CHALMERS, GÖTEBORGS UNIVERSITET

EXAM for
COMPUTATIONAL BIOLOGY

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

Time:	March 17, 2022, at $14^{00} - 18^{00}$
Place:	Johanneberg
Teachers:	Kristian Gustafsson, 070-050 2211 (mobile), visits once around 15^{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) Give two examples of biological systems, one where a time delay model is a suitable model and one where a discrete growth model is suitable.

b) In the Lotka-Volterra model for prey and predator populations N and P

$$\dot{N} = N(a - bP)$$
$$\dot{P} = P(cN - d)$$

predators have limitless appetite. Explain what this means and how one can model a limited appetite.

c) Write down a growth model for the concentrations in the spontaneous reaction from substrate S to product P with rate constant k, $\mathrm{S} \xrightarrow{\mathrm{k}} \mathrm{P}$.

d) Give examples of three distinct stochastic mechanisms or effects that we have described using stochastic models in the course.

e) Consider a striped snake skin pattern created by diffusion driven instability of morphogen concentrations. Do you expect the stripes to run from head to tail or like rings around the body? Motivate your answer.

f) Explain the mechanism for wave patterns in unstirred concentrations undergoing the Belousov-Zhabotinsky reaction in a vertical tube.

g) Explain what is meant by the phase of an oscillator. What does it mean that several oscillators are phase locked in the Kuramoto model?

h) The autocorrelation function of a time series $x_0, x_1 \dots x_{N-1}$ is defined by

$$C(k) = \frac{\sum_{n=0}^{N-k-1}(x_n - \langle x \rangle)(x_{n+k} - \langle x \rangle)}{\sum_{n=0}^{N-1}(x_n - \langle x \rangle)^2} \,.$$

Explain how $C(k)$ can be used to detect periodic oscillations in data.

2. Continuous VS discrete fishery management models [8 points] Consider the following continuous growth model for how a fish population of $N(t)$ individuals at time t is affected by fishing

$$\dot{N} = rN\left(1 - \frac{N}{K}\right) - EN \,. \tag{1}$$

Here E is a fishing rate with $0 \leq E < r$. r and K are positive parameters denoting the growth rate and carrying capacity in absence of fishing.

a) By rearranging the terms, show that Eq. (1) can be rewritten as a continuous logistic growth model, $\dot{N} = r_E N(1 - N/K_E)$, with effective growth rate r_E and carrying capacity K_E. Express r_E and K_E in terms of the original parameters r, K and E. Verify that $r_E > 0$.

b) Using the result in subtask a), write down the stable steady state N^* of Eq. (1). Define the yield as the removal rate of the population due to fishing. Which fishing rate E gives the maximal yield in the steady state of the model (1)? What is the corresponding maximal yield?

If generations of the fish population do not overlap, it may be better to consider the following discrete growth model for fishing

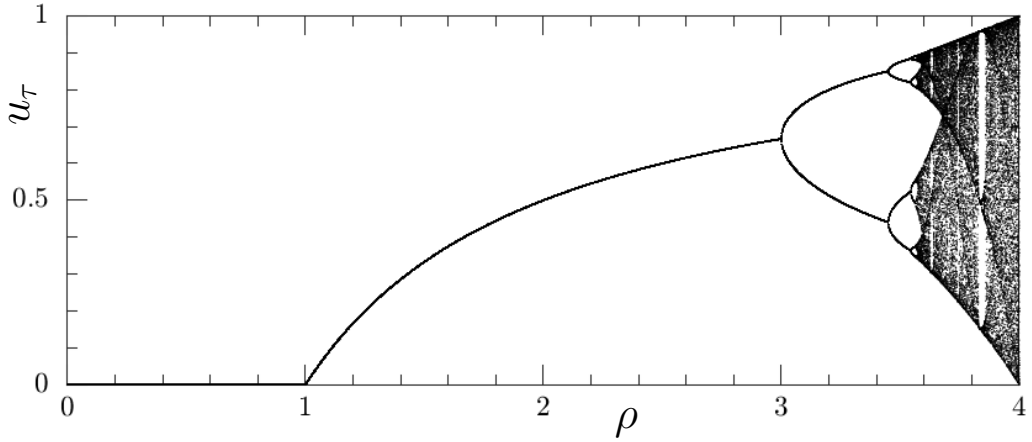
$$N_{\tau+1} = N_\tau + rN_\tau\left(1 - \frac{N_\tau}{K}\right) - EN_\tau \,. \tag{2}$$

Assume that the time units are scaled so that r and E takes the same values in Eqs. (1) and (2).

c) Verify that the steady state N^* in subtask b) is a steady state also of Eq. (2). Contrast the range of stability of N^* in the two models.
Hint: To simplify the calculation, you can rewrite Eq. (2) in terms of the parameters r_E and K_E of subtask a).

Find subtask d) on next page

d) The plot below shows the bifurcation diagram for the dimensionless logistic map $u_{\tau+1} = \rho u_\tau(1-u_\tau)$ with population size u and parameter ρ.



Verify that the range of stability in subtask c) is consistent with this diagram. What happens outside the range of stability of the steady state N^*? Discuss whether there are situations when fishing can have a beneficial effect on the population dynamics according to the model (2).

3. Travelling waves [10 points] The following is a model for spread of a population with concentration $n(x,t)$ in a one-dimensional habitat:

$$\frac{\partial n}{\partial t} = rn\left(1-\frac{n}{K}\right)\left(\frac{n}{A}-1\right) + D\frac{\partial^2 n}{\partial x^2}\,. \tag{3}$$

Here, r, K, A, and D are positive constant parameters. Assume that $A < K$.

a) Give an explanation of the first term $rn\left(1-\frac{n}{K}\right)\left(\frac{n}{A}-1\right)$ in Eq. (3). What does K and A signify?

b) Introduce dimensionless time τ, position ξ and population size u and rewrite the dynamics (3) in terms of these coordinates and a single dimensionless parameter of your choice.

c) Assume that u only depends on ξ and τ through the combination $z = \xi - c\tau$ and convert the partial differential equation in subtask b) to a dynamical system for u and $v = \frac{\mathrm{d}u}{\mathrm{d}z}$.
Hint: If you did not solve subtask b), you can instead use Eq. (3) with $r = K = D = 1$ and parameter A.

d) Find all fixed points in the system derived in subtask c) and determine their stability if $c < 0$ (**OBS: Negative sign of c**).

e) Sketch the shapes of possible travelling wave solutions with $c < 0$ (**OBS: Negative sign of c**) in the phase plane for u and v. Sketch the corresponding wave profiles $n(x,t)$ against t. In which direction does the wave travel?

4. Disease spreading of two variants of a disease [10 points] The SIRS model describes the spread of a non-lethal disease where immunity only lasts for a finite time:

$$\begin{aligned} \dot{s} &= -\beta si + \gamma r \\ \dot{i} &= \beta si - \alpha i \\ \dot{r} &= \alpha i - \gamma r \,. \end{aligned} \tag{4}$$

Here s, i and r denote the fractions of the total population that is susceptible (s), infective (i), or recovered (r). The parameters α, β and γ are positive.

a) Explain the meaning of the time scales associated with the inverse parameters: $1/\alpha$, $1/\beta$ and $1/\gamma$.

b) Verify that $s+i+r$ is conserved in Eq. (4), and use this fact to remove r from the equations, leaving two equations for s and i.

c) Show that this model has two steady states, one disease free and one endemic (you do not need to evaluate their stability). Give a condition for the endemic steady state to exist.

Assume that the SIRS model (4) successfully describes the spread of a disease, for example the Delta variant of COVID-19. Now assume a new variant of the disease forms, for example Omicron. Assume that the new variant gives similar symptoms (same α and γ), but it is more infectious, and each variant only gives partial immunity to the other variant.
To simplify, assume further:

- One individual can at most be infected by one variant at a time
- Individuals do not die, when they recover they become immune to the variant they were infected by and partially immune to the other variant.
- Both immunity and partial immunity is lost at the same rate.
- If a partially immune individual gets infected, it loses all immunity (i.e. it behaves as an infected susceptible).

d) Extend and modify the SIRS model (4) to model the spread of two variants of a disease in a population following the assumptions above.

e) Does your model in subtask d) have a steady state where one of the variants disappears due to the competition (you do not need to evaluate the stability of this steady state)? What is the form of the dynamics for the remaining disease?

5. Coalescent process [10 points]

a) **[5 points]** 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) **[3 points]** 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

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

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 $n = 2$.

c) **[2 points]** The two estimators

$$\hat{\theta}_1 = \Big(\sum_{j=1}^{n-1} \frac{1}{j} \Big)^{-1} S_n \quad \text{and} \quad \hat{\theta}_2 = \frac{1}{\binom{n}{2}} \sum_{i<j} \Delta_{ij}$$

have the same average for neutral genealogies. Explain qualitatively in a couple of sentences why the estimators tend to have different averages when selection is important.
Hint: Remember how selection tends to change genealogies.