CHALMERS, GÖTEBORGS UNIVERSITET

EXAM for
COMPUTATIONAL BIOLOGY

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

Time:	August 24, 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) Discuss similarities and differences in the solutions of the continuous logistic equation and its discrete counterpart, the logistic map, for small and large values of the growth rate r.

b) Consider the two growth equations

$$\frac{\mathrm{d}N}{\mathrm{d}t}(t) = rN(t) \quad \text{and} \quad \frac{\mathrm{d}N}{\mathrm{d}t}(t) = rN(t-T)\,.$$

with $r > 0$ and $T > 0$. Both equations can be solved using the following ansatz $N(t) = \sum_i A_i e^{\lambda_i t}$. Without doing any calculations, explain the difference in the spectrum of allowed λ_i for the two equations and what this difference implies in terms of oscillations of the solutions.

c) Assume a model for two competing species of sizes u_1 and u_2 with identical growth rate, carrying capacity and interaction rate $a > 1$

$$\begin{aligned}\dot{u}_1 &= u_1\,(1 - u_1 - au_2)\\ \dot{u}_2 &= u_2\,(1 - u_2 - au_1)\end{aligned}$$

Does the principle of competitive exclusion imply that both species survive in the long run since neither species has an advantage?

d) What does the law of mass action state? Explain the form of the law of mass action.

e) Discuss similarities and differences between molecular diffusion and turbulent diffusion.

f) The continuity equation was derived in the lectures. It reads

$$\frac{\partial}{\partial t} n(\boldsymbol{x}, t) = f(\boldsymbol{x}, t) - \boldsymbol{\nabla} \cdot \boldsymbol{j}(\boldsymbol{x}, t) .$$

Explain the difference in the origin of the two right-hand side terms. Give examples of what f and $\boldsymbol{j}$ could represent in a biological system.

g) Explain what a morphogen is and what role it has in diffusion-driven instability.

h) Explain the difference between epidemic and endemic diseases.

2. Nutrient cycling in a lake [10 points] The following is a simple model for the concentration n of a nutrient (phosphorus) in a lake

$$\dot{n} = L - Sn + \frac{Rn^2}{M^2 + n^2} . \tag{1}$$

Here L, S, R and M are positive parameters. The first term L models nutrient input from the surrounding due to natural sources and human sources such as agriculture. The second term $-Sn$ models loss due to sedimentation and outflow (uptake from plants and animals is assumed to be a much smaller contribution). The last term models nutrient recycling in the lake due to emission from sediments or decomposition of dead plants and animals.

a) Explain why the functional form is different in the first two terms on the right-hand side in Eq. (1).

b) Introduce dimensionless units to write Eq. (1) in terms of as few dimensionless parameters as possible.

Lakes are commonly classified to be either oligotrophic (low level of nutrients, clear blue water) or eutrophic (high level of nutrients, high biological production with frequent algal blooms). In what follows, consider the case $R = M = 1$, $S = 1/2$ and general values of L.

c) Using a geometric approach, show that there exist a range of L where the nutrient concentration has stable steady states at two significantly different levels, corresponding to oligotrophic and eutrophic lakes.

d) Assume a lake that is initially in the oligotrophic state. Use a geometric approach to show that the lake can change drastically to an eutrophic state due to an increased use of agricultural fertilizers that slowly increase L to a higher level.

e) Starting from the eutrophic state in subtask d), is it, within the model, possible to revert the lake back to the oligotrophic state by slowly reducing the use of agricultural fertilizers?
Hint It may be instructive to analytically analyze the case when $L = 0$.

3. Deterministic and stochastic SIS models [10 points] Consider the SIS model for the evolution of S susceptibles and I infectives

$$\begin{aligned} \dot{S} &= -\beta \frac{SI}{S+I} + \gamma I \\ \dot{I} &= \beta \frac{SI}{S+I} - \gamma I \end{aligned} \tag{2}$$

with positive parameters β and γ.

a) Explain the form of the different terms in Eq. (2).

b) Use that the total population is conserved in Eq. (2) to derive an equation for $\dot{S}$ in terms of S.

c) Evaluate the fixed points of the dynamics in subtask b), and determine their stability. What is the long-term state of the system (2)?

d) Formulate a stochastic version (Master equation) of the SIS model for the probability distribution $\rho_S(t)$ to have S susceptibles at time t.

e) Discuss qualitative differences in the solutions after long times in the deterministic and stochastic models.

4. The Kuramoto model [10 points] Consider a number N of coupled oscillators with phases $\theta_1, \theta_2, \ldots \theta_N$ with the following time evolution

$$\dot{\theta}_i = \omega_i + \frac{K}{N} \sum_{j=1}^{N} \sin(\theta_j - \theta_i)\,. \tag{3}$$

Here ω_i are constant natural angular velocities of the oscillators and $K > 0$ is constant. Define the complex order parameter as

$$r(t) e^{\mathrm{i}\,\psi(t)} = \frac{1}{N} \sum_{j=1}^{N} e^{\mathrm{i}\,\theta_j(t)}\,, \tag{4}$$

where r and ψ are real.

a) Commonly, oscillators are described using a second order differential equation for θ. The oscillators in Eq. (3) have only the first time derivative of θ. Why are there no second order time derivatives in Eq. (3)?

b) Show that Eq. (3) can be rewritten using the order parameter as

$$\dot{\theta}_i = \omega_i + K r(t) \sin(\psi(t) - \theta_i(t))\,.$$

c) Assume that for large times $r(t) = \text{const.}$ and $\psi(t) = \Omega t$, with constant Ω. Show that the dynamics in subtask a) takes the following form in a frame rotating with angular velocity Ω:

$$\dot{\theta}'_i = \omega'_i - K r \sin(\theta'_i)\,,$$

where $\theta'_i(t)$ is the phase in the rotating frame. What is the form of ω'_i?

d) Consider the case where the natural angular velocities of all oscillators in the rotating frame have the same magnitude, $|\omega_i| = \omega$, but half are positive, $\omega_i' = +\omega$ for $i = 1, \ldots, N/2$ and half are negative, $\omega_i' = -\omega$ for $i = N/2+1, \ldots, N$ (assume N even). What is the long-term dynamics in this case? Does it depend on the parameters?

e) For the case in subtask d), use the definition (4) to find an expression of the order parameter r in the steady state in terms of the parameters K and ω_0. Does at least one solution exist for all parameter values? If not, explain what happens.

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,

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

where $T_{\text{tot}}^{(n)}$ 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 $n = 2$.