BE/APh 161: Physical Biology of the Cell

Justin Bois

Caltech

Winter, 2019

Donna and Benjamin M.
Rosen Bioengineering Center

This document was prepared at Caltech with financial support from the Donna and Benjamin M. Rosen Bioengineering Center.

8 Dynamics of gene expression and genetic switches

Last time, we looked at the equilibrium statistical mechanics of control of gene expression. In particular, we saw that separation of time scales enables us to describe the rate of transcription with thermodynamic models.

$$\text{rate of transcription} = \alpha_m P_\text{b}, \tag{8.1}$$

where P_b is the equilibrium probability that the polymerase is bound and r is a phenomenological rate constant for getting the polymerase going along the transcript. In this lecture we will use those results to write dynamical equations for the concentrations of mRNA and protein in a cell or cell population.

Sections 19.3.2-4 of *PBoC2* discuss a master equation approach to these dynamics. We will not treat them in this class, but being familiar with them will help you in many contexts of biological physics.

8.1 Basic dynamical equations for gene expression

The dynamics of mRNA transcription can be written in the form

$$\frac{d\text{m}}{\text{d}t} = -\gamma_m m + \alpha_m f(r,a), \tag{8.2}$$

where m is the mRNA concentration, r and a are the concentrations of a regulators of gene expression (repressors and activators, respectively, and they may be the gene product of interest), and $f(r,a)$ is some dimensionless function, almost always proportional to P_b, computed using methods of statistical mechanics. The first term, $-\gamma_m m$, on the right hand side represents the decrease in mRNA concentration due to the combined effects of dilution by cell growth and natural degradation due to the finite lifetime of mRNA. Similarly, we may write the dynamical equations for protein production.

$$\frac{d\text{p}}{\text{d}t} = -\gamma_p p + \alpha_p m, \tag{8.3}$$

where we consider protein degradation and production by translation of the mRNA present. Defining dimensionless time to be $\tau = \gamma_p t$, these equations may be written as

$$\frac{\gamma_p}{\gamma_m}\frac{d\text{m}}{\text{d}\tau} = -m + \frac{\alpha_m}{\gamma_m} f(r,a), \tag{8.4}$$

$$\frac{d\text{p}}{\text{d}\tau} = -p + \frac{\alpha_p}{\gamma_p} m. \tag{8.5}$$

Typically, and especially in bacteria, proteins are much more stable than mRNAs. mRNAs typically have a half life of five minutes in *E. coli* (BNID 106869) and proteins have a half life of 20 hours (BNID 111930).[11] Thus, $\gamma_p \ll \gamma_m$. This implies that the left side of the mRNA dynamical equation (8.4) is close to zero. Thus, mRNA dynamics quickly come to steady state, and as far as protein dynamics are concerned,

$$m = \frac{\alpha_m}{\gamma_m} f(r,a). \tag{8.6}$$

[11] See also the entry in *CBBTN*.

Substituting this expression into the dynamical equation for protein concentration gives

$$\frac{d\mathrm{p}}{\mathrm{d}t} = -\gamma_p p + \frac{\alpha_p \alpha_m}{\gamma_m} f(r,a). \tag{8.7}$$

Again, we see that a separation of time scales, in this case the time scales of mRNA and protein degradation, leads to simplified expressions. For notational convenience, we will define $\gamma = \gamma_p$ and $\alpha = \alpha_p \alpha_m / \gamma_m$ going forward, giving

$$\frac{d\mathrm{p}}{\mathrm{d}t} = -\gamma p + \alpha f(r,a). \tag{8.8}$$

8.2 Sudden repression

To investigate the dynamics of the concentration of a gene product, we will begin with the special case of a gene under control of a single repressor. Recall that for this case,

$$P_b = \frac{\rho}{1+\rho} \frac{1}{1 + R/N_{\mathrm{NS}}\, \mathrm{e}^{-\beta \Delta E_{\mathrm{rd}}}} \tag{8.9}$$

If we define the concentration of repressors in the cell to be $r = R/V_{\mathrm{cell}}$, and define $K = N_{\mathrm{NS}} \mathrm{e}^{\beta \Delta E_{\mathrm{rd}}}/V_{\mathrm{cell}}$, we have

$$P_b = \frac{\rho}{1+\rho} \frac{1}{1 + r/K}. \tag{8.10}$$

We absorb the factor $\rho/(1+\rho)$ into α, which gives our dynamical equation for the concentration of protein p,

$$\frac{d\mathrm{p}}{\mathrm{d}t} = -\gamma p + \frac{\alpha}{1 + r/K}. \tag{8.11}$$

Imagine that initially the repressor concentration is zero, and then suddenly jumps to r_0 at time $t = 0$. That is, $r(t) = 0\theta(t)$, where $\theta(t)$ is the Heaviside function. If the protein level is initially at steady state, then, for $t < 0$, we have

$$-\gamma p + \alpha = 0, \tag{8.12}$$

giving $p = \alpha/\gamma$ for $t < 0$. So, we have our dynamical equation to solve with initial condition.

$$\frac{d\mathrm{p}}{\mathrm{d}t} = -\gamma p + \frac{\alpha}{1 + r_0/K}, \tag{8.13}$$

$$p(0) = \alpha/\gamma. \tag{8.14}$$

This is a first order linear differential equation, which may be solved by integrating factor to give

$$p(t) = \frac{\alpha}{\gamma} \frac{1 + (r_0/K)\mathrm{e}^{-\gamma t}}{1 + r_0/K}. \tag{8.15}$$

In looking at this expression, we first see that the decay rate of the protein sets the time scale of the dynamics, as time only appears multiplied by γ. The result is an exponential decay to the new repressed setpoint of gene expression level. We are essentially just waiting for the existing protein to degrade or get diluted.

Computing the ratio

$$\frac{p(t\to\infty)}{p(t=0)} = \frac{1}{1+r_0/K}, \tag{8.16}$$

gives exactly the fold change we expect from repression. Generally speaking, we can absorb the factor $\rho/(1+\rho)$ into the parameter α, and take $f(r,a)$ to be the fold change we computed from statistical mechanical theory of gene expression regulation.

$$\frac{d\mathrm{p}}{\mathrm{d}t} = -\gamma p + \alpha \cdot \text{fold change}. \tag{8.17}$$

8.3 A synthetic genetic switch

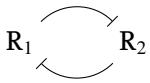
We have worked out how to incorporate repression (and also activation, since we know how to compute the fold change for activation) into the dynamical equations for protein levels. In a sense, we have mathematized the cartoon without first drawing the cartoon. Usually when we denote repression, we use a flat arrowhead. So, if repressor R represses production of species X, we write

$$\mathrm{R} \dashv \mathrm{X}.$$

Similarly, if an activator A activates species X, we connect A to X with an arrow.

$$\mathrm{A} \longrightarrow \mathrm{X}.$$

We now consider an interesting situation. We have two repressors, 1 and 2, which are *mutually repressive*. In this case, we draw the interactions between the gene products as



Such a collection of genes that control each other is called a **genetic circuit**. This circuit with mutual repressors was developed by https://doi.org/10.1038/35002131 in 2000, and is termed a "toggle switch" for reasons that will become clear as we proceed with our analysis of it.

8.3.1 Dynamical equations for the toggle switch

For the toggle switch, we will assume that n_1 repressors may bind to the promoter region of repressor 2, and n_2 repressors may bind to the promoter region of repressor 1. We assume further that the energy of binding m repressors is m times the energy of binding a single repressor. That is, all repressor bindings have the same energy. So, each binding event has the same dissociation constante, K. It can be shown the the fold change in the probability that the polymerase is bound for n repressor binding sites is

$$\text{fold change} = \frac{1}{(1+r/K)^n}. \tag{8.18}$$

This is consistent to what we derived in the last lecture for $n=1$ and $n=2$.

Though we need not make these assumptions for the following analysis, for simplicity, we will take $K_1 = K_2 \equiv K$, $\alpha_1 = \alpha_2 \equiv \alpha$, and $n_1 = n_2 \equiv n$. We assume further that both repressor 1 and

repressor 2 have the same decay rate γ, which would be the case if the repressors are fairly stable and the decay rate in concentration is set by dilution. Then, our dynamical equations are

$$\frac{\mathrm{d}r_1}{\mathrm{d}t} = -\gamma r_1 + \frac{\alpha}{(1+r_2/K)^n}, \tag{8.19}$$

$$\frac{\mathrm{d}r_2}{\mathrm{d}t} = -\gamma r_2 + \frac{\alpha}{(1+r_1/K)^n}. \tag{8.20}$$

As we proceed to analyze these dynamical equations, we can first nondimensionalize them. We can do so by defining

$$\tilde{r}_1 = r_1/K, \tag{8.21}$$

$$\tilde{r}_2 = r_2/K, \tag{8.22}$$

$$\tilde{t} = \gamma t, \tag{8.23}$$

$$\tilde{\alpha} = \alpha K/\gamma, \tag{8.24}$$

$$\tag{8.25}$$

The resulting dimensionless equations are

$$\frac{\mathrm{d}\tilde{r}_1}{\mathrm{d}\tilde{t}} = -\tilde{r}_1 + \frac{\tilde{\alpha}}{(1+\tilde{r}_2)^n}, \tag{8.26}$$

$$\frac{\mathrm{d}\tilde{r}_2}{\mathrm{d}\tilde{t}} = -\tilde{r}_2 + \frac{\tilde{\alpha}}{(1+\tilde{r}_1)^n}. \tag{8.27}$$

We will henceforth drop the tildes for notational convenience, proceeding with the understanding that we are dealing with the dimensionless versions of the respective variables and parameters.

8.3.2 Fixed points and nullclines

In analyzing the differential equations, we will first compute the **nullclines**, which are the curves in the r_1-r_2 plane where the repsective time derivatives vanish. There are two nullclines, one for $\mathrm{d}r_1/\mathrm{d}t = 0$ and one for $\mathrm{d}r_1/\mathrm{d}t = 0$. The nullclines are defined by

$$\frac{\mathrm{d}r_1}{\mathrm{d}t} = 0 = -r_1 + \frac{\alpha}{(1+r_2)^n}, \tag{8.28}$$

$$\frac{\mathrm{d}r_2}{\mathrm{d}t} = 0 = -r_2 + \frac{\alpha}{(1+r_1)^n}. \tag{8.29}$$

Solving these equations gives

$$r_1 = \frac{\alpha}{(1+r_2)^n}, \tag{8.30}$$

$$r_2 = \frac{\alpha}{(1+r_1)^n}. \tag{8.31}$$

A steady state is achieved when these nullclines cross, since this implies that $\dot{r}_1 = \dot{r}_2 = 0$, where the over-dot implies time differentiation. A steady state, or crossing of nullclines, is called a **fixed point**.

By symmetry a fixed point exists when $r_1 = r_2$. The fixed point is given by

$$r_1 = \frac{\alpha}{(1 + r_1)^n}, \tag{8.32}$$

which can be rearranged to give

$$r_1(1 + r_1)^n = \alpha. \tag{8.33}$$

The right hand side is constant and the left hand side is a monotonically increasing function of r_1 for positive r_1 (the only physically allowed values) and $n \geq 0$ (which is true by construction). Furthermore, the left hand side grows from zero without bound, thereby crossing the horizontal line at α exactly ones. Therefore, the fixed with $r_1 = r_2$ is unique and is given by the solution to (8.33). We will investigate this fixed point in the next section.

I will not work through the mathematics here, but it can be shown that in addition to the fixed point with $r_1 = r_2$, two other fixed points exist, one with $r_1 > r_2$ and one with $r_2 > r_1$, both with the same $|r_1 - r_2|$, provided $n > 2$, and

$$\alpha > \frac{n^n}{(n-1)^{n+1}}. \tag{8.34}$$

It can be further shown that there are no other fixed points.

A plot of the nullclines along with fixed points are shown in Fig. 17. For the parameter values used in making the plot, there are three fixed points.

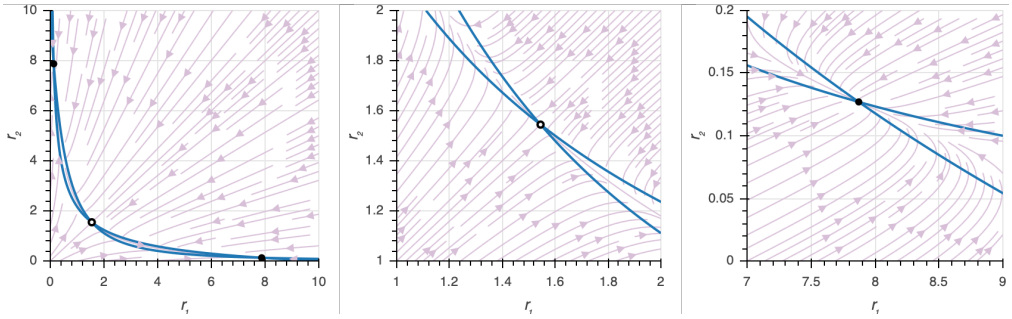


Figure 17: Left, phase portrait of the toggle switch system for parameters $\alpha = 10$ and $n = 2$. Stable fixed points are shown with solid circles and the unstable fixed point as a solid circle. The arrows indicate the how r_1 and r_2 change throughout the r_1-r_2 plane. The center and right plots are details of specific fixed points.

8.3.3 Linear stability analysis

We now turn to analysis of the $r_1 = r_2$ fixed point. We will take the approach of **linear stability analysis**. We first give an introduction to the technique of linear stability analysis generically. The basic idea is that we approximate a nonlinear dynamical system by its Taylor series to first order near the fixed point and then look at the behavior of the simpler linear system. The Hartman-Grobman

theorem (which we will not derive here) ensures that the linearized system faithfully represents the phase portrait of the full nonlinear system near the fixed point.

Say we have a dynamical system with variables $\mathbf{u}$ with

$$\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t} = f(\mathbf{u}), \tag{8.35}$$

where $f(\mathbf{u})$ is a vector-valued function, i.e.,

$$f(\mathbf{u}) = (f_1(u_1, u_2, \ldots), f_2(u_1, u_2, \ldots), \ldots). \tag{8.36}$$

Say that we have a fixed point $\mathbf{u}_0$. Then, linear stability analysis proceeds with the following steps.

1) Linearize about $\mathbf{u}_0$, defining $\delta\mathbf{u} = \mathbf{u} - \mathbf{u}_0$. To do this, expand $f(\mathbf{u})$ in a Taylor series about $\mathbf{u}_0$ to first order.

$$f(\mathbf{u}) = f(\mathbf{u}_0) + \nabla f(\mathbf{u}_0) \cdot \delta\mathbf{u} + \cdots, \tag{8.37}$$

where $\nabla f(\mathbf{u}_0) \equiv \mathsf{A}$ is the Jacobi matrix,

$$\nabla f(\mathbf{u}_0) \equiv \mathsf{A} = \begin{pmatrix} \frac{\partial f_1}{\partial u_1} & \frac{\partial f_1}{\partial u_2} & \cdots \\ \frac{\partial f_2}{\partial u_1} & \frac{\partial f_2}{\partial u_2} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}. \tag{8.38}$$

Thus, we have

$$\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t} = \frac{\mathrm{d}\mathbf{u}_0}{\mathrm{d}t} + \frac{\mathrm{d}\,\delta\mathbf{u}}{\mathrm{d}t} = f(\mathbf{u}_0) + \mathsf{A} \cdot \delta\mathbf{u} + \text{higher order terms}. \tag{8.39}$$

Since

$$\frac{\mathrm{d}\mathbf{u}_0}{\mathrm{d}t} = f(\mathbf{u}_0) = 0, \tag{8.40}$$

we have, to linear order,

$$\frac{\mathrm{d}\,\delta\mathbf{u}}{\mathrm{d}t} = \mathsf{A} \cdot \delta\mathbf{u}. \tag{8.41}$$

2) Compute the eigenvalues, λ of A.

3)
- If $\mathrm{Re}(\lambda) < 0$ for all λ, then the fixed point $\mathbf{u}_0$ is linearly stable.
- If $\mathrm{Re}(\lambda) > 0$ for any λ, then the fixed point $\mathbf{u}_0$ is linearly unstable.
- If $\mathrm{Re}(\lambda) = 0$ for one or more λ, with the rest having $\mathrm{Re}(\lambda) < 0$, then the fixed point $\mathbf{u}_0$ lies at a bifurcation.

So, if we can assess the dynamics of the linearized system near the fixed point, we can get an idea what is happening with the full system.

8.3.4 Linear stability analysis for the toggle switch

To perform linear stability analysis for the toggle switch, we begin by writing the linearized system. We note that we are considering the fixed point where $r_1 = r_2 \equiv r_0$. We have also already derived in (8.33) that $r_0(1 + r_0)^n = \alpha$.

$$\frac{\mathrm{d}\,\delta r_1}{\mathrm{d}t} \approx -1 - \frac{n\alpha\,\delta r_2}{(1 + r_0)^{n+1}}, \tag{8.42}$$

$$\frac{\mathrm{d}\,\delta r_2}{\mathrm{d}t} \approx -1 - \frac{n\,\alpha\,\delta r_1}{(1+r_0)^{n+1}}. \tag{8.43}$$

We can write this in matrix form as

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix}\delta r_1\\ \delta r_2\end{pmatrix} = \mathsf{A}\cdot\begin{pmatrix}\delta r_1\\ \delta r_2\end{pmatrix}, \tag{8.44}$$

with

$$\mathsf{A} = -\begin{pmatrix}-1 & -\frac{n\,\alpha}{(1+r_0)^{n+1}}\\ -\frac{n\,\alpha}{(1+r_0)^{n+1}} & -1\end{pmatrix}. \tag{8.45}$$

To compute the eigenvalues of A, we compute the characteristic polynomial,

$$(1+\lambda)^2 - \frac{n^2\,\alpha^2}{(1+r_0)^{2n+2}} = 0. \tag{8.46}$$

Using the fact that at the fixed point, $\alpha = r_0(1+r_0)^n$, this can be re-written as

$$(1+\lambda)^2 - \frac{n^2 r_0^2}{(1+r_0)^2} = 0. \tag{8.47}$$

This is solved to give

$$\lambda = -1 \pm \frac{n r_0}{1+r_0}. \tag{8.48}$$

The eigenvalues are both real. One eigenvalue, $\lambda = -1 - nr_0/(1+r_0)$, is always negative. The other eigenvalue is positive if

$$\frac{n r_0}{1+r_0} > 1. \tag{8.49}$$

A **bifurcation** occurs when the behavior of a dynamical system changes qualitatively for a small change in parameter values. The fixed point at $r_1 = r_2$ goes from being stable to unstable as the quantity $nr_0/(1+r_0)$ goes from being less than one to greater than one. So, the bifurcation in this system occurs when

$$r_0 = \frac{1}{n-1}. \tag{8.50}$$

Recalling again that $\alpha = r_0(1+r_0)^n$, we have a bifurcation at

$$\alpha = \frac{n^n}{(n-1)^{n+1}} \tag{8.51}$$

We have restricted ourselves to positive integer values of n. WE note that both eigenvalues are negative if $n = 1$, since the larger eigenvalue is $\lambda = -1 + nr_0/(1+r_0)$. So, we have arrived at the conditions for the $r_1 = r_2$ fixed point to be unstable, $n \geq 2$ and $\alpha > n^n/(n-1)^{n+1}$.

8.3.5 Mutual repression as a toggle switch

We have shown that the fixed point at $r_1 = r_2$ is unstable. This means that a small perturbation away from $r_1 = r_2 = r_0$ will result in the system moving away from the fixed point. This is shown in the

phase portrait of the system, shown in Fig. 17. The arrows indicate how the system evolves. As clear in the phase portrait, the system goes toward the instable fixed point, and then turns toward a high r_1/low r_2 state, or vice versa, depending on what the initial values of r_1 and r_2 were. As can be seen in the detailed portrait around the $r_1 < r_2$ fixed point, the system evolves toward the fixed point. It is stable.

As there are two stable fixed points, the system can serve as a toggle. It has two stable states, one with high r_1 and one with low r_2. Small perturbations will not move the system from these stable states. However, a large perturbation will move the system to the other fixed point. This can be done in practice by adding inducers, such as IPTG, to cells that have this circuit architecture.

The toggle is akin to a light switch. If you balance the switch exactly in the middle, it is a fixed point, but if you bump it ever so slightly, the switch will flip up or down. Once down, say, you can jiggle the switch gently, and it will remain down. If you push hard on the switch, it will push up, and then stay there.

This demonstrates how a cell can combine gene regulation architectures to get *function*. In this case, a cell can create a **bistable** switch, which can bring the cell stably into one or two different states. Importantly, we found that at least two repressors need to be able to regulate