CHALMERS, GÖTEBORGS UNIVERSITET

EXAM for
COMPUTATIONAL BIOLOGY

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

Time:	June 10, 2022, 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) For what population sizes do you expect the Malthus growth model to be most suitable? Why is the form of the Malthus growth model often valid for these population sizes, also when additional factors (crowding, competition for resources, etc) are taken into account?

b) Explain how to draw a cobweb plot for a map $N_{\tau+1} = F(N_\tau)$ and explain why this is the right procedure to obtain the solution orbit N_1, N_2, ..., starting from N_0.

c) Explain the principle of competitive exclusion and give an example of a possible application of it.

d) Assume that ammonia (NH_3) is spilled on the floor two meters from where you stand. Estimate the time scale until a significant amount of NH_3 molecules reach your nose, if they are solely transported by three-dimensional molecular diffusion through air with diffusion coefficient $2 \cdot 10^{-5}\,\mathrm{m}^2/\mathrm{s}$. Comment on the result, is it realistic?

e) Fisher's equation in one spatial dimension is given by

$$\frac{\partial}{\partial t} n(x,t) = r n(x,t) \left(1 - \frac{n(x,t)}{K}\right) + D \frac{\partial^2}{\partial x^2} n(x,t)\,.$$

Explain how a travelling wave solution to Fisher's equation can travel much faster into an initially empty region than the solution to the diffusion equation (obtained by letting $r = 0$ in Fisher's equation).

f) The Kuramoto model consists of N oscillators with phases θ_i and natural angular frequencies ω_i

$$\dot{\theta}_i = \omega_i + \frac{K}{N} \sum_{j=1}^{N} \sin(\theta_j - \theta_i) \,.$$

Explain how a fraction of these oscillators can be synchronized despite having different angular frequencies.

g) Consider the stochastic SIS model for disease spreading. Let λ_n denote the rate of infection and μ_n the rate of recovery for n infectives. Write down a Master equation for the distribution $\rho_n(t)$ for observing n infectives at time t. Assume that each infection event results in at most one individual becoming infected.

h) Discuss and contrast the effects of measurement noise and dynamical noise on a time series which is generated by linear (Malthus) decay (negative growth rate).

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

$$\dot{N} = -dN(t) + bN(t-T)(k - bsN(t-T)) \,. \tag{1}$$

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 (1). What is the significance of the parameters k, s and T?

b) Find all steady states ($N(t) = \text{const.}$) of the Eq. (1). 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.

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 η can be approximated by the form

$$\frac{\mathrm{d}\eta}{\mathrm{d}t} \approx C_1\eta(t) + C_2\eta(t-T) \,. \tag{2}$$

What are the expressions for the coefficients C_1 and C_2?

d) Use the ansatz $\eta(t) = Ae^{\lambda t}$ in Eq. (2) to derive an equation for λ.

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?

3. Effect of spruce budworms on a forest [10 points] In the lectures the spruce budworm model was introduced for the fast growth of budworms in a forest under predation. To instead make a model for the effect of budworms feeding on the foliage (i.e. leaves and branches) of the forest, assume a constant budworm population of size B at one of its stable steady states: either outbreak (large B) or refuge (small B). Let $S(t)$ denote the average surface area of the branches of a tree and let $H(t)$ denote a general measure of the forest's health. A growth model for S and H is

$$\begin{aligned} \dot{S} &= r_S S\left(1 - \frac{S}{K_S}\frac{K_H}{H}\right) \\ \dot{H} &= r_H H\left(1 - \frac{H}{K_H}\right) - P\frac{B}{S} \end{aligned} \tag{3}$$

where r_S, r_H, K_S, K_H and P are positive parameters.

a) Interpret the terms in Eq. (3) biologically.

b) Introduce dimensionless units to write the system (3) in as few dimensionless parameters as possible.

c) Show, for example geometrically, that the system has two positive fixed points if B is small and no positive ones if B is large.

d) Discuss how the spruce budworm population affects the forest according to the model. Discuss whether the results are realistic.

4. Diffusion driven instability [8 points] Consider the following reaction-diffusion equation system in one spatial dimension for a version of the SI model of suspeptibles $S(x,t)$ and infectives $I(x,t)$

$$\begin{aligned} \frac{\partial S}{\partial t} &= b - d_S S - \beta SI + D_S \frac{\partial^2 S}{\partial x^2} \\ \frac{\partial I}{\partial t} &= \beta SI - d_I I + D_I \frac{\partial^2 I}{\partial x^2} \end{aligned} , \tag{4}$$

where b, d_S, d_I, β, D_S and D_I are positive parameters. For simplicity you can let $d_S = d_I = b = D_S = 1$ throughout this problem (corresponding to a suitable dedimensionalisation that you do not need to derive).

a) Show that if $b > d_I d_S/\beta = 1$, the system (4) has a positive homogeneous steady state which is stable.

In the lectures, we showed that small perturbations $\delta u(x,t) \equiv u(x,t) - u^*$ and $\delta v(x,t) \equiv v(x,t) - v^*$ from a stable homogeneous steady state (u^*, v^*) for the concentrations u and v following a generic reaction-diffusion system, can be decomposed using the ansatz

$$\begin{pmatrix} \delta u \\ \delta v \end{pmatrix} = e^{\lambda t + \mathrm{i}kx} \begin{pmatrix} \delta u_0 \\ \delta v_0 \end{pmatrix}$$

yielding the following equation:

$$0 = [\lambda - \mathbb{K}] \begin{pmatrix} \delta u_0 \\ \delta v_0 \end{pmatrix} , \text{ where } \mathbb{K} = \mathbb{J}(u^*, v^*) - k^2 \begin{pmatrix} D_u & 0 \\ 0 & D_v \end{pmatrix} .$$

Here D_u and D_v are the diffusion coefficients of u and v respectively.

b) Does the system (4) have a diffusion-driven instability for any parameter values $b_c(k)$ for which space-dependent perturbations with a single wave number k becomes unstable?

c) Assume a system with a diffusion-driven instability that is unstable to perturbations with wave numbers in a range $k_- \leq k \leq k_+$. Initially, the concentration is obtained by a small random perturbation from the homogeneous stable steady state. The domain is one-dimensional with length L and no-flux boundary conditions. What is the minimal length L required to have a diffusion-driven instability in this system?

5. Coalescent process [8 points]

a) The coalescent process is a model for neutral sample genealogies, consistent with the Fisher-Wright model. Describe the coalescent process in its simplest form, for a sample of size n from a large population, $N \gg n$, and derive the following distribution of the time T_j to the next coalescent event, given that there are j ancestral lines:

$$P(T_j) = \lambda_j \exp(-\lambda_j T_j) \quad \text{with} \quad \lambda_j = \frac{1}{N}\binom{j}{2} . \tag{5}$$

b) Tajima suggested a test for selection by comparing whether a genetic mosaic is compatible with a neutral sample genealogy, or not. The test is based upon two different estimators for the mutation parameter $\theta = 2N\mu$ that are derived from the following equations

$$\langle S_n \rangle = \theta \sum_{j=1}^{n-1} \frac{1}{j} \quad \text{and} \quad \left\langle \frac{1}{\binom{n}{2}} \sum_{i<j} \Delta_{ij} \right\rangle = \theta . \tag{6}$$

Here $\langle S_n \rangle$ is the average number of single-nucleotide polymorphisms (SNPs) in the sample of size n, and Δ_{ij} is the number of SNPs between two individuals in the sample, i and j. Derive the two relations in Eq. (6) using the coalescent process.
Hint: For the first relation, use that the number S_n of SNPs in a given genealogy for n individuals is Poisson distributed,

$$P(S_n = j) = \frac{(\mu T_{\rm tot}^{(n)})^j}{j!} \exp(-\mu T_{\rm tot}^{(n)}) ,$$

where $T_{\rm tot}^{(n)}$ is the total branch length of the genealogy. Compute the expected number of SNPs, and then average over genealogies. For the second relation, compute $\langle \Delta_{ij} \rangle$ by considering $n = 2$.