb{U}}$ et $(r)_{\mathbb{U}} \neq (t)_{\mathbb{U}}$, alors il existe $s \in \mathcal{A}^{\mathbb{M}}$ tel que $G(\partial F(s, t))$ a un circuit positif élémentaire.

Remarque 3.3. — Comme $G(\partial F(s, t))$ contient un circuit élémentaire positif mais comme $s_{\partial U} = t_{\partial U}$, le circuit positif est localisé sur U .

4. Perspectives

Dans le Théorème 3.2, la localisation du circuit est induite par les états stationnaires, donc dépendante de la dynamique du système. Mais dans la plupart des modèles étudiés par les biologistes, les circuits sont localisés sur au plus deux cellules. Il serait intéressant d'étudier le nombre maximum de cellules sur lesquelles le circuit positif s'étend.

De plus, ce modèle positionne les cellules sur un réseau, donc régulièrement dans l'espace. Une perspective serait de les répartir sur une structure moins ordonnée mais suffisamment régulière pour définir une dynamique locale, par exemple le pavage de Penrose.

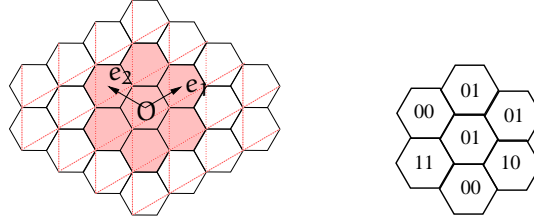
Références

- [1] A. Crumière, *Circuits de rétroaction dans les réseaux génétiques de régulation intercellulaires*, PhD Thesis, 2008.
- [2] C. Soulé, *Graphic requirements for multistationarity*, *ComplexUs*, 1 :123–133, 2003.
- [3] É. Remy, P. Ruet, and D. Thieffry, *Graphic requirements for multistability and attractive cycles in a Boolean dynamical framework*. *Advances in Applied Mathematics*, 41(3) : 335-350, 2008.
- [4] A. Richard, *Modèle formel pour les réseaux de régulation génétique et influence des circuits de rétroaction*, PhD Thesis, 2006.
- [5] R. Thomas, *On the relation between the logical structure of systems and their ability to generate multiple steady states and sustained oscillations*, *Series in Synergetics*, 9 :180-193. Springer, 1981.

ANNEXE : Un exemple de réseau hexagonal

On étudie un système constitué de deux gènes par cellule, *i.e.*, $I = \{a, b\}$, qui ont deux niveaux d'expression, ainsi $\mathcal{A} = \{0, 1\} \times \{0, 1\}$. Les cellules sont localisées sur un plan. Pour des raisons biologiques, les cellules sont représentées par des hexagones qui pavent le plan. Ce *pavage hexagonal* \mathbb{M} est généré par les vecteurs \mathbf{e}_1 et \mathbf{e}_2 , plus précisément $\mathbb{M} = \mathbb{Z}\mathbf{e}_1 + \mathbb{Z}\mathbf{e}_2$. Soit $(x_1, x_2) \in \mathbb{Z}^2$, (x_1, x_2) sont les coordonnées de la cellule $x \in \mathbb{M}$ dans le pavage selon une origine arbitraire O et la base $(\mathbf{e}_1, \mathbf{e}_2)$ (voire Figure 1).

Le voisinage d'une cellule quelconque x de coordonnées (x_1, x_2) est donc : $x + \mathbb{V} = \{(x_1, x_2), (x_1, x_2 + 1), (x_1 + 1, x_2 + 1), (x_1 - 1, x_2), (x_1 + 1, x_2), (x_1 - 1, x_2 - 1), (x_1, x_2 - 1)\}$. L'état local $s_{\mathbb{V}}$, représenté en Figure 1, est composé de 7 cellules hexagonales où les deux nombres dans chaque cellule sont les niveaux d'expression des deux gènes contenus dans chaque cellule. Cet état est mathématiquement représenté par


 FIGURE 1. (Gauche) Localisation (Droite) x_V

la matrice suivante : $\begin{pmatrix} (0,0) & (0,1) \\ (1,1) & (0,1) & (0,1) \\ (0,0) & (1,0) \end{pmatrix}$. Plus généralement pour tout état $s \in$

\mathcal{A}^M , l'état $s(x + \nabla)$ est une matrice 3×3 avec deux trous :

$$s(x + \nabla) = \begin{pmatrix} & s(x_1, x_2 + 1) & s(x_1 + 1, x_2 + 1) \\ s(x_1 - 1, x_2) & s(x_1, x_2) & s(x_1 + 1, x_2) \\ s(x_1 - 1, x_2 - 1) & s(x_1, x_2 - 1) & \end{pmatrix}$$

Par la suite, pour des raisons d'espace, nous supposons qu'il y a un seul gène dans chaque cellule, *i.e.* $I = \{a\}$ avec toujours deux niveaux d'expression, *i.e.* $\mathcal{A}_a = \{0, 1\}$. Une partie de la dynamique locale de ce système est donnée dans le tableau ci-dessous :

| | | | |
|---|---|---|---|
| $f \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = 0$ | $f \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} = 0$ | $f \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} = 1$ | $f \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = 0$ |
| $f \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} = 1$ | $f \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} = 0$ | $f \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix} = 0$ | $f \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix} = 0$ |
| $f \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} = 1$ | $f \begin{pmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} = 1$ | $f \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix} = 0$ | $f \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix} = 1$ |
| $f \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} = 0$ | $f \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} = 1$ | $f \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix} = 1$ | $f \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} = 1$ |
| $f \begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} = 1$ | $f \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} = 1$ | $f \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} = 1$ | $f \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} = 1$ |
| $f \begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} = 1$ | $f \begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} = 0$ | $f \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} = 0$ | $f \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} = 1$ |
| $f \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} = 1$ | $f \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} = 0$ | $f \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} = 1$ | $f \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} = 1$ |

Le graphe d'interaction suivant est construit par calcul des dérivées :

$$G\left(\partial F\left(\begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}\right)\right) = \text{Diagram}$$

F a deux états stationnaires r et t répondant aux conditions du Théorème 3.2 (une partie uniquement de chaque état stationnaire est représentée). L'ensemble \mathbb{U} est constitué des deux cellules non colorées entourées par les cellules colorées et $\partial\mathbb{U}$ est constitué des huit cellules colorées. Donc il existe un état s tel que $G(\partial F(s, t))$ a un circuit positif élémentaire.

$$r = \text{Diagram} \quad t = \text{Diagram} \quad G\left(\partial F\left(\begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}\right)\right) = \text{Diagram}$$

FIGURE 2. (Gauche) Etats stationnaires (Droite) Circuit positif entre les gènes des deux cellules centrales

Anne Crumière

Institut de Mathématiques de Luminy (UMR 6206),
Campus de Luminy, Case 907 13288 MARSEILLE Cedex 9.

E-mail : crumiere@iml.univ-mrs.fr