MATH 442: Mathematical Modeling

Spring 2016

Instructor:	Wolfgang Bangerth Blocker Bldg., Room 507D bangerth@math.tamu.edu http://www.math.tamu.edu/~bangerth
Instructor:	Juliane Dannberg Blocker Bldg., Room 609B dannberg@math.tamu.edu http://www.math.tamu.edu/~dannberg
Lecture:	Tuesdays + Thursdays, 9:35–10:50am Blocker Bldg., Room 122
Office hours:	Thursdays, 2:00–4:00pm or by appointment

Writing assignment 4

Assignment

The following pages provide a number of ideas for projects for this writing assignment. You can choose any of them, or you can draw up one of your own in consultation with me – in the latter case, please send me a draft of a project assignment in the same style as the ones below by Sunday, 11/22.
All of the project ideas are purposefully open ended. The different parts are suggestions for directions, but you can go in different directions. What I want to see is that you creatively pursue ideas of your own inspired by the project prompts. Make this a research project that involves modeling and whose end product is a written report.

Specifics

- Write your report in latex, for example using http://www.overleaf.com. Use an appropriate style from their very long list of styles. Since you will have formulas, some of which will be long, a single-column style is probably most appropriate.
- We will do a peer review session during class on Tuesday 4/21/2016, so have a draft ready by that time.
- The final version of your assignment is due Wednesday 4/27/2016 5p.m. You need to submit it before this date through the eCampus "Turninit Assignments" function.
- If you have not previously done so for one of your previous projects, you will have to give a presentation on this assignment. The date for these presentations will be announced in time.

Grading

Grading will be based on the following rubric:

- *Sophistication of the model(s) and their description 20%:* This part evaluates how far you push your models in the quest to describe a situation in realistic and complete ways. This may include adding additional terms to model classes you have already seen in class, or considering larger numbers of actors, or other approaches. It also measures how well you describe your approaches.

- *Creativity 10%:* This part measures your creativity in coming up with a direction of investigation of your own. These research directions will clearly be inspired by the project ideas below, but they do not need to slavishly follow them. Rather, I want you to creatively think about directions worth investigating.

- *Insight 20%:* This part assesses the level of insight you provide. This may include descriptions of why a particular model yields a certain behavior, how well you understand the ways in which reality and your model differ, or how one would need to modify models to make them more realistic (even if this is not feasible for the current project).

- *Completeness 30%:* Your report should be complete in that you present an arc of thought that starts with a description of the problem setting, the derivation of a model, its application, and further developments. In other words, you are telling a story. This part evaluates how well you cover everything that contributes to this arc, as opposed to leaving holes where every reader would think "Why did they not investigate this obvious question?".

- *Structure, style, language, and clarity 20%:* As a text, your report should be easy to read, i.e., it should be written with a reasonable subdivision into subsections, in proper English, and at an appropriate level of technicality suitable for your audience. It should include formulas where appropriate, show Maple commands if necessary for someone like you to repeat a computation, should have axes in graphs appropriately labeled, etc. Put yourself in the shoes of your reader if you think about whether to include a detail or not.

Why I'm assigning this project

The project ideas on the next pages are projects that demonstrate many aspects you will encounter in actual industrial research or development. The first has to do with ways to simulate problems for which the model is actually straightforward, but for which you need to spend a bit of time to find the correct parameters and for which computations become slow and cumbersome pretty quickly. The second is a more conceptual one where I'd like you to think about situations that require tools you're probably not yet familiar with. These are qualities common to many modeling challenges; of course, if you choose a different project, these are the qualities I also want to see from your own idea.

Project idea 1: The solar system (A computational challenge)

In this project, you will investigate the motion of the bodies in our solar system. We know that planets move around the sun in an orderly fashion: by and large, everyone stays on her orbit for long times without interfering with other planets too much. Of course, it is this kind of stability that is necessary for a planet to be habitable – if a solar system tosses about planets every few million years, then life will not be able to develop there.
But why is this so? This project is about investigating computationally some of the conditions under which the solar system would not be as stable as it looks to us.
You already know how multiple bodies in a solar system interact with each other gravitationally. Let's assume that there is a single central star (which we denote by an index one, for convention) and $N-1$ other bodies in this system, with masses $m_1, \ldots, m_N$, respectively. Let their three-dimensional positions at time t be $\mathbf{r}_i(t), i = 1 \ldots N$, then Newton's laws tell us that they move according to the equations

$$m_i \frac{d^2}{dt^2}\mathbf{r}_i(t) = \mathbf{F}_{i,\text{total}}(t),$$

where

$$\mathbf{F}_{i,\text{total}}(t) = \sum_{\substack{j=1\\ j\neq i}}^{N} G \frac{m_i m_j}{|\mathbf{r}_{ij}(t)|^3}\mathbf{r}_{ij}(t)$$
$$\mathbf{r}_{ij}(t) = \mathbf{r}_j(t) - \mathbf{r}_i(t).$$

To make this model complete, we need initial positions $\mathbf{r}_i(0)$ and initial velocities $\mathbf{v}_i(0) = \frac{d}{dt}\mathbf{r}_i(0)$.

Part 1. For a time of your choice (what we'll call the *initial time*), find out the position and velocity of the eight planets as well as Pluto and the five largest asteroids Ceres, Vesta, Pallas, Hygiea and Interamnia. Start with the Wikipedia pages on the planets and asteroids, and Wikipedia pages they refer to; at the bottom of each page there is a list of external links from which you should find your way to various websites that list the *proper orbital elements* of each of these bodies. (Using Google to find the positions of celestial bodies at your initial time will also yield a number of websites.)
For the formulas above you also need the masses of each of these 15 bodies, which are readily available from Wikipedia. Provide all this data as part of your report.
You may assume that at your initial time the sun is at the origin of the coordinate system; in order to ensure that the system as a whole does not move, choose the initial velocity of the sun, $\dot{\mathbf{r}}_1(0)$ in such a way that the *total momentum* $\sum_i m_i \dot{\mathbf{r}}_i(0) = 0$, i.e., choose the following initial conditions for the sun:

$$\mathbf{r}_1(0) = 0,$$
$$\dot{\mathbf{r}}_1(0) = -\frac{1}{m_1}\sum_{i=2}^{N} m_i \dot{\mathbf{r}}_i(0).$$

The right hand side of the second equation contains only quantities you have already found.
Note: This part may seems easy and quick, but it might take a while to have all this information available and converted into the necessary three-dimensional coordinates for initial position and velocity. Do not leave it for the last minute!

Part 2. Use Maple (or any other programming system) to numerically solve the differential equations above for the system of 15 objects for at least 1,000 years and if you can for the longest time for which you can let your simulation run. Plot the trajectories of the planets over this time horizon and verify that they move on stable orbits that move at best by very small amounts. You may have to play with the kind of numeric solver and accuracy in dsolve to achieve this.

Part 3. The Borg (of Star Trek fame) are growing old and get ready to settle down. For reasons unknown to us they find Jupiter a friendly place. However, they still have a glimmer of evil in them so the spaceship with which they arrive slows down Jupiter by a factor of fifty before it drops off its colonists and gets the heck out of here. Re-run your simulation where Jupiter's initial velocity is only one tenth of what it was in Part 1 and observe the effect on the solar system. (Also correspondingly adjust the initial velocity of the sun.) What happens? Plot again the trajectories of all solar system bodies over your time horizon of at least 1,000 years. Interpret your findings.
Note: In addition to a plot that shows the entire trajectories at once, try to also generate a plot that zooms in to only show the fraction of the universe close to the sun, maybe the innermost 2 or 3 billion kilometers or even smaller. This should provide a more accurate intuition of what is going on here.

Part 4. In a bout of over-colonization, humanity had also played with the stability of the solar system: They found themselves a nice habitable Class M planet in the form of Mars, but it needed to get closer to Sun to become cozy and warm. Consequently, Mars was towed towards Sun and dropped off (at the initial time of your simulation) at a point exactly 1.2 times as far away from Sun as Earth along the line Sun-Earth. However, due to a last-minute tractor beam malfunction, its initial velocity – while perpendicular to the line Sun-Earth-Mars – was only one tenth that of Earth at the initial time.
Compared to the simulations made in Part 2, all that has changed now is the initial position and velocity of Mars, in much the same way as what changed in Part 3. Simulate this situation as well and determine what happens in the long run to the solar system. Compare the outcome to that of Part 3.

Part 5. Another way to disturb the orbits of the planets is to assume that there was a second big object in the solar system. While Jupiter is big (318 times as massive as Earth), it is still fairly small compared to the sun: it has less than 1,000th of the Sun's mass. Repeat the computations from Part 2 under the assumption that Jupiter had half the mass of the sun. At this mass Jupiter would be a star in itself, making the solar system a binary system with a bunch of planets thrown in. Following your computations, what can you say about the stability of planetary orbits in binary systems?
In your computations, again don't forget to adjust the initial velocity of the sun.

Bonus part (worth an additional 5% on the project). If you feel challenged, consider the interaction of stars in a dense star cluster like Messier 92 (look it up on Wikipedia!). To this end populate your simulation with not just 15 objects, but with 1,000 or 10,000 stars with randomly chosen locations in a volume representative of Messier 92. Their masses should equal one tenth to 1000 times the mass of the Sun, again randomly chosen. Equip each star with a velocity that appears reasonable (for example a velocity that allows stars to cross the star cluster in 100–1000 years). Plot trajectories of all your stars for a representative time horizon and interpret what you find.
Caution: If enough stars are involved, compute times can become fairly long. Consequently, try this with only a handful of stars at first and then increase their number as you become comfortable with the correctness of results.

Notes:

- The solar system is relatively flat, i.e., pretty much everything orbits in a plane. When you visualize orbits, make sure your plots use the same range for x/y/z to make this flatness of the solar system obvious.
- The solar system spans vast differences, and when you plot the orbit of Neptune, you will find that the orbits of Mercury and Venus are barely visible. You may therefore occasionally want to generate separate plots for the inner and outer solar system bodies.
- When you do longer simulations, investigate the `numpoints` parameter to `dsolve` to ensure that your curves all look smooth.

- If you think names like “Quaoar”, “Makemake” or “Haumea” sound cool, add a few dwarf planets to your mix!

Project idea 2: Spreading disease (A modeling challenge)

This project is about investigating the spread of pests and diseases. You will realize that this assignment has a lot of parts. You are not expected to work on all of them (though of course you can). Choose those parts that suit your talents and see how far you get (obviously, your final report should contain sufficiently many interesting parts to qualify for the "Sophistication" and "Completeness" parts of the grading rubric).

Pests and chronic diseases

Let's consider things like pests or chronic diseases: these are organisms (let's collectively call them *pathogens*) that infect hosts (e.g. plants or humans) but for which the host's defenses are not strong enough to completely eradicate them. This can be seen everywhere: Plants can be infected with aphids; trees by bark beetles; and humans by herpes simplex (HSV), hepatitis C (HCV) or human papilloma viruses (HPV). In all of these cases, infections are chronic. This is opposed to, for example, the common flu that also infects us but after some initial training period the human immune system is able to target flu viruses with sufficient specificity to completely purge them from the body, returning the population of viruses in the body to zero.
For chronic infections, let us consider the following interactions with the hosts:

- In the early stages of an infection, the pathogen population will grow rapidly.
- As the population becomes larger, host defenses become stronger and competition for available resources fiercer, limiting population growth.
- The host will not be able to completely purge the pathogen from its system, leading to a chronic infection: an equilibrium is found between pathogen reproduction and pathogen death due to lack of resources or immune system attacks.
- A small fraction of the pathogen population is able to escape the host and infect other hosts. Infection can only happen if there is an avenue for infection between the two hosts. For example, aphids can infect another plant if it is close enough by riding the wind (or being dispersed by ants); beetles can fly a certain distance; and HSV, HCV or HPV can jump from one host to another through the exchange of bodily fluids.
- The number of pathogens that per time unit leaves a host is a constant fraction of the total population on this host. For example, each day 10% of aphids may leave a plant in search for greener pastures.
- The number of pathogens that arrive at host i from host j is a fraction that depends on some distance relationship between hosts i and j. For example, if you have the flu and sneeze (you release, say, 10^9 viruses) in a room in which I am as well, there is an avenue of transmission. If I stand on the other side of the room, I will only inhale 10^3 of them whereas if we stood closer to each other it may be 10^5. This distance needs not be physical: some pairs of people are more likely to exchange HSC, HCV and HPV due to their behavior than other pairs of people, regardless of their physical distance.

Part 1 (The simplest model). Consider that there are just two hosts and that there is a method of infection between the two of them. Derive a set of differential equations for the number of pathogens $x_1(t), x_2(t)$ in the two hosts that satisfy the general principles above. Describe the individual terms and coefficients in your equations as well as what initial conditions need to be posed to make the model complete.

Part 2 (A graph based model). Now consider the more complicated case where there are N hosts and we are interested in the number of pathogens $x_i(t), i = 1 \dots N$ for each host. Complete the same tasks as for Part 1. To do so, you should consider the hosts as the vertices of a graph and possibilities for transmission as the edges; the likelihood of transmission (e.g. the distance in the introduction above) would be an edge weight. For the moment you can assume that the edges and edge weights of the graph are given somehow, i.e., use symbols instead of concrete values (we will consider particular graphs below).

Part 3-a (Aphids in the nursery). Let us now consider a nursery in which plants are arranged on a neat x-y-lattice at equal intervals – let's say there are M_x plants in the x-direction at a distance of $\Delta x = 0.3m$ each, and M_y in the y-direction at a distance of $\Delta y = 0.4m$ each, for a total of $N = M_x M_y$. On each plant lives a population of aphids (the population could be zero). Our hypothetical aphids are terrible fliers: they can drift with the wind for a maximal distance of $0.55m$: if they don't land on a plant by then they will fall to the ground and starve; if they do land on a plant they join the local population (or start the population if the number of aphids had been zero before).
Given these flight characteristics, describe in words and formulas the graph that describes this situation, i.e. state what nodes you have and what edges connect them. How many edges does the graph have? Visualize this graph for modest numbers M_x, M_y. Make reasonable assumptions (and explain them) for edge weights that describe the likelihood of transmission along an edge.

Part 3-b (Visualization). Use a computer program to find a graphical representation of this graph for $M_x = M_y = 20$. There are many that you can probably find on the internet for this task. One that is available on the calclab machines is the command line program `neato` (see `http://www.graphviz.org/pdf/neatoguide.pdf`); there are certainly also ways to do this in Maple. The input to any such program is a description of the sets of nodes and edges; the output is an image file. Remember that the graph itself only consists of nodes and edges and has nothing to do with the geometric location of these nodes in the real world. Part of the grading of this problem will be the artistic quality of the image.

Part 3-c (Numerical solution). Take the model you derived in Part 2 and the graph you got from Part 3-a and implement a program that solves these equations numerically, assuming that initially only a single plant has aphids. Your model likely has several parameters (e.g. the fraction of aphids that leave their plant's population per unit of time; the fraction of aphids that actually arrive at a destination; initial values on the single infested plant; etc.); choose reasonable values for this system and describe how you arrived at them. Document the numerical solution by plotting the aphid population on a variety of plants throughout the nursery and interpret any patterns you may see. Compare your results with what you would intuitively expect to see.

Part 3-d (Visualization). Try to visualize the aphid population on all plants at once at various times of your simulation. Alternatively you could produce a movie that shows populations. Again, part of the grading of this problem will be the artistic quality. Interpret your results.

Part 4-a (Chronic diseases on random graphs). Now consider chronic diseases and how they spread among humans. The model will be essentially the same as in Part 2 but since humans are not stationary like plants, they prove to be far better vehicles for the spread of diseases (witness, for example, the spread of the SARS or ebola viruses). On the other hand, the graphs on which diseases spread are also more complicated. In particular, it is much more time consuming to determine such graphs in the real world because it requires interviewing actual people and investigating their habits.
As a consequence, the spread of diseases is often investigated on artificially generated graphs. (In research, these graphs often have millions or more vertices, but we won't go this far for this project.) A popular approach to this end are *scale-free networks* (a special case of random graphs; scale-free graphs are often used to describe the contact networks people have with their acquaintances). Find appropriate literature that describes what these networks are and how one can construct such randomly generated graphs; summarize your findings in your own words. State how the model derived in Part 2 will now look and explain how you would expect a disease to spread from a single infected person; compare your expectations with those for the aphid model.

Part 4-b (Sample graphs). Use one of the methods described in the literature to generate a scale-free network with 400 vertices, and one with 10,000 vertices. The most commonly used algorithms allow you to specify the minimum number of edges incident on every vertex; make sure your graph has at least 2 edges

on every vertex. (If the minimum number is one, you end up with a random tree, which is a particular case of random graphs, but one not very interesting for our case.)
Visualize the two graphs you generated. (Note: Visualizing graphs with 10,000 vertices is not difficult, but it is difficult to do in a way that makes it easy to understand what is happening. The goal is not just to produce *a* picture, but a picture that leads to insight.)

Part 4-c (Numerical solution). Take the model you derived in Part 2 and one of the graphs you got from Part 4-b and implement a program that solves these equations numerically, assuming that initially only a single person is affected. As in the aphid model, choose and document reasonable coefficients for your model. Document the numerical solution by plotting the pathogen population in several people and interpret any patterns you may see. Compare your results with what you would intuitively expect to see.

Part 4-d (Visualization). Try to visualize the pathogen populations in all humans at once. Again, part of the grading of this problem will be the artistic quality of the image. Interpret your results.

As mentioned above, you are not expected to work on *all* parts outlined above – that would be a far larger project than required for this assignment. Pick and choose what you find interesting, keeping in mind that your report should contain sufficiently many interesting parts to qualify for the "Sophistication" and "Completeness" parts of the grading rubric.