Lecture on

Dynamics in gene regulation

Robert G. Endres*
Department of Life Sciences & Centre for Integrative Systems Biology and Bioinformatics, Imperial College, London SW7 2AZ, United Kingdom

A lot of effort in systems biology aims to understand how genes are regulated. This will in turn describe how cells respond to changing external conditions. Important related questions are epigenetic variability, i.e. that genetically identical cells show different phenotypes for increased fitness, and how cells switch between these different phenotypic states. Here, I will show two examples of such gene-regulatory pathways: (1) a generic bistable gene switch similar to the lac system in *E. coli*, and (2) competence in *B. subtilis*, an excitatory gene circuit activated during starvation. We will discuss fixed points and their stability using graphical approaches and linear stability analysis.

Contents

I. INTRODUCTION: EPIGENETICS, BISTABILITY, AND THE LAC SYSTEM

Populations of cells such as bacteria often exhibit strong epigenetic variation, i.e. different phenotypes despite the fact that cells have identical genetic material. This is thought to increase the fitness of the cells to sudden changes in environmental conditions at the population level [1]. While sudden changes may kill a large number of cells, the population as a whole is likely to survive as some cells happen to be equipped with the right proteins for the new environment. Once the environmental conditions are more favorable again, the population will re-grow with its epigenetic variation. This is a kind of bet-hedging strategy. Another advantage of epigenetic variation is specialisation, so subpopulations can fulfill different tasks, e.g. as part of a community in a biofilm.

The epigenetic variation across a population is sometimes described as multistability, or bistability, if only two phenotypic states are prevalent. This means that cells can be in different states and express different genes accordingly. A related example is the phenomenon of *diauxie* (double growth), discovered in 1942 by Jacques Monod and honored with the Nobel Prize in 1965 (shared with Francois Jacob). Fig. 1 shows that bacterial cells, exponentially growing on a combination of two sugars, the preferred nutrient glucose and the inferior sugar galactose. First, cells only metabolize glucose and only when depleted, do they grow on galactose with a lower growth rate. In between the two growth phases there is a lag phase, which was later found to be caused by the induction of the new enzymes necessary for utilising galactose. Already in 1957, by Novick and Weiner showed with experiments on induction kinetics in a chemostat that there are two subpopulations of induced and uninduced cells. This all-or-none enzyme induction is typical for bacteria in that bacteria only make enzymes in time when needed to maximise growth.

The described work led to the first models of gene regulation, especially in terms of understanding the Lac operon for utilising sugar lactose and how diauxie is achieved (pioneered by Francois Jacob). The schematic in Fig. 2 shows how the cell responds to both glucose and lactose (here analogue TMG). Glucose suppresses inlet of lactose

* E-mail: r.endres@imperial.ac.uk

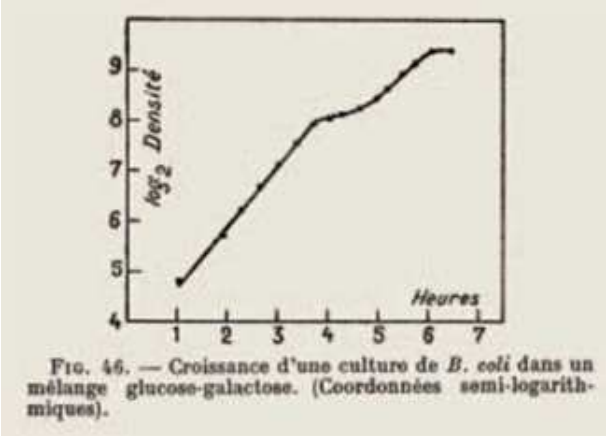

FIG. 46. — Croissance d'une culture de *B. coli* dans un mélange glucose-galactose. (Coordonnées semi-logarithmiques).

FIG. 1: Seminal diauxie curve of bacterial growth on two sugars by Jacques Monod.

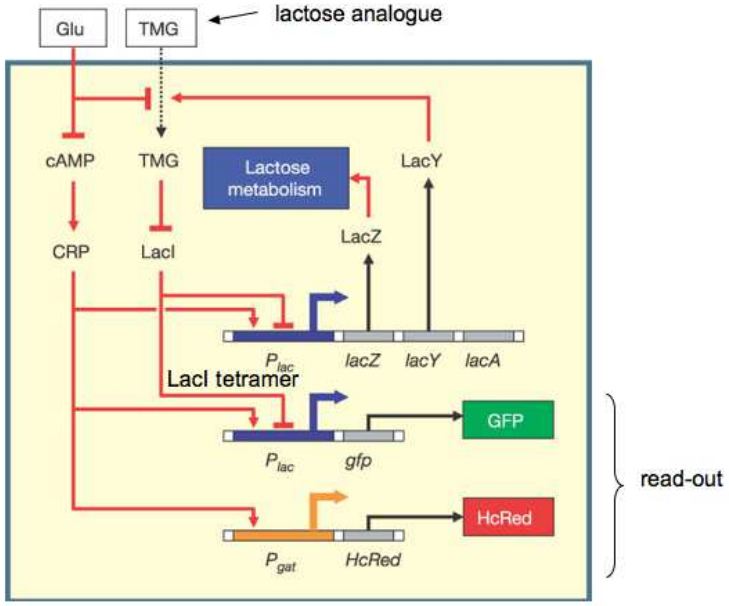

FIG. 2: Glucose and lactose update pathways, and their interconnectiveness [3].

and inhibits activator CRP of the Lac operon. Once glucose is used up, lactose can be utilised as it is activated by CRP and the repression by LacI is lifted, leading to diauxie. More recent experiments on the Lac operon with a fluorescent reporter impressively showed that cells are indeed either induced or uninduced depending on the lactose concentration (Fig. 3A) and that this depends on the history of lactose. The history or memory dependence gives rise to hysteresis, i.e. two distinct state curves shown in Fig. 3B [3].

Hysteresis, a hall mark of bistability and well-known from magnetism (Fig. 3C), means that the transition from the first to the second state occurs at a different value of some trigger molecule or inducer than the transition from the second to the first. This provides stability to prevent accidential switching. In the Lac system, the history dependence comes from the number of expressed membrane permeases LacY, which allow the inducer to enter the cell. Once a cell is fully induced, the number of permeases is only reduced by cell division. Hence, it takes about 20-30 minutes to dilute their number by a factor 1/2.

Theoretically, it is known that bistability can be achieved in a simple genetic pathway in (at least) two different ways (see Fig. 4). (1) Positive feedback and cooperativity, and (2) double negative feedback and cooperativity. Positive feedback, indicated by a + sign, means that the protein product is also a transcription factor, which activates its own

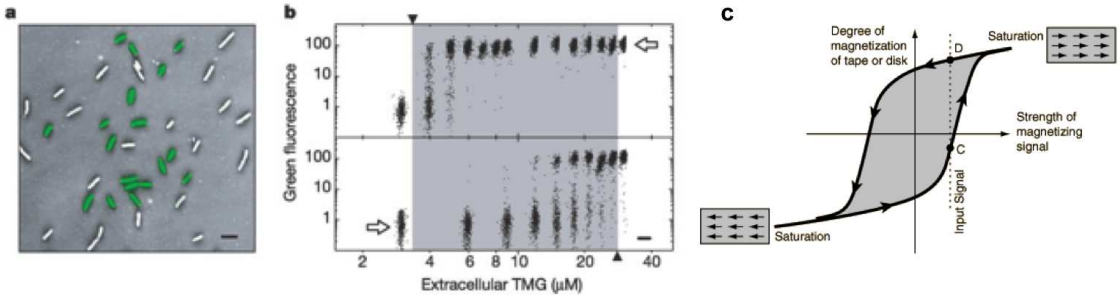

FIG. 3: Bistability of lac system in *E. coli* to gratuitous inducer TMG [3] (a). Hysteresis in lac system [3] (b) and magnetism (c).

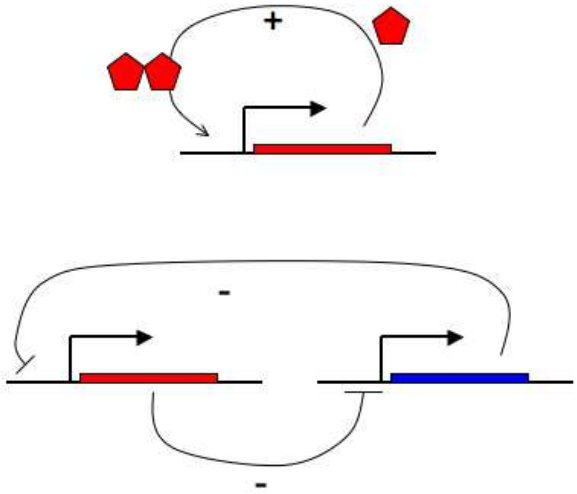

FIG. 4: There are (at least) two ways of achieving bistability: (top) positive feedback and cooperativity, and (bottom) double negative feedback and cooperativity.

expression. Cooperativity means that more than one protein is necessary, e.g. a dimer. This provides sensitivity to the concentration of the protein. Taken together, bistability in example (1) is achieved: if there is little protein available, it is not sufficient to express the gene properly. In contrast, if a sufficient amount is available, e.g. by an upward fluctuation of its expression, even more protein is expressed, leading to a full transition to the large-number state. For instance, in the bistable Lac system shown in Fig. 2, we have positive feedback through LacY and inlet of lactose, and cooperativity by tetrameric LacI. In contrast, negative feedback, indicated by a - sign, leads to suppression of the other gene. From two mutually suppressing genes, a toggle switch can be made: if protein 1 is available in large number, protein 2 is suppressed; if protein 2 happens to be available in large number, protein 1 is efficiently suppressed.

A variation of bistability is excitability, which is well known from neural action potentials. In excitability, the second state is only temporalily induced, and the cell reverses eventually again. A well known example in bacteria is competence (Fig. 5). Competence means that cells under starvation can enter a genetic programme which allows them to take up DNA from the environment, either to incooperate this foreign DNA into their own DNA for increased fitness, or as a food source. Another genetic programme, which can be activated under shortage of nutrients, is sporulation. However, this programme is a one way street - cells entering sporulation beyond a certain check point do not reverse back again but are fully committed to complete sporulation.

In the following we describe the mathematics of a generic bistable gene switch (switch 1 from Fig. 4) and subsequently the biology and mathematics of excitability is described using competence as an example.

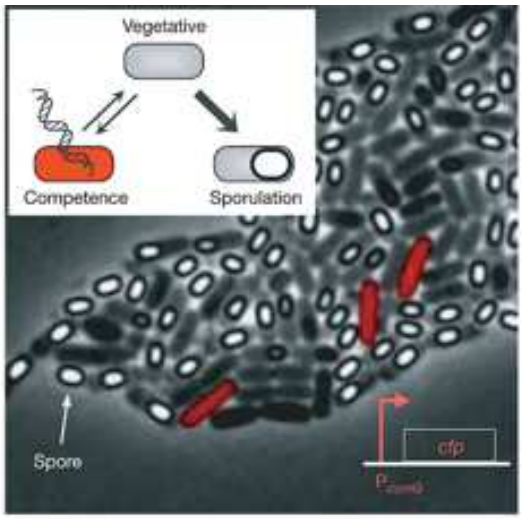

FIG. 5: Competence and sporulation in starved *B. subtilis* [5].

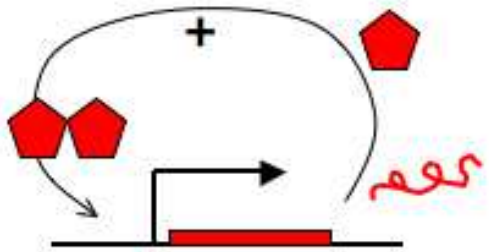

FIG. 6: Schematic of mathematical model for bistability for mRNA and protein based on positive feedback and cooperativity.

II. MATHEMATICS OF BISTABLE GENE SWITCH

In the example of Fig. 6, we have mRNA and its protein, with their dynamics given by the following equations

$$\text{mRNA} \qquad \dot{m} = \frac{p^2}{1+p^2} - bm \tag{1}$$

$$\text{protein} \qquad \dot{p} = m - ap, \tag{2}$$

where $\dot{m} = dm/dt$ describes the time derivative of the mRNA concentration m, $\dot{p} = dp/dt$ describes the time derivative of the protein concentration p, and b and a are the rates of degradation of mRNA and protein, respectively.

Our first goal is to find the fixed points of the composit dynamics, i.e. the possible steady states, and to characterise their stability. For instance, a stable fixed point is a steady state which is stable with respect to a small deviation or perturbation, so that the dynamics always return back to the fixed point. Setting $\dot{m} = 0$ and $\dot{p} = 0$, we can plot the nullclines shown in Fig. 7. Depending on the slope of the straight line from the protein dynamics, we can have 1, 2, or 3 intersections or fixed points. For large protein degradation a, there is only one fixed point, i.e. no protein.

To obtain the intersections, we solve the two equations for p by setting them equal to one another to eliminate m, i.e.

$$ap = \frac{p^2}{b(1+p^2)} \rightarrow p_1^* = 0 \tag{3}$$

$$abp^2 - p + ab = 0 \tag{4}$$

$$p_{2,3}^* = \frac{1 \pm \sqrt{1-4a^2b^2}}{2ab}. \tag{5}$$

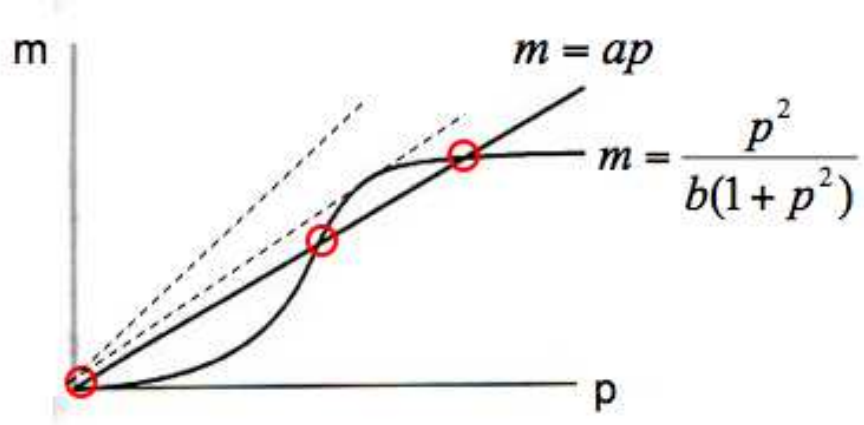

FIG. 7: Nullclines show up to three fixed points.

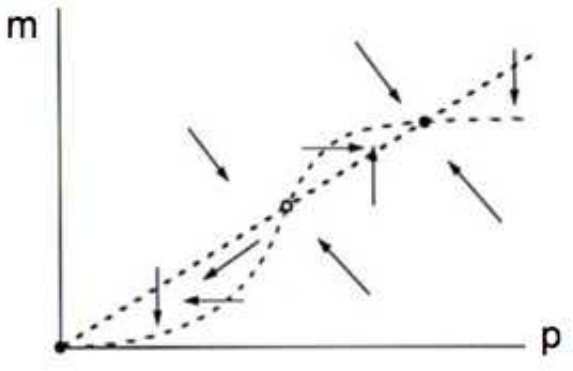

FIG. 8: Flow diagram illustrates stability of fixed points.

Hence, one solution is $p^* = 0$ as expected, while the other two solutions are given by $p^*_{2,3}$, provided that the discriminant is positive

$$1 - 4a^2b^2 > 0. \tag{6}$$

The two solutions collapse to one, if

$$1 - 4a^2b^2 = 0 \tag{7}$$

$$\text{i.e. } a_c = \frac{1}{2b} \tag{8}$$

$$p^*_c = \frac{1}{2a_c b} = 1, \tag{9}$$

defining the critical degradation rate a_c and corresponding protein concentration p^*_c.

The nullclines provide valuable information about the phase portrait, i.e. the vector field, which indicates the direction of the flow of mRNA and protein concentrations in time at any point in the m-p concentration plane. As an example, imagine a starting point for the dynamics in the left upper corner in Fig. 8. At this point, the direction of $\dot{m}$ is down and the direction of $\dot{p}$ is to the right, since

$$\dot{m} = \frac{p^2}{1+p^2} - bm < 0 \tag{10}$$

$$\dot{p} = m - ap > 0. \tag{11}$$

Conducting this kind of analysis for many points in the plane leads to the vector field shown in Fig. 8 and a rather

comprehensive illustration of the dynamics.

Next, we want to classify the fixed points according to their stability. While the stability can be inferred from Fig. 8, here we use *linear stability analysis* to formally see how the system behaves for small perturbations [4]. Generally, the dynamical equations, now written as

$$\dot{m} = f(m,p) \tag{12}$$
$$\dot{p} = g(m,p), \tag{13}$$

are nonlinear. However, for an analysis near a fixed point (m^*, p^*) we can linearise these equations and introduce new variables describing the deviations from the steady states. Exemplified for the dynamics of the mRNA, we write $m(t) = m^* + n(t)$, which provides the dynamics of the small deviation n via

$$\dot{n} = \dot{m} = f(m^* + n, p^* + q) \tag{14}$$
$$= f(m^*, p^*) + n\frac{\partial f}{\partial m}|_{(m^*,p^*)} + q\frac{\partial f}{\partial p}|_{(m^*,p^*)} + O(n^2, q^2, nq), \tag{15}$$

where the first term on the right-hand side of Eq. 15 is zero at a steady state, and higher order terms are neglected from now on. The full linearised equations for the mRNA and protein concentrations can be written as

$$\underbrace{\begin{pmatrix}\dot{n}\\ \dot{q}\end{pmatrix}}_{\dot{\vec{x}}} = \begin{pmatrix}\frac{\partial f}{\partial m} & \frac{\partial f}{\partial p}\\ \frac{\partial g}{\partial m} & \frac{\partial g}{\partial p}\end{pmatrix}_{(m^*,p^*)} \cdot \underbrace{\begin{pmatrix}n\\ q\end{pmatrix}}_{\vec{x}}, \tag{16}$$

where q is the small deviation in protein concentration $q(t) = p(t) - p^*$. To analyse the stability, we first look at the simpler case of a single molecular species, e.g. mRNA. In this 1-d example, the linearised dynamics are described by

$$\dot{n} = \frac{\partial f}{\partial m}|_{m^*} \cdot n = f' \cdot n \tag{17}$$

with its solution in time given by

$$n(t) = n_0 e^{f't}. \tag{18}$$

When f' is positive, deviations grow exponentially in time, while when it is negative, deviations are dampened out with time. These two outcomes reflect an unstable and a stable fixed point, respectively. The time scale for these outcomes is given by $1/f'$.

In higher dimensions, e.g. in our 2-d example of mRNA and protein, the matrix of first derivatives, i.e. the Jacobian, is given by

$$A = \begin{pmatrix}\frac{\partial f}{\partial m} & \frac{\partial f}{\partial p}\\ \frac{\partial g}{\partial m} & \frac{\partial g}{\partial p}\end{pmatrix}_{(m^*,p^*)} = \begin{pmatrix}-b & \frac{2p^*}{[1+(p^*)^2]^2}\\ 1 & -a\end{pmatrix} \tag{19}$$

and needs to be diagonalised first to proceed with our analysis. Diagonalisation means that we need to find a new local coordinate system, in which the matrix has only diagonal matrix elements with all the off-diagonal matrix elements being zero. If the original dynamics were $\dot{\vec{x}} = A \cdot \vec{x}$, then the new dynamics are

$$\underbrace{U\dot{\vec{x}}}_{\dot{\vec{y}}} = \underbrace{(UAU^{-1})}_{B}\underbrace{U\vec{x}}_{\vec{y}} \tag{20}$$

with matrix U explained below. Hence, our decoupled dynamics are given by

$$\dot{\vec{y}} = B\vec{y} = \begin{pmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{pmatrix}\vec{y} \tag{21}$$

with eigenvalues λ_1 and λ_2. This effectively corresponds to two independent 1-d problems, which each have a simple exponential solution as shown above. Now, λ_1 and λ_2 can be complex with real and imaginary parts, such that

$\lambda_i = Re(\lambda_i) + iIm(\lambda_i)$. The sign of the real part determines the stability, and the imaginary part produces oscillations around the steady state. The solution for the dynamics of the original linearised equations is given by

$$\vec{x}(t) = a\vec{v}_1 e^{\lambda_1 t} \; + \; b\vec{v}_2 e^{\lambda_2 t} \tag{22}$$

with the coefficients a and b determined from the initial conditions. However, to achieve this solution, we have to solve the eigenvalue equation $A\vec{v} = \lambda\vec{v}$, where $\vec{v}$ is an eigenvector and λ is the corresponding eigenvalue. To obtain the eigenvalues, we need to solve the following determinant

$$\det(A - \mathbb{I}\lambda) = \begin{vmatrix} -b-\lambda & \frac{2p^*}{[1+(p^*)^2]^2} \\ 1 & -a-\lambda \end{vmatrix} = 0, \tag{23}$$

which leads to the characteristic polynom, i.e. the following quadratic equation in λ

$$\lambda^2 + (a+b)\lambda + ab - \frac{2p^*}{[1+(p^*)^2]^2} = 0 \tag{24}$$

with general solutions

$$\lambda_{1,2} = \frac{-(a+b) \pm \sqrt{(a+b)^2 - 4\left\{ab - \frac{2p^*}{[1+(p^*)^2]^2}\right\}^2}}{2}. \tag{25}$$

Once the eigenvalues λ_1 and λ_2 are calculated, the corresponding eigenvectors $\vec{v}_1$ and $\vec{v}_2$ can be determined from the eigenvalue equation. Using $U^{-1} = (\vec{v}_1 \; \vec{v}_2)$ as the matrix with the columns given by the two eigenvectors, this then leads to the general solution in Eq. 22. Here, however, the explicit dynamics are not important. Note, that we can also introduce the trace (sum of diagonal matrix elements) and the determinant of Jacobian A via

$$\mathrm{tr}(A) \; = \; -(a+b) \tag{26}$$

$$\det(A) \; = \; ab - \frac{2p^*}{[1+(p^*)^2]^2}. \tag{27}$$

This allows us to write the eigenvalues more compactly as

$$\lambda_{1,2} = \frac{\mathrm{tr}(A) \pm \sqrt{\mathrm{tr}(A)^2 - 4\det(A)}}{2}. \tag{28}$$

For our bistability problem with three fixed points, we obtain two stable fixed points since both their eigenvalues are real (discriminant > 0) with negative real parts (det> 0,tr< 0), and one saddle point, since one of its eigenvalues is real and positive and the other one real and negative (det< 0). Generally, one can also obtain other types of fixed points, i.e. unstable fixed points (det> 0, tr> 0), stable spirals for complex eigenvalues (discriminant < 0) with negative real parts (det> 0,tr< 0), unstable spirals (same but with positive real parts, i.e. tr> 0), and limit cycles (oscillators) for purely imaginary eigenvalues (tr$= 0$).

As an illustrative example, we show that $p^* = 0$ is a stable fixed point. Setting $p^* = 0$ in Eq. 25 and simplifying the right-hand side, we obtain $\lambda_1 = -a$ and $\lambda_2 = -b$. Since decay rates a and b are positive, both eigenvalues are negative, resulting in a stable fixed point per definition.

To summarise the result of the linear stability analysis, we include the saddle point and the two basins of attraction in the phase portrait. In Fig. 9, these are shown in red and green, respectively.

III. COMPETENCE IN *B. SUBTILIS*

Our understanding of bistability will guide us through the next example on excitability in competence. As mentioned, cells of the species *B. subtilis* can temporally enter this state during starvation, allowing them to take up DNA.

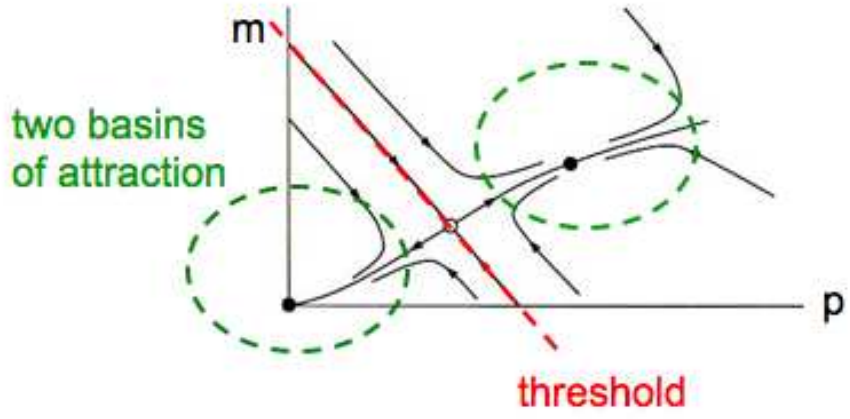

FIG. 9: Basins of attraction. Note molecule concentrations need to be positive.

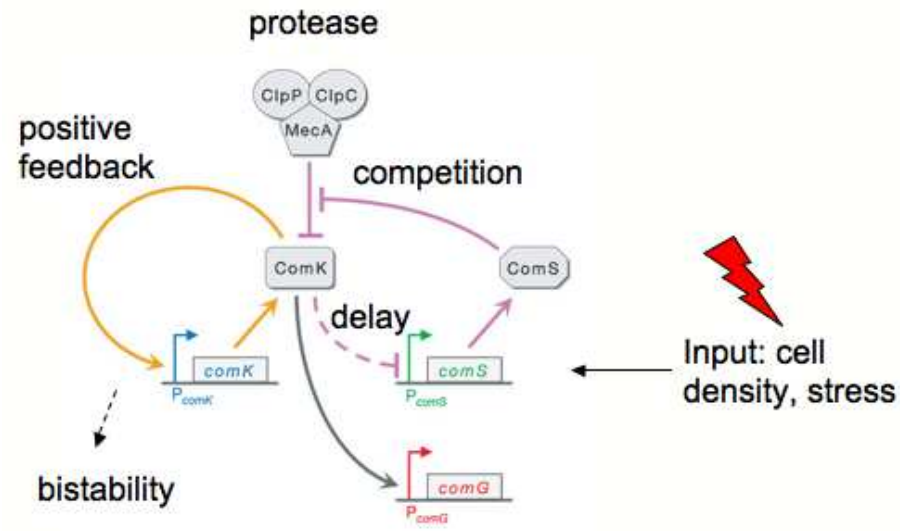

FIG. 10: Excitability in competence [5].

Important questions are *Is the transition to competence deterministic or stochastic?*, *What triggers the transition?*, *What determines the observed 20-hour long competence time?*, and *How do cells return back to the original state?*

In a Nature paper by the Elowitz lab at Caltech, these questions were addressed using a combination of experiments and modelling [5]. The minimal genetic network considered by the researchers is shown in Fig. 10. There is the master regulator ComK, a tetrameric transcription factor, which expresses the large comG operon with its hundreds of competence-related genes. It also activates itself (yellow feedback loop). Together with the cooperativity, this strongly suggests bistability with all other things neglected. So far, I described the entry into competence. To exit competence, ComK is degraded by a protease complex. To free up this protease complex for ComK degradation, its substrate competitor ComS needs to be suppressed. This is achieved by ComK, which represses expression of operon comS. As this is an indirect repression involving intermediate steps, it occurs with some time delay, allowing cells to stay in the competence state for a while.

Single-cell experiments show nicely that this conceptual model works. In Fig. 11, expression from promoter comG is shown in red and expression from promoter comK is shown in blue. When both are expressed together, the result is purple cells! In any case, these experiments show that ComG and ComK are co-expressed as predicted by the model. In Fig. 12, expression from comS is shown in green. These data show that ComG and ComS are mutual exclusive, i.e. that there is anticorrelation between the two (panel c), supporting the idea of the role of ComS as an indicator of exit

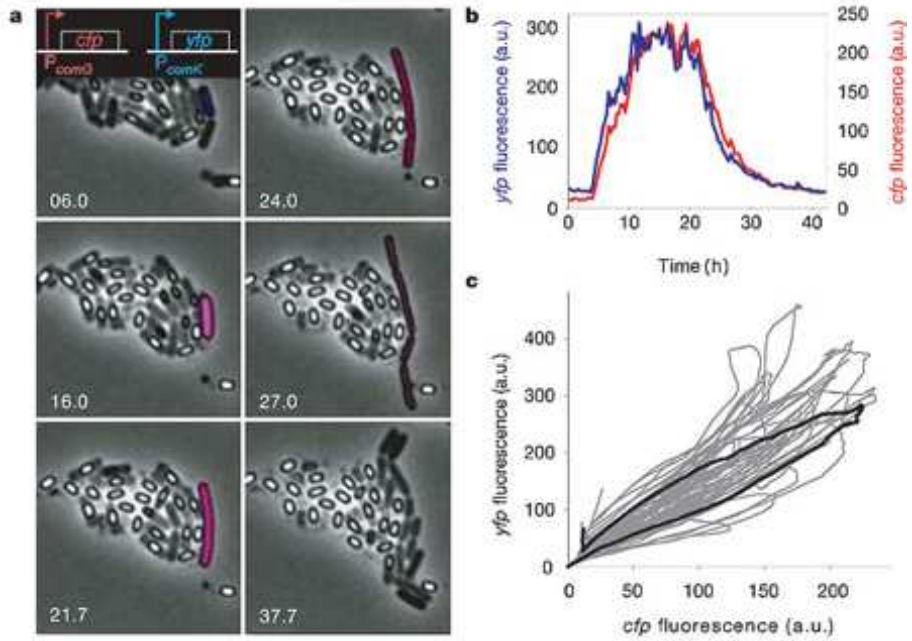

FIG. 11: ComK (blue) and comG (red) are strongly correlated [5].

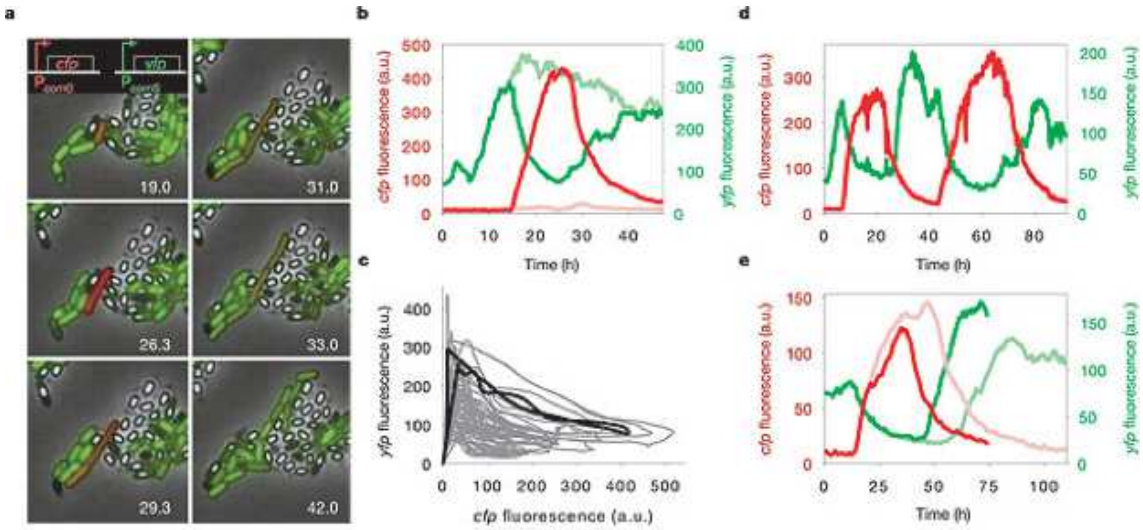

FIG. 12: ComK (red) and comS (green) are anticorrelated [5]. Entry and exit of competence are stochastic and memoryless.

from competence. The figure further shows that the entry into competence is purely stochastic (panel b), that cells have no memory of having entered competence before (panel d), and that the exit is purely stochastic as well (panel e).

To further support and quantify the conceptual model, a mathematical model was proposed. A minimal version of

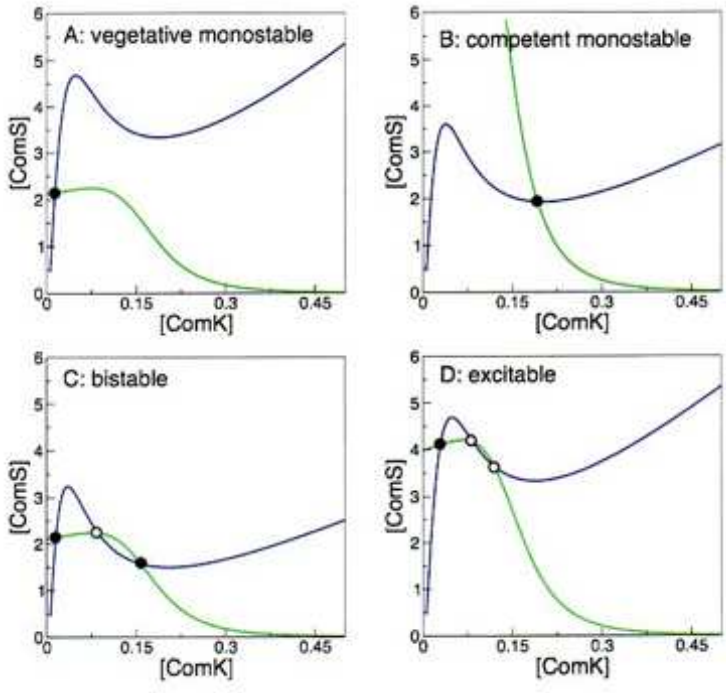

FIG. 13: Different model outcomes for different parameters based on nullclines [5].

the model only contains two variables, ComK and ComS, with their dynamics given by

$$\text{ComK} \quad \dot{K} = \underbrace{a_K}_{\text{basal expression}} + \underbrace{\frac{b_K K^n}{k_0^n + K^n}}_{\text{self-activation}} - \underbrace{\frac{K}{1+K+S}}_{\text{competitive degradation}} \tag{29}$$

$$\text{ComS} \quad \dot{S} = \underbrace{\frac{b_S}{1+(K/k_1)^p}}_{\text{repression by K}} - \underbrace{\frac{S}{1+K+S}}_{\text{competitive degradation}} + \underbrace{\zeta(t)}_{\text{noise}}. \tag{30}$$

A noise term is added already but is not used for the next figure. Figure 13 shows phase portraits with the nullclines of the two proteins for different model parameters. Depending on the paremeters, one can obtain a single vegetative state (panel A), a costly competent monostable state (panel B), a costly bistable state (panel C), as well as an excitable state, including a stable non-competent state at low ComK concentartions, a saddle node at intermediate ComK concentrations, and an unstable spiral at high ComK concentrations.

To understand the dynamics of competence better, in particular of entry and exit, the dynamics are simulated by using a noise term (see Eq. 30). Fig. 14 shows multiple trajectories from these simulations, demonstrating that large excursions in phase space lead to a long competence time as the trajectory has to go around the spiral fixed point to return to the non-competent state. This excursion is similar to the stereotypical response and refractatory period in an action potential (Fig. 15), which can be modelled with another two-species circuit - the FitzHugh-Nagumo model.

To increase the parameter regime in which the excitable solution is found, a delay in ComS repression is found to be necessary

$$\dot{S} = \frac{b_S}{1+(K(t-\tau)/k_1)^p} - \frac{S}{1+K+S} + \zeta(t), \tag{31}$$

where τ is the delay time. Fig. 16 shows how the delay helps to destabilise the spiral, necessary for the desired excitable behavior. Not the current velocity around the spiral is used, but a value from the past, which points further outwards and hence prevents the trajectory from converging to the centre.

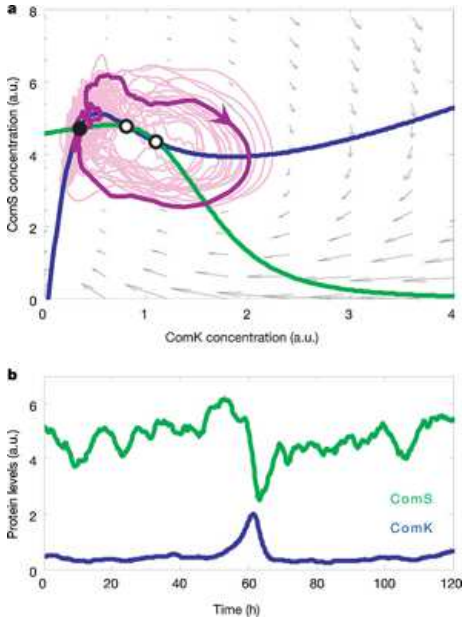

FIG. 14: Stochastic model trajectory shows large excursion in phase space [5].

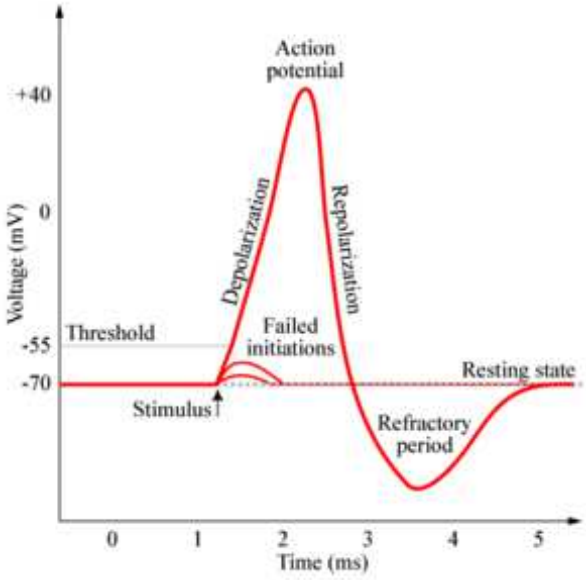

FIG. 15: Another example of excitability - the action potential.

[1] Kuipers OP, Veening JW. Phenotypic variation in bacteria: the role of feedback regulation. Nature Reviews Microbiol. **4**: 259 (2006).
[2] Veening JW, Smits WK, Kuipers OP. Bistability, epigenetics, and bet-hedging in bacteria. Annu. Rev. Microbiol. **62**: 193 (2008).
[3] Ozbudak EM, Thattai M, Lim HN, Shraiman BI, Van Oudenaarden A. Multistability in the lactose utilization network of *Escherichia coli*. Nature **427**: 737 (2004).
[4] Strogatz SH, Nonlinear Dynamics and Chaos, Westview (1994).
[5] Süel GM, Garcia-Ojalvo J, Liberman LM, Elowitz MB. An excitable gene regulatory circuit induces transient cellular differentiation. Nature **440**, 545 (2006).

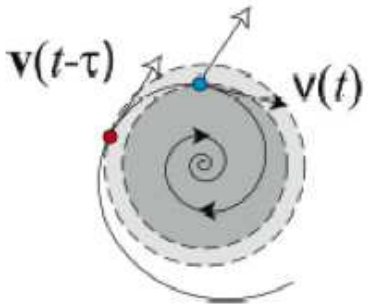

FIG. 16: Illustration of how to make spiral more unstable due to delay.