# Modelling in Biology

**Briefly explain the concept of bifurcation of a dynamical system.**

A bifurcation occurs when, changing the value of one or several parameters, there is a qualitative change in the number or stability of the attractors of a dynamical system.

**Briefly explain the following types of bifurcations:** **A) Transcritical bifurcation** **B) Pitchfork bifurcation**

A) The *transcritical bifurcation* is characterised by an exchange of the stability properties of the fixed points at the critical bifurcation value, e.g., what was stable becomes unstable and vice versa.

B) The *pitchfork bifurcation* is characterised by the change of stability of a fixed point and the simultaneous appearance of two new fixed points at the critical bifurcation value.

**i)What are the distinguishing characteristic of scale-free networks ?**

The distinguishing characteristic of scale-free networks is that their degree distribution follows a power law relationship defined by P(k) ∝ k- γ. In words, some nodes act as “highly connected hubs” (high degree), but most nodes have a low degree. The scale-free model has a systematically shorter average path length than a random graph (thanks to the hub nodes).

**ii) Explain the Barabasi-Albert’s model on how to create scale-free networks.**

2. The network begins with:  
*i)* an initial network of m0 (>1) nodes   
*ii)* Growth: New nodes are added to the network one at a time  
*iii)* Preferential attachment: Each new node is connected to m of the existing with a probability proportional to the number of links that the existing node already has. Formally, the probability pi that the new node is connected to node i is pi = where ki is the degree of node i (rich gets richer)

**iii) Define and explain the clustering coefficient of a network.**

The clustering coefficient Ci for a vertex vi is given by the proportion of links between the vertices within its neighbourhood divided by the max number of links that could possibly exist between them. The clustering coefficient for the whole system is the average of the clustering coefficients.

**iv) Explain the concept behind network motifs.**

First studied by Uri Alon, they are defined to be a small sets of nodes that are connected in a specific wiring diagram and appear in the network in a significantly higher frequency than we would expect by random. These basic local elements/structures of the network usually have a functional role in the network that promotes their ubiquitous nature.

**v) Link this concept to the gene expression systems discussed in the first part of the course by giving an example and discussing it.**

Link to part 1: For example one common functional element that can be seen in many biological (and other) networks is the feed-forward-loop . In transcription networks this 3-node sub-graph is quite common, and has a role in providing robustness and resistance to noise in gene regulation: Suppose that we want gene Z to be expressed only when transcription factor X is expressed. We would not like a momentary fluctuation of X’s values to affect Z’s expression. We only want Z to be expressed if X’s levels are up for a significant period of time. For that purpose we add transcription factor Y. In order for Z to be expressed, both transcription factors, X and Y, have to be expressed. When X is first expressed it causes expression of Y. Y’s levels gradually go up, and only when there are sufficient levels of both X and Y, Z is being expressed. On the other hand, when X’s signal shuts down we want the system to immediately respond by shutting down Z’s expression. This goal is achieved by the direct connection between X and Z: as soon as X’s levels decline Z’s expression is stopped.

**Define an Erdos-Rényi network, and give an algorithm for creating an ˝ adjacency matrix A representing a directed network of N nodes, with average out-degree ¯kout. Forbid self-edges.**

In an Erdos-Rényi network, each possible edge is randomly included with a ˝ fixed probability p, independent of all other edges. The probability is chosen to give the appropriate average degree. Algorithm to create network:   
∗ Define N × N adjacency matrix A, and set entries to zero.   
∗ Loop over all pairs of nodes i, j, excluding i = j.   
∗ Set A*ij* = 1 with probability p = kout/(N − 1).   
∗ Use a random number generator to do this (generate uniform random number x in range 0 to 1; if x < p, set Aij = 1).

**What is the probability of any two specific edges both being present in an Erdos-Rényi network if each edge is present with a probability p?**

The probability of any given pair of edges being present is p 2 , since they are independent.

**A directed Erdos-Rényi network has 26 nodes, no self-edges and an average out-degree ¯kout = 5. Estimate the expected number of pairs of nodes that are connected by both possible edges.**

For each node, there are 25 possible outwards connections. Thus 1/5 of the possible edges are present and p = 1/5. Thus both edges between a pair of nodes will be present with a probability of q = p2 = 1/25. There are N(N − 1)/2 = 13 × 25 possible pairs of nodes in this network, and thus we would expect qN(N − 1)/2 = 13 pairs of nodes connected by both edges.

**A charged protein in water undergoes stochastic, overdamped motion. In one dimension, in the presence of a potential V (x), drag coefficient γ and at temperature T, the stochastic process can be modelled using *the Fokker-Planck equation* or *the Langevin equation* is a random function. Discuss the meaning of these equations, explaining how two different equations can both be used to describe the same process. Identify the terms that give rise to drift and diffusion in both cases, justifying your reasoning**

The Fokker-Plank equation describes the evolution of the entire probability distribution over time, P(x, t). The Langevin equation produces a single sample trajectory x(t) from the distribution of possible trajectories. Thus they provide two different perspectives on the same process. Generating enough sample trajectories would in principle allow the reproduction of P(x, t). In the Fokker-Planck equation, the first term leads to changes in P(x, t) over time in response to the presence of a non-zero force -dV/dx. Similarly, the first term of the Langevin equation causes the sample trajectory x(t) to move in the direction of forcing. These are then “drift” terms. In the Fokker-Planck equation, the second term causes P(x, t) to decrease when the second derivative is negative, and increase when the second derivative is positive - reproducing the tendency of diffusion to cause particles to move from areas of high concentration to nearby areas of low concentration. In the Langevin equation, the second term gives a random component to the motion of the particle, leading to Brownian-like motion and diffusion.