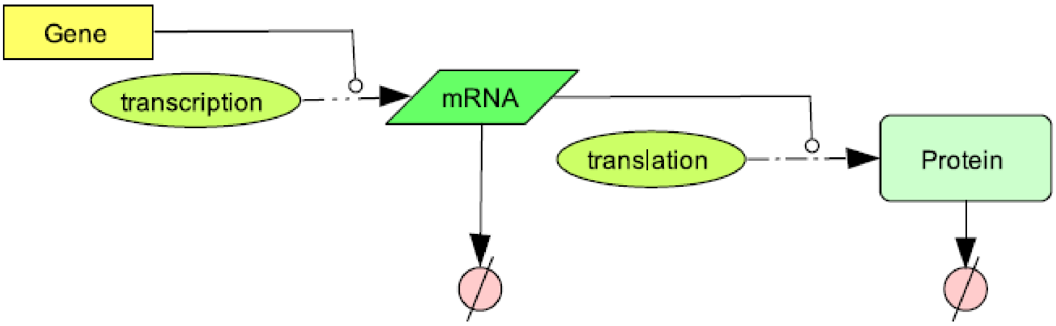
John Sy
Bioengineering
14 February 2008

Synthetic Biology
Practical 2 Write Up

Part I: Constitutive Gene Expression

The CellDesigner model used to simulate constitutive expression is shown below.



From the law of mass action, we can derive the following differential equations to model the system (P = protein).

$$\frac{d[mRNA]}{dt} = k_1 - d_1[mRNA]$$

$$\frac{d[P]}{dt} = k_2[mRNA] - d_2[P]$$

To analyze the system further, we can calculate the stead state concentrations of mRNA within the system as well as protein by setting $\frac{d[mRNA]}{dt} = 0$ and $\frac{d[P]}{dt} = 0$. We obtain that the equilibrium concentrations are as follows in terms of the parameters.

$$[mRNA]_{eq} = \frac{k_1}{d_1}$$

$$[P]_{eq} = \frac{k_1 k_2}{d_1 d_2}$$

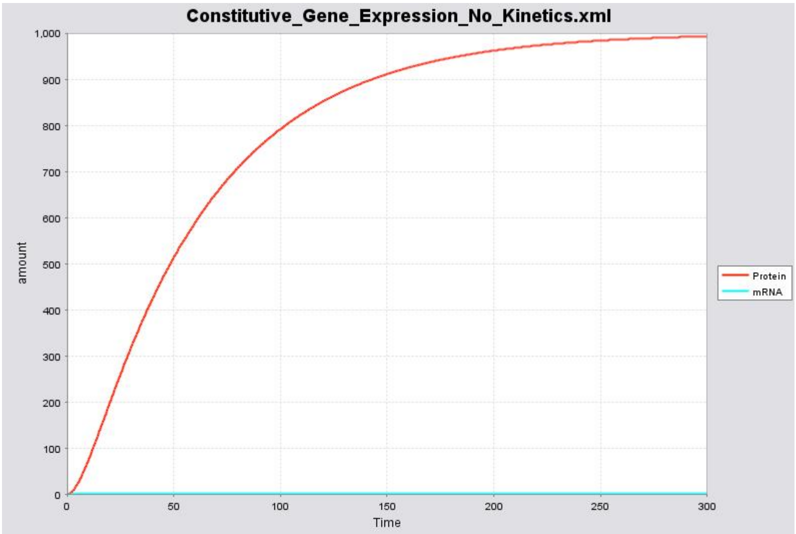
Given some values for the typical half-life for mRNA and protein, we can derive the values of d_1 and d_2. From the previous practical, we found that the half-life of the reaction is related to the rate constant for a first order reaction as per the equation below.

$$d = \frac{\ln 2}{t_{1/2}}$$

The half-life of mRNA is approximately 5 minutes, and the corresponding degradation, d_1, value is 0.139 min^{-1}. Similarly, the half-life of protein is approximately 40 minutes, giving d_2 = 0.017 min^{-1}.

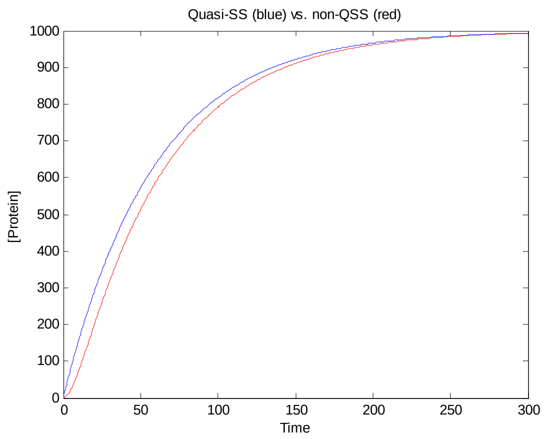
We can also calculate the values of k_1 and k_2 knowing the equilibrium values of mRNA and protein in the cell. This can be done by considering the steady state equations above and solving for the appropriate parameter. The average number of mRNA molecule per gene in E. coli is 2.5, giving a k_1 value of 0.348 (average transcription rate). The average number of protein per gene in E. coli is 1000, giving a k_2 value of 6.790 (average translation rate).

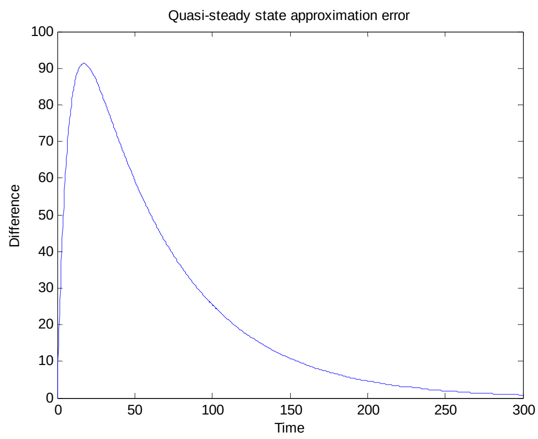
Below is the simulation output of constitutive expression. As is seen, the protein levels reach a steady state after about 300 seconds, corresponding to the equilibrium concentration of proteins that was given. mRNA levels are also at equilibrium, but at 2.5 molecules per gene, so in the graph below, the dynamics of mRNA levels are not easily seen.



Let us now explore the quasi-steady state approximation, where we assume that the mRNA concentration is constant but the levels of protein are not. We thus assume that $\frac{d[mRNA]}{dt} = 0$ and substitute $[mRNA]_{eq} = \frac{k_1}{d_1}$ into the expression governing the rate of change of protein to be left with $\frac{d[P]}{dt} = \frac{k_1 k_2}{d_1} - d_2[P]$. From the equation that was given in the practical sheet, it is clear that $s = \frac{k_1 k_2}{d_1}$ and $d = d_2$. In the figure below, we plot both the quasi-steady state approximation (in blue) and the original model without making the approximation (in red). To gain a better understanding of what is going on, it is helpful to calculate the difference between the two curves at each point in time as shown in the figure below.

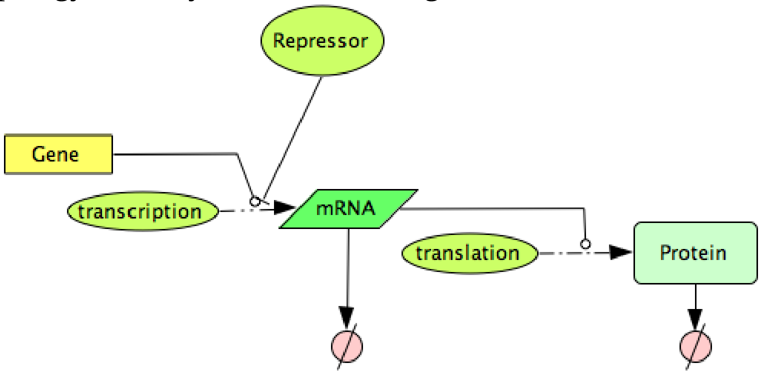
As we can see, at longer time periods, the difference between the two approaches zero, and so we can say that only at long time periods does the quasi-steady state approximation hold. If we are concerned with short time scales under 150 minutes, it is difficult to say for certain that the quasi-steady state approximation holds and we should be using the original model. However, one advantage of using this model is that we can reduce one of the variables so that we only have one equation rather than two. This is important if we are considering a multi-dimensional model and we want to reduce our model. It is also advantageous to take the quasi-steady state approximation if we are only concerned with the steady-state values and not the transient states.





Part II: Activated and Repressed Gene Expression

We define the topology of the system with the diagram below.



From the law of mass action and modelling the repressor through a Hill equation, we obtain the equations below that describe the system:

$$\frac{d[mRNA]}{dt} = \frac{k_1 k_m{}^n}{k_m{}^n + R^n} - d_1[mRNA]$$

$$\frac{d[P]}{dt} = k_2[mRNA] - d_2[P]$$

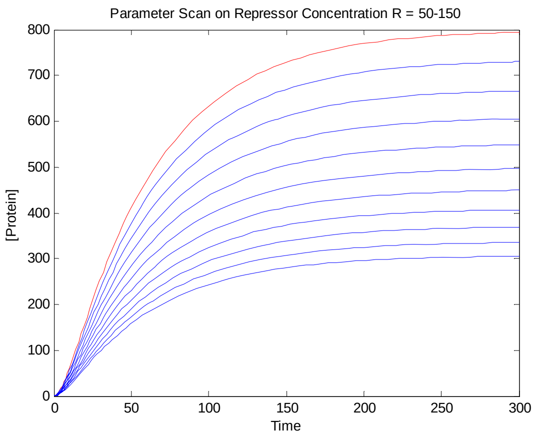
This system suddenly becomes non-linear and much more difficult to analyze analytically. We resort to good numerics to investigate this system (as run in CellDesigner).

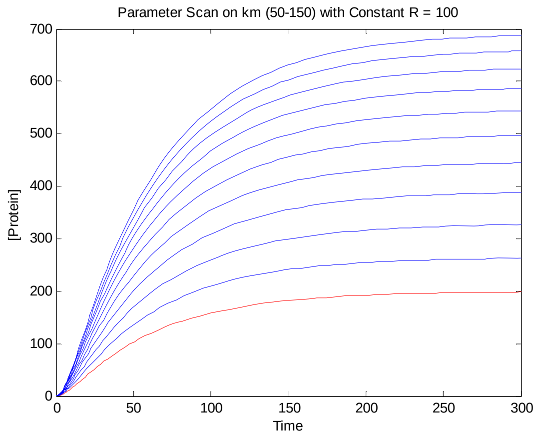
To compute the transfer function, we only consider the steady state such that $\frac{d[mRNA]}{dt} = 0$ and $\frac{d[P]}{dt} = 0$. This gives us that the steady state value of mRNA and protein is:

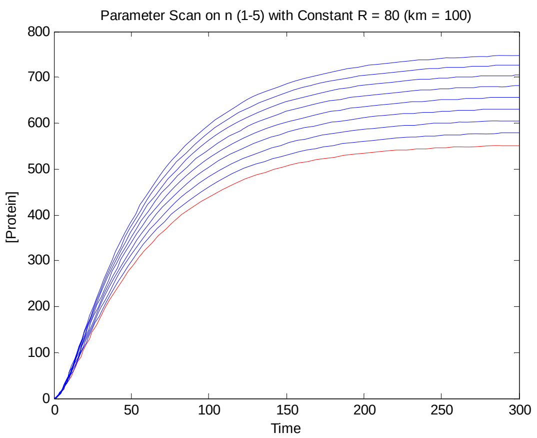
$$[mRNA]_{eq} = \frac{k_1}{d_1} \times \frac{k_m{}^n}{k_m{}^n + R^n}$$

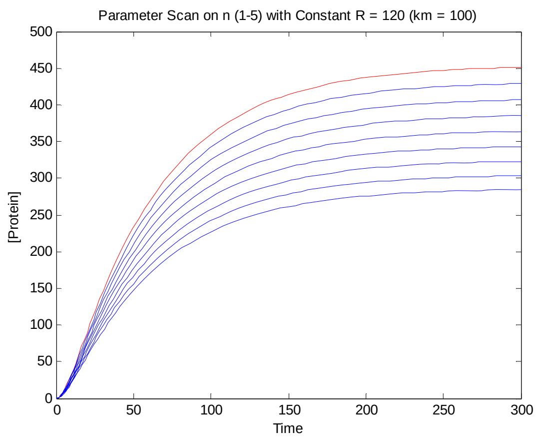
$$[P]_{eq} = \frac{k_1 k_2}{d_1 d_2} \times \frac{k_m{}^n}{k_m{}^n + R^n}$$

From the equations above, we can see that the value of k_m and n have no effect on the system if there is no repressor in the system. An ordinary increase of repressor is expected to decrease the steady state concentration, and that is what is seen in the top left graph below. Once k_m is on the order of R is there an effect on the steady-state values of the system. Increasing the value of k_m will tend to increase the steady state concentration for a given value of R as shown in the top right graph below. The variability of n is more complex to analyze and it depends on the relative side of k_m to R. If R is less than k_m, increasing the value of n will tend to increase the steady state concentration (see bottom left), but if R is greater than k_m, increasing the value of n will tend to lower the steady state concentration (see bottom right). These relationships are easily illustrated in the figures below. The red lines on the graph show the minimum value of the parameter.

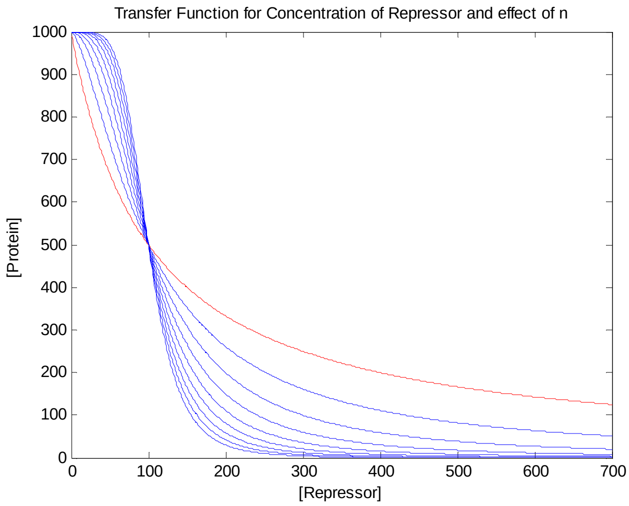








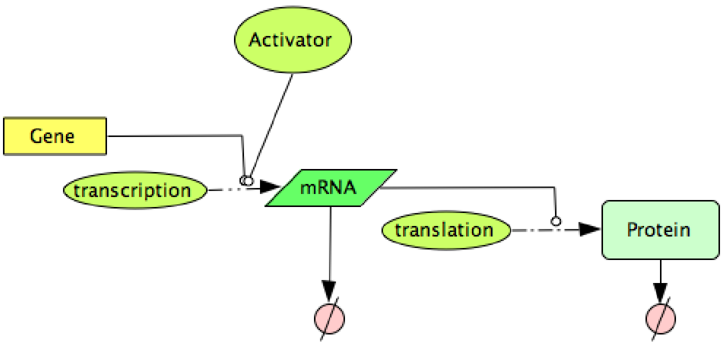
We can also investigate the transfer function between repressor concentration and protein concentration with the effect of n. This is shown in the figure below. The red line shows the transfer function when n = 1 and becomes more sigmoidal with increasing values of n. Increasing values of n will require that the repressor concentration is increased to a certain amount before a noticeable difference in the steady state will be seen. The point at which all of the lines meet is at the value of k_m, as expected from the analysis above.



One application of where a repressor circuit may be useful is when we want to create a genetic inverter such that a high input of our molecule will cause transcriptional repression to occur, lowering the output.

Similarly to the above analysis, we can also analyze an activator circuit.

We first define the topology of the system below.



Following a similar Hill-like equation, we can derive the rate equations for the system taking into account the activator.

$$\frac{d[mRNA]}{dt} = \frac{k_1 A^n}{k_m{}^n + A^n} - d_1[mRNA]$$

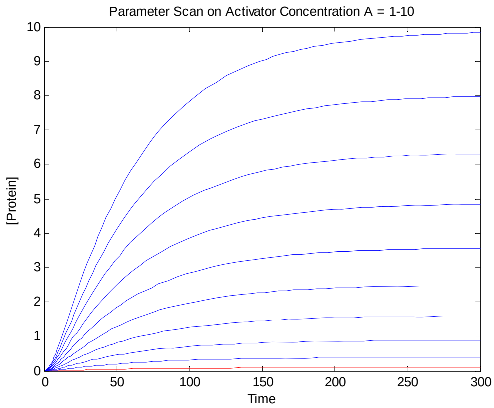
$$\frac{d[P]}{dt} = k_2[mRNA] - d_2[P]$$

To calculate the transfer function, we again set both differential equations equal to zero and solve for the steady-state protein equilibrium.

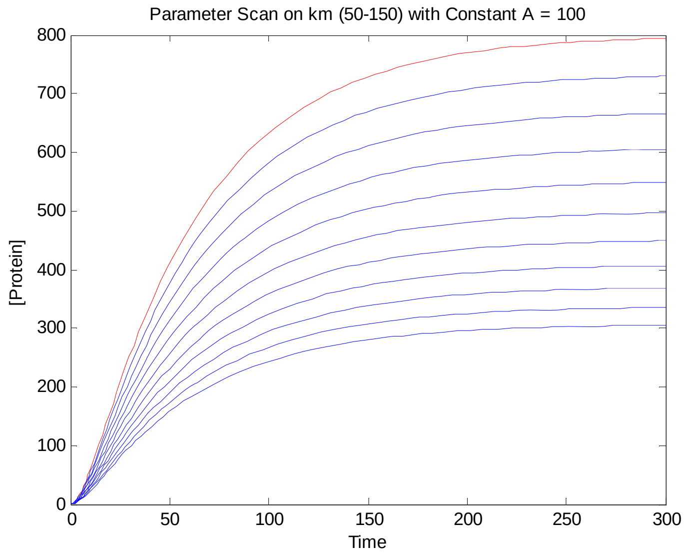
$$[mRNA]_{eq} = \frac{k_1}{d_1} \times \frac{A^n}{k_m{}^n + A^n}$$

$$[P]_{eq} = \frac{k_1 k_2}{d_1 d_2} \times \frac{A^n}{k_m{}^n + A^n}$$

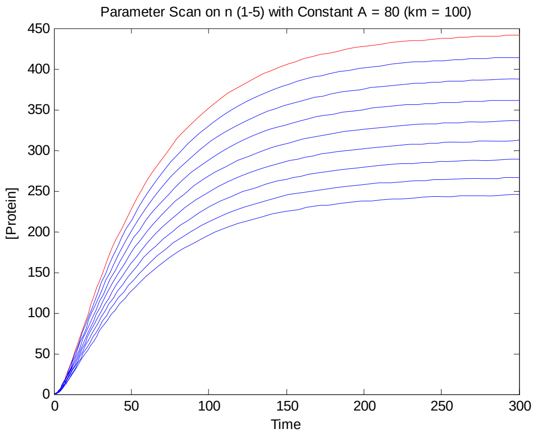
As is seen in the above equations, the dependence on activator concentration, A, is now quite strong as long as it is not around the value of k_m. Without activator, there will be no expression of both mRNA and protein, but as soon as there is activator in the system, then gene expression will occur. As activator concentration increases towards the value of k_m, the protein equilibrium approaches $[P]_{eq} = \frac{k_1 k_2}{d_1 d_2} \times \frac{1}{2}$, and as activator concentration increases even further to much larger values above k_m, then $[P]_{eq} = \frac{k_1 k_2}{d_1 d_2}$. The figure below shows the time evolution of the protein for the system. The red line corresponds to the activator concentration of 1.

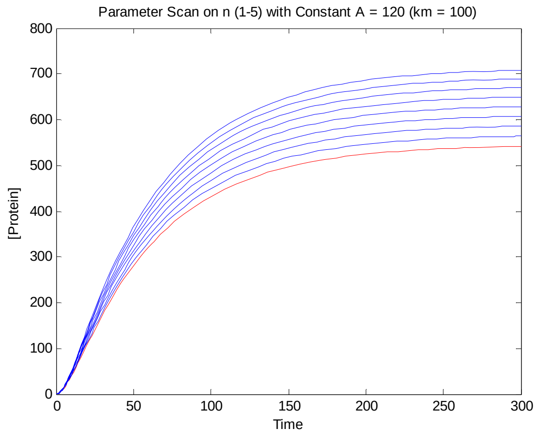


Similar to the system with the repressor, the effect k_m is only seen when it is close to the activator concentration. The effects of k_m are seen in the figure below. As expected, an increase in k_m for a given activator concentration will result in a decrease in the steady-state concentration of protein.

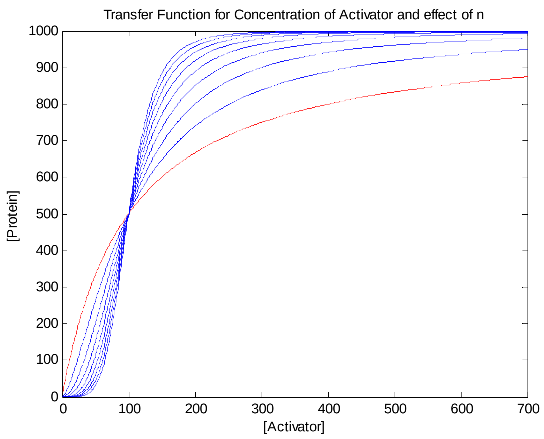


Similar, the value of n will also only be seen when there is a difference between the value of k_m and activator concentration. Below are two scenarios where k_m>A and k_m<A and the effect of different values of n. The effects are opposite to the repressor results. If A<k_m then increasing n will result in a decrease in steady-state concentration. The opposite is true if A>k_m.





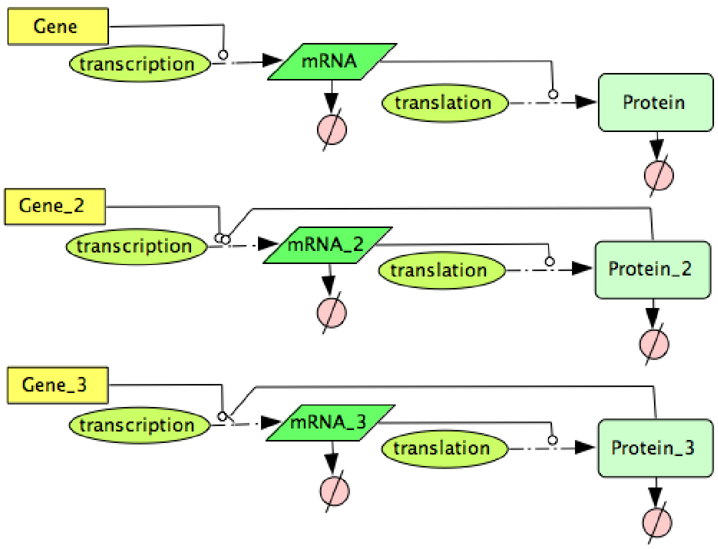
We can also investigate the transfer function (protein concentration as a function of activator concentration) and also see the effects of n from there. This is shown in the figure below. The red line shows the transfer function with n = 1 and as the value of n is increased, the transfer function becomes more sigmoidal, characteristic of increasing cooperativity. Also as n increases, it takes a much higher concentration of activator for the steady state concentration of protein to be at a high level.



An activator circuit might be helpful if we want gene expression to turn on at a certain time point and we can induce expression with the addition of an activator, similar to a switch which is initially turned off.

Part III: Positive and Negative Feedback

We again define the topology of the network in CellDesigner with one circuit that is constitutive, one that has positive feedback from its products, and on that has a negative feedback from its products.



For the Constitutive Circuit:

$$\frac{d[mRNA]}{dt} = k_1 - d_1[mRNA]$$

$$\frac{d[P]}{dt} = k_2[mRNA] - d_2[P]$$

For the Repressor Circuit:

$$\frac{d[mRNA]}{dt} = \frac{k_1 k_m^{\ n}}{k_m^{\ n} + [P]^n} - d_1[mRNA]$$

$$\frac{d[P]}{dt} = k_2[mRNA] - d_2[P]$$

For the Activator Circuit:

$$\frac{d[mRNA]}{dt} = \frac{k_1[P]^n}{k_m^{\ n} + [P]^n} - d_1[mRNA]$$

$$\frac{d[P]}{dt} = k_2[mRNA] - d_2[P]$$

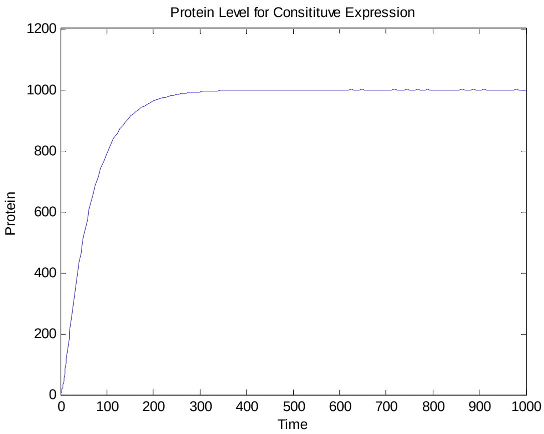
Given the values of our parameters, we can calculate the fixed points of the system to get a glimpse of the behaviour of the system. The fixed points will tell us at what values we can expect our system to reach at steady state if they are attracting fixed points.

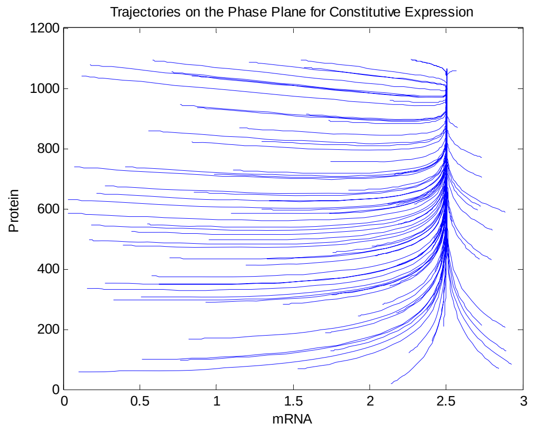
For the constitutive circuit (given the parameters), the fixed point is at (2.5, 1000). Thus at steady state, the concentration of mRNA will reach 2.5 molecules per gene and concentration of protein will be 1000 molecules per gene. What we see here is that our steady state value reaches the values that we predicted the cell to have (on average they have 2.5 mRNA molecules per gene and 1000 protein molecules per gene).

For the repressor circuit, this system has become non-linear and from matlab, we see that there is only one real fixed point and two imaginary fixed points to the system. This fixed point occurs at (0.5, 200). Because there is a repressor circuit, we see that the steady-state concentration is slightly lower for both mRNA and protein than the constitutive expression.

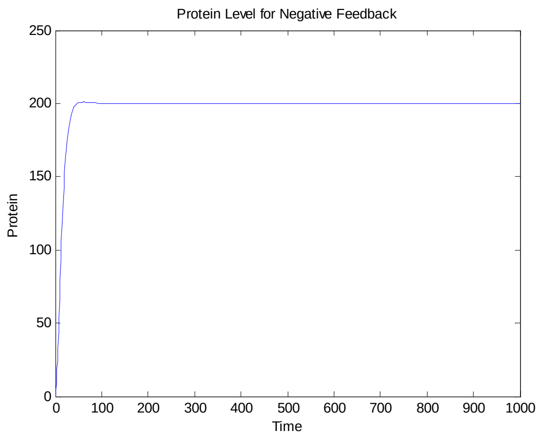
For the activator circuit, we now have three fixed points in the system. (0, 0), (2.5, 990), and (0.025, 10.1). For the fixed point (0, 0), the eigenvalues of the Jacobian matrix evaluated at the fixed point are negative, and thus show that this is a stable fixed point. For the fixed point (2.5, 990), the eigenvalues of the Jacobian matrix evaluated at the fixed point are also negative, and thus show that this is a stable fixed point. For the fixed point (0.025, 10.1), oen of the the eigenvalues of the Jacobian matrix evaluated at the fixed point is negative and one is positive, and thus show that this is a saddle node.

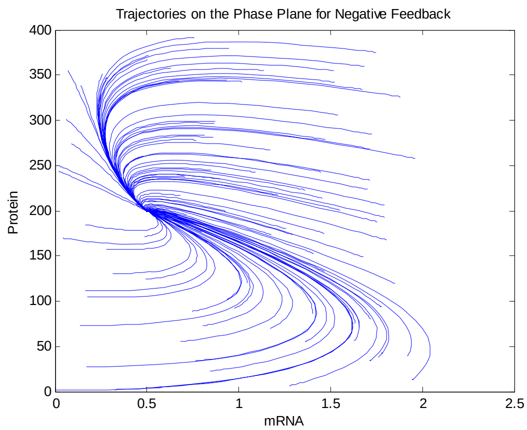
For constitutive expression, the time evolution with initial conditions of (0, 0) is shown below along with a plot of random trajectories derived from random initial conditions on the phase plane clearly attracted to the fixed point at (2.5, 1000).



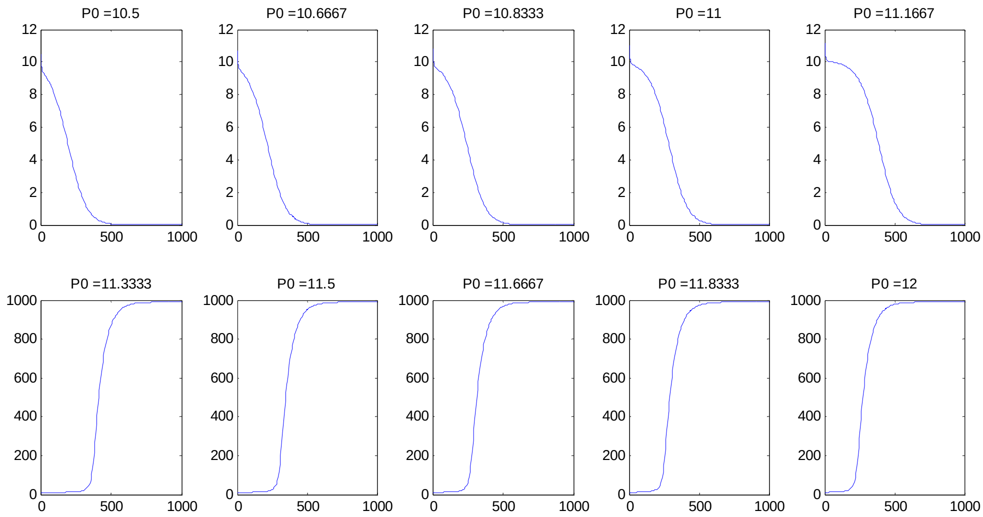


For the negative feedback circuit, the time evolution is shown again with initial conditions of (0, 0) as well as a plot of random trajectories derived from random initial conditions on the phase plane also showing that the trajectories are attracted to the fixed point (0.5, 200). Although the trajectories seem to almost be developing a cycle, there is no evidence of a limit cycle. However, since this is a non-linear problem, this does not exclude the existence of a limit cycle at different parameter values.

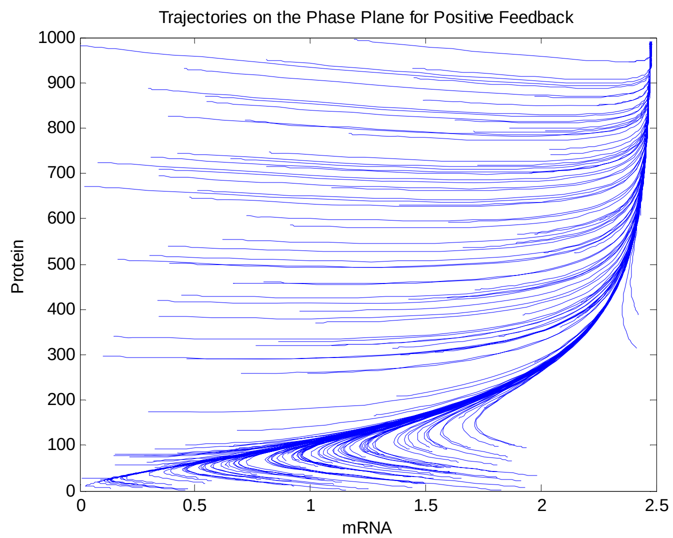




For the positive feedback circuit, the time evolution with initial conditions of (0, 0) is not very exciting since we are starting with a fixed point. If we start at various initial protein concentrations while keeping the mRNA concentration at 0, then we do see a change in the behaviour of the system around $P_0 = 11.2$ as shown in the figure below showing the effect of the saddle node in its ability to drive the system from one fixed point to another.

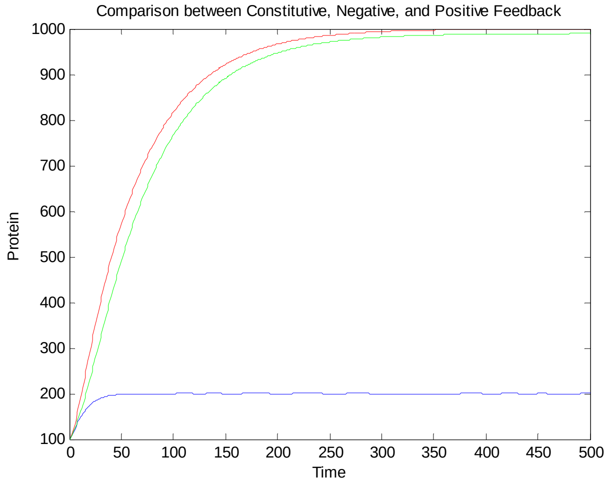


We can also visualize this effect on the phase plane where we randomly select many initial conditions and plot their trajectories in the figure below. What is interesting to note is that at low values of protein (P_0 less than 100, $mRNA_0$ less than 1.5), the trajectories are strongly pulled towards the (0.025, 10.1) fixed point, but are then driven towards the (2.5, 990) fixed point as they get close to the saddle node. All other initial conditions besides (0, 0) seem to go to the (2.5, 990) fixed point. What is not clearly evident here is the presence of the (0, 0) fixed point which should be stable as well. It would seem that the proximity of the saddle node would keep most trajectories within initial conditions above (0.025, 10.1) away from the (0, 0) fixed point and drive it towards the other fixed point.



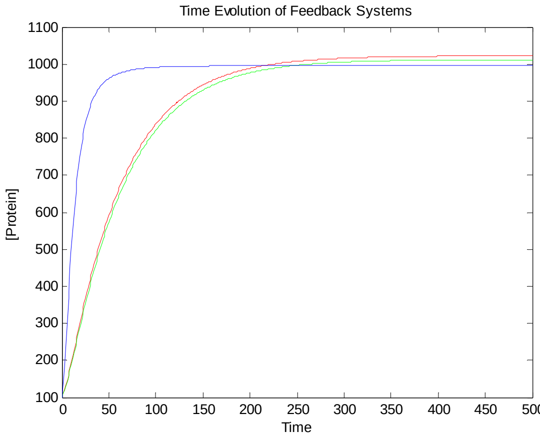
But what are the biological implications of this behaviour? This shows that there needs to be a certain level of protein already present in the cell (either injected or due to leaky gene expression) to stimulate the system to be "fully on". Under that certain level, there will be no expression at all and the protein will just get degraded before it can continuously activate the system.

Comparing the three different models is quite difficult as the behaviour varies depending upon initial conditions. To avoid the problem of the saddle node in the positive feedback loop, we will select an initial condition that is well clear of the saddle node. To this effect, we have selected the initial conditions (1, 100). The figure below shows the time evolutions of protein concentrations for each system.



The red line above shows constitutive expression, blue line shows the negative feedback system, and the green line shows the positive feedback system. Comparing the time it takes to reach steady state, the negative feedback loop is the fastest, reaching steady state in less than 50 minutes. The other two systems take about 300-350 minutes to reach steady state. Comparing the steady state level, both positive feedback and constitutive expression levels give approximately the same steady state while the negative feedback system shows a much lower steady state.

The comparison between the three systems is made slightly easier when all of them have the same steady state value. In the figure below, the three systems have calculated parameters such that the steady state value is around 1000. The blue line corresponds to the negative feedback, red line to constitutive expression, and green line to positive feedback. We can again see that the negative feedback system is faster at reaching the steady state value than the other two systems. Both the positive feedback and constitutive expression show about the same time to reach steady-state.



Once the system has reached steady state, which is more robust to pertubations in the system? To answer this question, we can look again at the magnitude of the eigenvalues at the steady fixed point. The eigenvalues tell us the speed to which the trajectories will approach the fixed point since they give the value of the exponential. For constitutive expression with the new values, the eigenvalues are -0.138 and -0.017 for the fixed point (2.5, 1022). For the negative feedback system, the eigenvalues are -0.0409 and -0.1141 for the fixed point (2.54, 955). For the positive feedback system, the parameters that have been chosen now have different fixed points. There are now only two fixed points, instead of three. The (0,0) fixed point has now become a saddle node since one of the eigenvalues is positive and the other negative. The fixed point at (2.48, 932) is a stable fixed point with eigenvalues -0.138 and -0.017, the same eigenvalues as for the constitutive expression showing that both the constitutive and positive expression have similar robustness. The negative feedback system has slightly lower eigenvalues, but might be because of the change in parameter values. If the lower eigenvalues are correct, this would mean that it takes a longer time to get back to the steady state value than for the other two systems.