

## Exact Methods for Computing Biological System Dynamics

Drew Endy (<http://mit.edu/indy/>)

### *Goals for Today*

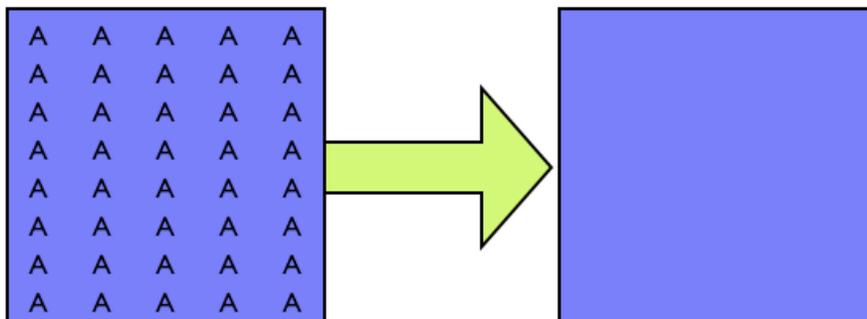
**A. When to use computational methods?**

**B. When to use exact methods?**

**C. What is the underlying physics models used by exact methods?**

### Introduction

1. Biological systems exist through time. So far in 20.181 you've developed an understanding of computational methods for representing the behavior of biological systems on long time scales (e.g., evolution) and very short or instantaneous time scales (e.g., protein structure).
2. Here, we're going to think about the behavior of biological systems on the time-scale of an individual, and the computational methods that can be used to represent their behavior.
3. Such computational methods are useful for both analysis and design of systems. For example, given a model of an existing biological system, and a description of its current state, you can compute the expected future behavior (i.e., transitions from state to state) of the system. Similarly, given a specification of the desired behavior for some notional biological system, you can compute the behavior of possible biological systems, helping to select a design that encodes the desired solution.
4. Now, before we get too caught up in "biological systems" this and that, let me state directly that there's nothing magical about the methods that we'll develop. In fact, you'll see that they are based on computational methods that were developed for simple chemical and physical systems. Whether or not such methods accurately depict the inner workings of biological systems is a topic that you should think about!
5. Let's consider an extraordinarily simple system, one involving first-order decay. A system containing molecule A decays into oblivion (below).



6. How will this system behave over time? Develop model for exponential decay based on classical chemical kinetics.

7.  $d[A]/dt = -k[A]$ . Now, how to solve this? We can integrate and get analytical solution!

8.  $d[A]/[A] = -kdt$

9.  $\ln[A]$  (from  $A_0$  to  $A_t$ ) =  $-kt$  (from  $t_0$  to  $t_t$ )

10.  $\ln[A_t/A_0] = -kt$

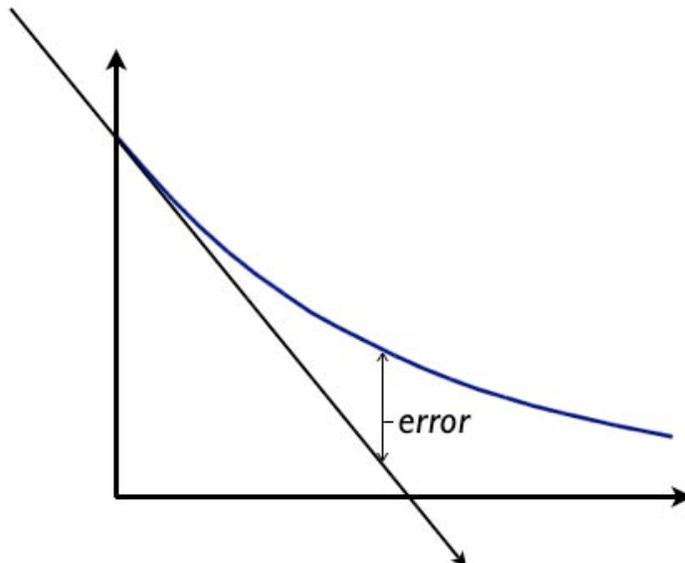
11.  $A_t = A_0(\exp[-kt])$

12. **Question A.** What happens if we can't develop an analytical solution? For example, consider more complicated cases like bacteriophage T7 gene expression, yeast pheromone response, and so on.

13. For many systems that we care about it will not be trivial (and oftentimes impossible) to develop analytical solutions describing the future state of the system. If we want to determine the expected behavior of such systems we will need to use computational methods.

14. Most computational methods are approximate, in that they imperfectly estimate the future state of a system given a description of the system and its current state. For example, consider a numerical solution to our simple first-order decay system. One easy-to-consider method (Euler's method) will compute the expected future behavior of the system using just the rate equation and the current system state. The imperfection in the estimation is called numerical error.

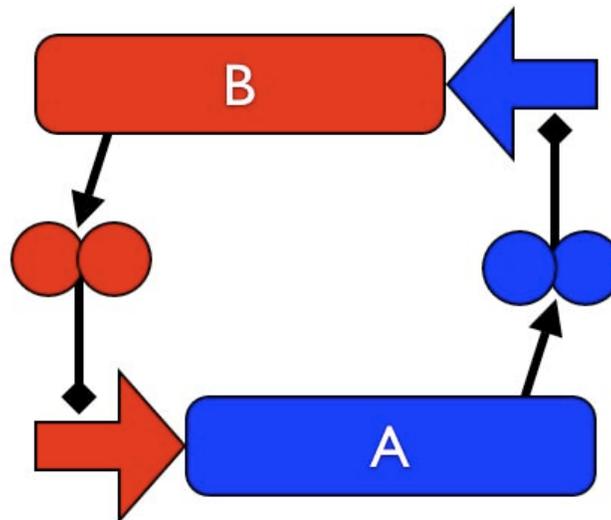
15. Sometimes numerical error is manageable, other times numerical error is a show stopper. For example, in biological systems you can usually tell when you've a problem with numerical error when the concentrations of the species in your system are negative (a physically impossible situation)!!



16. Some sources of numerical error can be dealt with by reducing the “time-step” used in the method (or by adapting the time step as a function of how quickly the system is changing). Also, advanced methods of numerical integration can be deployed which are more accurate. You can study these techniques in a course on numerical methods (e.g., in graduate school).

17. However, there is a form of error that cannot be fixed by using ever more accurate (but still approximate) numerical methods. The problem arises in systems whose behavior is sensitive to the exact ordering of individual chemical reactions. What does this mean?

18. Consider the following system, in which proteins A and B are both being produced and cross repress one another:



Mathematically, a simple form of this system can be stated as follows:

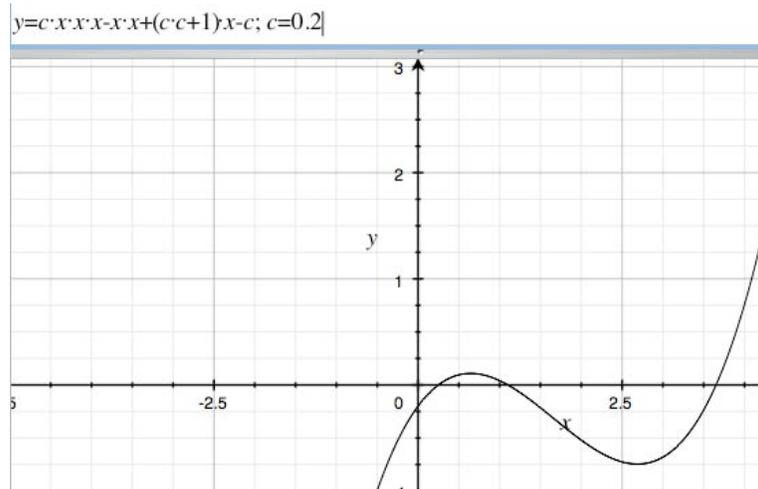
$$d[A]/dt = 1/(k+[B]^2) - [A], \quad d[B]/dt = 1/(k+[A]^2) - [B]$$

19. What's this system going to do? For example, does it have a steady state?

$$[A]_{ss} = 1/(k+[B]_{ss}^2), \quad [B]_{ss} = 1/(k+[A]_{ss}^2).$$

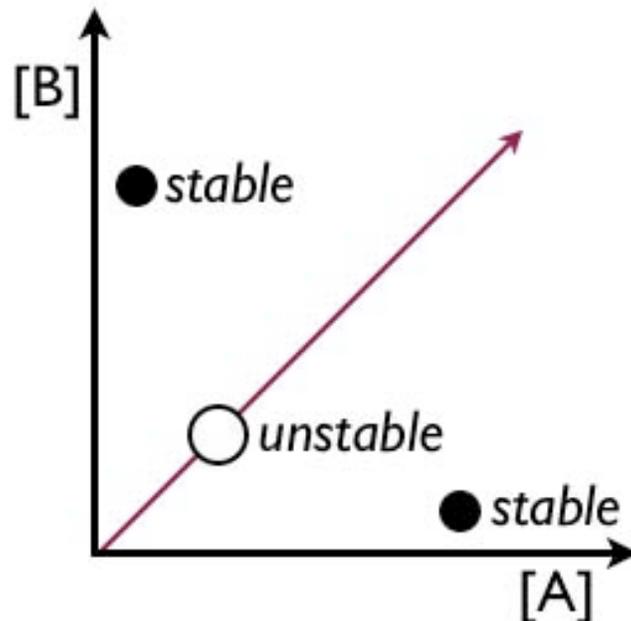
At least by symmetry we can see that there is some steady state where  $[A]=[B]$

20. But, if we solve for this steady state (e.g., by substituting  $[B]_{ss}$  into the  $[A]_{ss}$  equation) it's apparent that we have a cubic equation, suggesting up to three possible steady states. Here's the cubic:  $0=k[A]^3 - [A]^2 + (k^2+1)[A]-k$ . Now, for some values of  $k$  there are in fact three steady states (next page).



21. **Question B.** So, what's this system going to do?

22. Let's go back and look at the rate equations. Starting at a state of  $[A]=[B]=0$ , consider what happens if a molecule of A is produced first. The rate of loss of A increases and the rate of producing B decreases. Importantly, the rate of producing B decreases more than the rate of losing A. A similar behavior we result given any single A or B production event from the steady state of  $[A]=[B] \neq 0$ . This results in what's called a "bistable" system (i.e., a system with two stable steady states, aka "fixed points", and one "unstable" steady state). (graph below)



22. While we can do a lot of analytical work on this system, what if we wanted to compute the system's dynamic behavior? How could we do this?

23. For starters, we could try to use a continuous numerical method, of the sort we introduced above. But, consider what would happen if we used such a method starting from a state of  $[A]=[B]=0$ . Basically, the system would walk a “knife-edge” balance to the unstable fixed point and then... stay there! This would be very weird because there’s no way for the system to maintain itself at this fixed point in reality (it would be like a pencil balancing itself on its tip forever)!

24. The problem here is that our computational method is still approximating what’s going on inside the system. Specifically, we are only considering changes in the continuous concentration of species A and B. But, if A and B are actual molecules inside a cell then there are a finite positive integer number of each, and these numbers increment or decrement by one every time there is a production or decay reaction event, respectively. **To solve this problem we need to replace our approximate continuous numerical method with an exact discrete reaction event method.**

25. **Question C.** To develop an exact discrete reaction event method we need to consider the underlying physics model that might be used to represent individual reactions events [the method we’ll consider was developed by Daniel T. Gillespie and published in Journal of Computational Physics v22 p403-434 in 1976].

26. To begin we need to make some assumptions. The first assumption is that the reaction volume (e.g., the volume of the cell) is well mixed. [Is this a valid assumption? It depends on how quickly reactions events are happening]. The second assumption is that most collisions between molecules are non-productive (i.e., they do not result in a chemical reaction that converts substrates into products). [Note, although we won’t discuss it in 20.181, the condition of most collisions being non-reactive is required to establish an autonomous well-mixed system -- i.e., a system that mixes itself].

27. Given these two assumptions, “all” we need to do is compute the probability of any particular chemical reaction event taking place. Rather than derive the formula from scratch, let’s start with the simplest relevant equation and deconstruct it.

$$C_u = V^{-1} \pi d_{12}^2 (8kT / \pi m_{12})^{0.5} e^{(-u_u/kT)}$$

where:

C = average probability that a particular combination of molecules, u, will react

V = volume of the system

d = effective collision diameter between molecules 1 and 2, equal to  $(d_1+d_2)/2$

$(8kT\dots)$  = average apparent molecule-molecule velocity via Maxwell-Boltzman

u = activation energy for reaction u