**CHALMERS, GÖTEBORGS UNIVERSITET**

EXAM for
COMPUTATIONAL BIOLOGY

COURSE CODES: **FFR 110, FIM740GU, PhD**

| | |
|---|---|
| **Time:** | June 09, 2023, at $08^{30} - 12^{30}$ |
| **Place:** | Johanneberg |
| **Teachers:** | Kristian Gustafsson, 070-050 2211 (mobile), visits once around $10^{00}$ |
| **Allowed material:** | Mathematics Handbook for Science and Engineering |
| **Not allowed:** | any other written material, calculator |

---

Maximum score on this exam: 50 points (need 20 points to pass).
Maximum score for homework problems: 50 points (need 20 points to pass).
**CTH** $\geq$40 grade 3; $\geq$60 grade 4; $\geq$80 grade 5,
**GU** $\geq$40 grade G; $\geq$70 grade VG.

---

**1. Short questions [12 points]** For each of the following questions give a concise answer within a few lines per question.

a) Argue why it can be useful to analyze simple mathematical models, even though we can use a computer to solve more complicated models.

**Solution**
Lecture Notes 1.1

b) Explain what a cobweb plot is and how it is generated. Illustrate using an explicit example.

**Solution**
Lecture Notes 2.1

c) In the first problem set, you analyzed a time-delayed model with an Allee effect. Explain what the Allee effect is and give an example of a biological system where it may be important.

**Solution**
In systems where the population has a reduced reproduction or survival capacity for small population sizes, the population may go extinct if the population density becomes too small. This is the Allee effect. It may for example be important in systems where anti-predator strategies becomes inefficient in small groups.

d) In the third problem set, you were supposed to use the Gillespie algorithm to efficiently simulate a stochastic model. Explain how the Gillespie algorithm works.

**Solution**
See problem formulation in problem set 3.

e) What is meant by the phase of oscillation? Why is it useful?

**Solution**
The phase characterises the state of an oscillator (which fraction of its full oscillation cycle it has traversed). In reaction-diffusion equations it is possible to have local oscillatory reactions that are coupled spatially via diffusion. This results in a spatial distribution of oscillators with different phases. Using contour lines of the phase, we can describe how wave fronts propagate through the coupled oscillators (e.g. travelling waves, spiral waves or other waves).

f) In the Kuramoto model the number of oscillators, $N$, is assumed to be large. Discuss why this is assumed and what would be different if $N$ were not large.

**Solution**
Bernhard's lecture notes 7

g) The Lotka-Volterra model is given by

$$\begin{aligned} \dot{u} &= u(a - bv) \\ \dot{v} &= v(cu - d) \end{aligned} \tag{1}$$

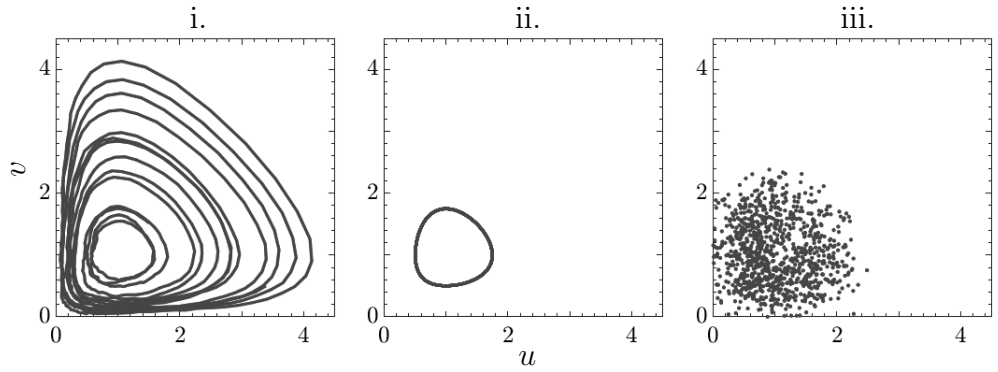
where $a$, $b$, $c$, and $d$ are positive parameters. Lotka derived these equations for a chemical reaction with $c = b$. Give an example of a chemical reaction that could give rise to Eq. (1) for two of its reactants.

**Solution**
Interpreting $b = c$ as the reaction rate at which the reactant $U$ with concentration $u = [U]$ reacts with $V$ with concentration $v = [V]$ and implementing the law of mass actions, one example reaction is

$$\begin{aligned} \mathrm{S} + \mathrm{U} &\xrightarrow{\mathrm{k_0}} 2\,\mathrm{U} \\ \mathrm{U} + \mathrm{V} &\xrightarrow{\mathrm{c}} 2\,\mathrm{V} \\ \mathrm{V} &\xrightarrow{\mathrm{d}} \mathrm{P} \end{aligned}$$

Here $S$ is assumed to have a constant concentration $s = [S]$, the reaction rate $k_0 = a/s$, and $P$ is some rest product.

h) This problem is not part of the course material this year.
The figures below show trajectories $(u(t), v(t))$ following the Lotka-Volterra dynamics (1) for three cases: without noise, with measurement noise, and with dynamical noise. Explain which figure corresponds to which case.




**Solution**
Since the noise-free Lotka-Volterra equations have closed orbit solutions, ii. must correspond to that case. Measurement noise only affects the measured value, but not the dynamics, meaning case iii. corresponds to that case. Dynamical noise affects the dynamics, giving different closed orbits in the Lotka-Volterra system, i.e. case i.

**2. Delay model of houseflies [12 points]** Consider the following model for the growth of a population of house flies of size $N$

$$\dot{N} = -dN(t) + bN(t-T)(k - bsN(t-T))\,. \tag{2}$$

Here $d$ is the per capita death rate and $b$ is the per capita rate of laying eggs. Furthermore, $k$, $s$ and $T$ are positive parameters.

a) Give a plausible explanation for the form of the system (2). What is the significance of the parameters $k$, $s$ and $T$?

**Solution**
The first term is a linear Malthus death term proportional to the population size, with per capita death rate $d$. The second term is a birth term that takes into account the time $T$ spent as an egg, i.e. at time $t$ the $bN(t-T)$ eggs laid at time $t-T$ hatches to contribute to the population. $k - bsN(t-T)$ is egg-to-adult survival ratio: $k$ is maximal fraction that survives and $s$ is a reduction in survival due to each additional produced egg.

b) Find all steady states ($N(t) = \text{const.}$) of the Eq. (2). Find a condition for one positive steady state to exist and use your explanation of the parameters in subtask a) to argue why the form of the condition is reasonable.

**Solution**
There are two steady states when $N^* = 0$ and when $N^* = (bk-d)/(b^2 s)$. The second steady state is positive if $bk > d$. This makes sense because then the egg-laying rate times the fraction of eggs surviving to adulthood is larger than the death rate (the second-order term $-b^2 sN(t-T)^2$ does not matter close to the transition of $N$ from negative to positive).

Let $b = 4$ and $k = d = s = 1$ in the subtasks below.

c) Show that close to the positive steady state, the dynamics of a small perturbation $\eta$ can be approximated by the form

$$\frac{\mathrm{d}\eta}{\mathrm{d}t} \approx C_1\eta(t) + C_2\eta(t-T)\,. \tag{3}$$

What are the expressions for the coefficients $C_1$ and $C_2$?

**Solution**

The positive steady state is $N^* = (bk-d)/(b^2s)$. Write $N = (bk-d)/(b^2s) + \eta$ and expand the dynamics (2) to first order in $\eta$:

$$\begin{aligned}
\frac{\mathrm{d}\eta}{\mathrm{d}t} &= -d\left[\frac{bk-d}{b^2s} + \eta(t)\right] + b\left[\frac{bk-d}{b^2s} + \eta(t-T)\right]\left(k - bs\left[\frac{bk-d}{b^2s} + \eta(t-T)\right]\right) \\
&= -d\left[\frac{bk-d}{b^2s} + \eta(t)\right] + \frac{bk-d}{bs}\left(k - bs\left[\frac{bk-d}{b^2s} + \eta(t-T)\right]\right) + b\eta(t-T)\left(k - \frac{bk-d}{b}\right) \\
&= (bk-d)\frac{1}{b^2s}[-d-(bk-d)+bk] - d\eta(t) + [-(bk-d)+bk-(bk-d)]\eta(t-T) \\
&= -d\eta(t) + [2d-bk]\eta(t-T)\,,
\end{aligned}$$

i.e. on the form (3) with $C_1 = -d$ and $C_2 = 2d - bk$. Using the specified parameter values, we have $C_1 = -1$ and $C_2 = -2$ (note that it is easier to solve this task by inserting $b = 4$ and $k = d = s = 1$ from the beginning).

d) Use the ansatz $\eta(t) = Ae^{\lambda t}$ in Eq. (3) to derive an equation for $\lambda$.

**Solution**

Inserting the ansatz gives

$$\begin{aligned}
\lambda Ae^{\lambda t} &= C_1 Ae^{\lambda t} + C_2 Ae^{\lambda(t-T)}\,. \\
\Rightarrow \lambda &= C_1 + C_2 e^{-\lambda T} = -1 - 2e^{-\lambda T}\,.
\end{aligned}$$

e) Find a condition on $T$ for which the steady state population of houseflies is unstable. What kind of long-term behavior do you expect?

**Solution**

Write $\lambda = \lambda' + \mathrm{i}\lambda''$. The system undergoes a bifurcation to unstable when $\lambda$ with maximal real part $\lambda'$ passes zero. Expressing the equations in subtask d) with $\lambda = \mathrm{i}\lambda''$ (so that $\lambda' = 0$) gives

$$\begin{aligned}
\mathrm{i}\lambda'' &= -1 - 2e^{-\mathrm{i}\lambda''T} = -1 - 2[\cos(-\lambda''T) + \mathrm{i}\sin(-\lambda''T)] \\
&\Rightarrow \begin{cases} 0 = -1 - 2\cos(\lambda''T) \\ \lambda'' = 2\sin(\lambda''T) \end{cases}
\end{aligned}$$

The first equation gives

$$\lambda''T = \operatorname{acos}\left(-\tfrac{1}{2}\right) + 2\pi n = \tfrac{2\pi}{3} + 2\pi n\,,$$

where $n$ is a positive integer. Inserting $\lambda'' T$ into the second equation gives

$$\lambda'' = \sqrt{3}$$

Thus, the eigenvalue with maximal real part passes zero when $T$ becomes larger than $T_c = \frac{1}{\sqrt{3}} \frac{2\pi}{3}$ (using $n = 0$ in the condition for $\lambda'' T$ above).

Since $\lambda''$ is non-zero, we expect growing oscillations for solutions starting close to the unstable steady state. Since the system does not have stable attractors we expect the dynamics to approach a limit cycle (could in principle also be a chaotic dynamics).

**3. Model for spreading of a chronic disease [12 points]** The following is a model for spreading of a contagious chronic disease

$$\begin{aligned} \dot{S} &= (1-p)b(S+I+R) - \beta S \frac{I}{S+I+R} - dS \\ \dot{I} &= \beta S \frac{I}{S+I+R} - dI \\ \dot{R} &= pb(S+I+R) - dR\,. \end{aligned} \tag{4}$$

Here $S$ is the number of susceptibles, $I$ is the number of infectives, and $R$ is the number of immune individuals. All parameters are positive, $b$ is birth rate, $d$ is death rate, $\beta$ is infection rate, and $p$ is the fraction of the population that is vaccinated at birth, $0 < p \leq 1$.

a) How is the fact that the disease is chronic modeled?

**Solution**
There is no removal term of infectives, i.e. the infectives never turn into recovered.

b) In the SIR model introduced in the lectures the total population size was conserved. Show that this is not the case for Eq. (4). Explain which part(s) of Eq. (4) causes this difference.

**Solution**
The change in the total population $N = S + I + R$ is

$$\dot{N} = \dot{S} + \dot{I} + \dot{R} = (b-d)N$$

The difference is that in Eq. (4) births and deaths are included in the terms proportional to $b$ and $d$, while in the SIR model introduced in the lectures, births and deaths are neglected (or perfectly balanced by letting $b = d$). Births and deaths must be included in Eq. (4) to model vaccinations at birth.

c) Consider the case of $b = d$ and find the vaccination fraction $p$ needed for the disease to die out if it starts with a small number of infectives.

**Solution**

When $b = d$, then $N = S+I+R$ is constant, and the equations simplify

$$\dot{S} = (1-p)bN - \beta\frac{SI}{N} - bS$$
$$\dot{I} = \beta\frac{SI}{N} - bI$$

with $R = N_0 - S - I$.

The disease dies out if the disease-free steady state $(I^*, S^*) = (0, 1-p)N$ is stable. The stability matrix becomes

$$\mathbb{J}(I^*, S^*) = \begin{pmatrix} -b & -\beta(1-p) \\ 0 & \beta(1-p) - b \end{pmatrix}$$
$$\mathrm{tr}\mathbb{J}(I^*, S^*) = \beta(1-p) - 2b$$
$$\det\mathbb{J}(I^*, S^*) = b(b - \beta(1-p))$$

We have stability if $\mathrm{tr}\mathbb{J}(I^*, S^*) < 0$ and $\det\mathbb{J}(I^*, S^*) > 0$. The first condition, $2b > \beta(1-p)$ is satisfied if the second condition $(b > \beta(1-p))$ is satisfied. The condition on $p$ is therefore $p > 1 - \frac{b}{\beta}$.

Altenatively, the condition can be obtained from solving $\dot{I} < 0$ at the fixed point $(I \approx 0)$:

$$\dot{I} = \beta\frac{SI}{N} - bI = [\beta(1-p) - b]I < 0$$

d) Modify the model (4) so that, instead of being vaccinated at birth, each susceptible chooses to vaccinate itself with a rate $P$. Consider the case $b = d$, and find a condition on $P$ such that the disease dies out if it starts with a small number of infectives.

**Solution**

If each susceptible vaccinates at a rate $P$, then the total removal rate of susceptibles is $-PS$. The modified model becomes

$$\dot{S} = b(S+I+R) - \beta S\frac{I}{S+I+R} - dS - PS$$
$$\dot{I} = \beta S\frac{I}{S+I+R} - dI$$
$$\dot{R} = -dR + PS\,.$$

The population is conserved when $b = d$, giving

$$\dot{S} = bN - \beta\frac{SI}{N} - bS - PS$$
$$\dot{I} = \beta\frac{SI}{N} - bI$$

The disease-free steady state is $(I^*, S^*) = (0, \frac{b}{b+P})N$. The stability matrix becomes

$$\mathbb{J}(I^*, S^*) = \begin{pmatrix} -b-P & -\beta\frac{b}{b+P} \\ 0 & \beta\frac{b}{b+P} - b \end{pmatrix}$$
$$\mathrm{tr}\mathbb{J}(I^*, S^*) = -(b+P) - [b - \beta\tfrac{b}{b+P}]$$
$$\det\mathbb{J}(I^*, S^*) = (b - \beta\tfrac{b}{b+P})(b+P)$$

Here $\det\mathbb{J}(I^*, S^*) > 0$ if $b - \beta\frac{b}{b+P} > 0$, which implies that $\mathrm{tr}\mathbb{J}(I^*, S^*) < 0$. The condition on $P$ becomes $P > \beta - b$.

Altenatively, by solving $\dot{I} < 0$ at the fixed point:

$$\dot{I} = \beta\frac{SI}{N} - bI = [\frac{\beta}{b+P} - 1]bI < 0$$

e) Compare your results in subtasks c) and d), starting from the same initial population size. Which method requires the smallest vaccination rate to keep the disease at low levels in the long run?

**Solution**

The instantaneous vaccination rate is $V_1 = bpN$ for the first method and $V_2 = PS$ for the second one. In the long run, the susceptibles in the second case approaches the fixed point value $V_2 = PS^* = \frac{bP}{b+P}N$.

The ratio of the vaccination rates is

$$\frac{V_1}{V_2} = \frac{p}{P}(b+P)\,.$$

The minimal vaccination rate for the two methods is obtained by inserting the minimal levels $p = (\beta - b)/\beta$ and $P = \beta - b$ (assuming $b < \beta$, otherwise vaccination is not needed) from subtasks c) and d)

$$\frac{V_1}{V_2} = 1\,.$$

Thus, the two methods have the same vaccination rate (although they have different ratio between number of susceptibles and recovered).

**4. Model for formation of fingerprints [8 points]** Recently, researchers identified two proteins that create fingerprints during human gestation. Ridges are formed by one protein promoting rising of the skin, while it is suppressed by the other protein. The interaction between the proteins is modeled by

$$\begin{aligned} \frac{\partial u}{\partial t} &= r\left(a + \frac{u^2}{v(1+Ku^2)}\right) - bu + \boldsymbol{\nabla}^2 u \\ \frac{\partial v}{\partial t} &= u^2 - cv - d\boldsymbol{\nabla}^2 v \end{aligned} \tag{5}$$

Here $u$ and $v$ are concentrations of the proteins. All parameters are positive: $r$ is a growth rate, $ra$ is a basal production, $b$ and $c$ are degradation rates, $d > 1$ is the ratio of diffusion coefficients, and $0 < K \ll 1$ is a small parameter.

a) Inspect Eq. (5) to determine which of $u$ and $v$ is the activator and inhibitor respectively. Motivate your answer.

**Solution**
Increased concentration of $v$ reduces the growth of both $u$ and $v$, while increased concentration of $u$ increases growth of $v$ and of $u$ for a range of values. In conclusion, $u$ is the activator and $v$ is the inhibitor.

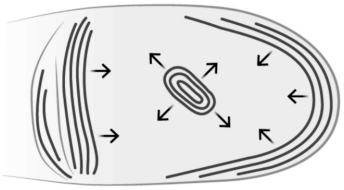
b) The parameter $K$ is introduced to stabilize the dynamics. Qualitatively explain how the dynamics is affected by the parameter $K$ (you do not need to do any calculations here).

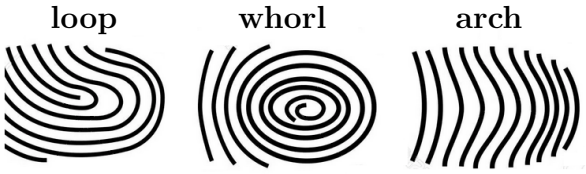
**Solution**
Since $K$ is small, the factor $1 + Ku^2$ is only significantly different from unity if $u$ is very large, larger than $1/\sqrt{K}$. This saturates the self-activation at high activator levels.

c) Consider the case $K = 0$ and show that the system (5) has a single homogeneous steady state $(u^*, v^*)$.

**Solution**
We search for homogeneous steady states by neglecting the spatial dependence in $u$ and $v$. The second equation gives $v^* = [u^*]^2/c$, which inserted in the first equation with $K = 0$ gives $0 = r\big(a+c\big) - bu^*$, with solution $u^* = \frac{r}{b}(a + c)$.

The sketch below shows an example of the early stage of the pattern formation on a fingertip and three typical patterns (loop, whorl and arch):






Ridges spread as waves from two initiation sites at the boundaries, and one initiation site at the center of the finger tip. The propagation and meeting of these waves determine the final fingerprint pattern. In particular, small initial fluctuations of the concentration at the center of the finger tip are important. Assume that this concentration initially is on the form

$$u(\boldsymbol{x}, t = 0) = u^*(1 + 0.1\exp[-\rho(\boldsymbol{x})^2])$$
$$v(\boldsymbol{x}, t = 0) = v^*$$

Here $\boldsymbol{x}$ is taken on the domain of the finger tip, with $\boldsymbol{x} = 0$ at the center of the finger tip. Moreover, $(u^*, v^*)$ denotes the homogeneous steady state in subtask c), assumed to be stable, and $\rho(\boldsymbol{x})$ is a scalar function of space.

d) Below, three examples of $\rho(\boldsymbol{x})$ are given

i. $\rho(\boldsymbol{x}) = \infty$

ii. $\rho(\boldsymbol{x}) = |\boldsymbol{x}|$

iii. $\rho(\boldsymbol{x}) = (|x+y| + 0.01)|\boldsymbol{x}|$

Discuss and sketch the resulting fingerprint pattern you expect for the cases i.–iii. Do the fingerprints show loop, whorl or arch patterns?

**Solution**

The initial condition $u(\boldsymbol{x}, t=0) = u^*(1 + 0.1\exp[-\rho(\boldsymbol{x})^2])$ is a small perturbation to the homogeneous steady state that decays the larger $\rho$ is. This decay localizes the perturbation close to $\rho = 0$.

For case i., $\rho(\boldsymbol{x})$, the perturbation is simply zero, making the central region homogeneous. The fingerprint grows from the boundaries, creating an arch-like structure.

For case ii., the initial condition is a dot-like region. This perturbation is expected to give rise to waves growing out symmetrically, which could give a pattern that is similar to the whorl.

For case iii., the initial condition is an elongated dot-like region, extended in the direction $y = -x$. This spatial extension may allow for a ridge escaping the trapping of the ridges from the boundaries, forming a pattern similar to the loop.

**5. Wright-Fisher model [6 points]** In the lectures, the steady-state population homozygosity for the haploid Wright-Fisher model was derived assuming the infinite-alleles model. For a large population and small mutation rate, the result is

$$F_2 = \frac{1}{1+\theta}\,. \tag{6}$$

Here $\theta = 2N\mu$, where $N$ is the number of haploid individuals in the population, and $\mu$ is the mutation rate per individual per generation. The infinite-alleles model assumes that a mutation always creates a new allelic type never encountered before.
Now consider a population of $N$ haploid individuals with only two possible allelic types, $A$ and $a$. Assume that mutations change $A$ to $a$ (or $a$ to $A$) with mutation rate $\mu$:

$$A \overset{\mu}{\longleftrightarrow} a\,. \tag{7}$$

a) Derive a formula for the steady-state population homozygosity for the case in Eq. (7).
*Hint*: consider how the population homozygosity $F_2^{(t+1)}$ in generation $t+1$ depends on $F_2^{(t)}$, solve for the steady state of this recursion, and take the limits $N \to \infty$ and $\mu \to 0$ keeping $\theta$ constant.

**Solution**
The recursion relation becomes

$$F_2^{(t+1)} = (1-\mu)^2\left[\frac{1}{N} + \left(1 - \frac{1}{N}\right)F_2^{(t)}\right] + 2\mu(1-\mu)\left(1-\frac{1}{N}\right)(1-F_2^{(t)})\,.$$

The first term is the same as for the case discussed in the lectures, i.e. the factor $(1-\mu)^2$ is the probability that no mutation occurred for either of the selected individuals, the first term in the brackets is the probability to pick the same individual, the second term is the probability to pick different individuals times the previous homozygosity of the previous generation. The second term gives the change in $F_2$ if one of the individuals has a mutation and the other not. The probability for this is $\mu(1-\mu)$ times two to count that either of the two individuals could have the mutation. This is multiplied by the probability that different individuals are chosen (otherwise they would have the same number of mutations) times the probability that they were different at the previous generation. The contribution for both individuals to have a mutation can be neglected because $\mu^2$ is negligible compared to $\mu$ as $\mu \to 0$.

Multiplying the equation for $F_2$ by $N$, neglecting $\mu^2$ and solving for the fixed point $F_2^{(t)} = F_2^*$ gives

$$\begin{aligned} NF_2^* &= 1 + (N-1)F_2^* + 2\mu N\left[-\frac{1}{N} - \left(1-\frac{1}{N}\right)F_2^* + \left(1-\frac{1}{N}\right)(1-F_2^*)\right] \\ &\approx 1 + (N-1)F_2^* + \theta\left[0 - F_2^* + (1-F_2^*)\right] \\ &\Rightarrow 0 = 1 - F_2^* + \theta\left[1 - 2F_2^*\right] \\ &\Rightarrow F_2^* = \frac{1+\theta}{1+2\theta} \end{aligned}$$

where the limit $N \to \infty$ was taken with constant $\theta$.

b) Compare the expression you obtained in subtask a) to Eq. (6). Which one is larger? Why? Analyse how the difference between the two expressions behaves in the limit of very small but non-zero $\theta$, and in the limit of very large $\theta$. Explain the two behaviours.

**Solution**

The two expressions are equal only if

$$\frac{1}{1+\theta} = \frac{1+\theta}{1+2\theta} \;\Rightarrow\; 1+2\theta = (1+\theta)^2 \;\Rightarrow\; \theta = 0$$

This implies that the second expression is larger, because it is larger when $\theta \to \infty$. This is reasonable because mutations contribute positively to the homozygosity.

For small $\theta$, both cases behave as $F_2^* \sim 1-\theta$ When there are no mutations, the homozygosity approaches unity as expected due to fixation. The first-order correction in $\theta$ is the same in the two cases. This is because the factor $1-F_2$ in the contribution due to mutations is small when homozygosity is high.

In the limit of large $\theta$, the first case approaches $F_2^* = 0$, while the second case approaches $F_2^* = 1/2$. When mutations are very frequent, no alleles are the same in the infinite alleles model, while the population is split randomly between the two types in the second case.