Solutions/answers to selected problems of the exam 16:th of March in Computational Biology A 2017

1. Pest control [2 points] Assume that a population of size $N(t)$ of fertile insects shows the following time evolution:

$$\dot{N} = bN - dN - cN^2\,, \tag{1}$$

where b, c and d are positive parameters and $b > d$.

a) Explain the form of Eq. (1). Give plausible interpretations for the parameters b, d, and c. How large is the carrying capacity of the environment?

Solution
The first and second term are describe changes due to linear birth of rate b and linear death of rate d. The third term is a higher-order death term proportional to the population size. It limits the growth rate due to finite resources in the system. There is no corresponding birth term, i.e. it is assumed that the birth rate is not affected by the population size.

The carrying capacity K is given by the population size at the non-zero fixed point, $N^* = (b-d)/c$, of the system: $K = (b-d)/c$

b) One method for pest control of insects is to release a number of sterile insects into a population. Assume that a population of sterile insects is introduced into a population of N fertile insects. The population size s of sterile insects is kept constant by a steady supply of new sterile insects balancing deaths.

Assume that the sterile insects show identical behaviour as fertile insects (equal mating rate and equal competition for resources), with the only exception that mating involving sterile insects results in failed births. Assume moreover that the sterile insects are male or female with equal probability, i.e. you do not need to take the sex of the insects into consideration.

Modify Eq. (1) to model how a number s of sterile insects affect the time evolution of N.

Solution
Eq. (1) must be modified in two ways. First, the birth rate is modified to take into account that only a fraction $N/(N+s)$ of births results in offsprings. Second, we need to add the population size of sterile insects to the death rate of N due to competition for resources. This gives the modified growth dynamics:

$$\dot{N} = \frac{N}{N+s}bN - dN - cN(N+s)\,. \tag{2}$$

c) Show that the ratio $\rho = N/(N+s)$ satisfies the following dynamics

$$\dot{\rho} = \rho((b\rho - d)(1-\rho) - cs)\,.$$

Solution
We have

$$\begin{aligned}\dot{\rho} &= \frac{\mathrm{d}}{\mathrm{d}t}\frac{N}{N+s} = \frac{\dot{N}}{N+s} - \frac{N\dot{N}}{(N+s)^2} = \frac{\dot{N}}{N}\rho(1-\rho)\\ &= [b\rho - d - c(N+s)]\,\rho(1-\rho)\\ &= [(b\rho-d)(1-\rho) - cs]\,\rho\,.\end{aligned}$$

where we used Eq. (2) and that $N+s = -s/(\rho-1)$.

d) Assume that $d=0$, $b=1$, and $c=1$. Under this assumption, determine the smallest number of sterile insects, s_c, needed to make the insect population go extinct for any allowed initial value of ρ. **Solution**
The dynamics has the flow

$$f(\rho) = \rho(\rho(1-\rho) - s)\,.$$

We want to find $s = s_\mathrm{c}$ such that $f(\rho) < 0$ for all allowed ρ ($0 \le \rho \le 1$). To this end, find the zeroes of $f(\rho)$: $\rho_1^* = 0$, $\rho_2^* = (1+\sqrt{1-4s})/2$, $\rho_3^* = (1-\sqrt{1-4s})/2$. If $s > 1/4$, there is no crossing and the flow is negative everywhere (can be seen by testing one value, for example $f(\rho = 1/2) = 1/2(1/4 - s) < 0$ if $s > 1/4$). If $s < 1/4$ two fixed points are formed at $\rho = 1/2$ and it is straightforward to check that the flow is positive in between them: $f(\rho = 1/2) = 1/2(1/4 - s) > 0$ when $s < 1/4$.

In conclusion, $s_\mathrm{c} = 1/4$.

As a side remark, in the current units, the carrying capacity is unity, i.e. we need to at least maintain a sterile population of size one quarter of the carrying capacity to make the population go extinct. The time until extinction goes to infinity as s approaches $1/4$ from above (c.f. ghost of a saddle point).

Alternative solution 1 The fixed point at $\rho = 0$ has $f'(0) = -s$, i.e. is always stable. The other fixed points do not exist when $s < 1/4 \Rightarrow s_\mathrm{c} = 1/4$.

Alternative solution 2 Flow is negative, $\dot{\rho} < 0$, if $\rho(1-\rho) - s < 0$, i.e. if $s < \rho(1-\rho)$. The curve $\rho(1-\rho)$ has maximum value $1/4$ at $\rho = 1/2 \Rightarrow s_\mathrm{c} = 1/4$.

2. SIRS model [2 points] A simple model for the spreading of influenza is the SIRS model (in contrast to the SIR model discussed in the lectures). The SIRS model has the following dynamics:

$$\begin{aligned}\dot{S} &= -rSI + \gamma R\\ \dot{I} &= rSI - \alpha I\\ \dot{R} &= \alpha I - \gamma R\,.\end{aligned} \tag{3}$$

The different population sizes correspond to susceptibles $S(t)$, infectives $I(t)$, and removed $R(t)$. Assume that the initial population size is $N = S(0)+I(0)$ and that $R(0) = 0$.

a) Give brief, plausible explanations of the different parameters r, α, and γ (all assumed to be positive).

Solution

The parameter r is the rate at which the influenza spreads due to encounters between susceptibles and infectives. The parameter α is the removal rate of infectives to the removed group. The parameter γ is the rate at which removed population becomes susceptible. Due to the parameter γ we expect the influenza to have small mortality rate (it would be strange if dead individuals suddenly became reanimated as susceptibles).

b) What are the differences between Eq. (3) and the SIR model discussed in the lectures?

Solution

The SIRS model has an additional removal term $(-\gamma R)$ of removed species and the same amount $(+\gamma R)$ is added to the susceptibles. This could model a situation where the removed population consists of immune individuals that slowly loose their immunity and become susceptible again with rate γ.

c) Show that Eq. (3) can be written as a two-dimensional system:

$$\begin{aligned}\dot{S} &= f(S, I)\\ \dot{I} &= g(S, I)\,.\end{aligned}$$

Explicitly write down $f(S, I)$ and $g(S, I)$.

Solution

Eq. (3) preserves the population size $\frac{\mathrm{d}}{\mathrm{d}t}[S+I+R] = 0 \Rightarrow S+I+R = N = \text{const.}$. Substitute this expression for R in the equation for S to get

$$\begin{aligned}\dot{S} &= \underbrace{-rSI + \gamma[N-S-I]}_{f(S,I)}\\ \dot{I} &= \underbrace{rSI - \alpha I}_{g(S,I)}\,.\end{aligned}$$

d) The system in subtask c) has two possible fixed points, determine these. Give a condition on the parameters for which the system has two biologically relevant (non-negative I and S) fixed points.

Solution

The system always has a fixed point at $(S_1^*, I_1^*) = (N, 0)$. There is an additional possible fixed point at $(S_2^*, I_2^*) = (\alpha/r, \gamma(Nr-\alpha)/(r(\alpha+\gamma)))$.

The system has two distinct biologically relevant fixed points if $N > \alpha/r$.

e) Determine the stability of the fixed points for the case where the system has one biologically relevant fixed point, and for the case where the system has two biologically relevant fixed points. Discuss the long-term fate of the system for these two cases.

Solution

The Jacobian is

$$\mathbb{J}(S,I) = \begin{pmatrix} -rI-\gamma & -rS-\gamma \\ rI & rS-\alpha \end{pmatrix} .$$

For the first fixed point

$$\begin{aligned} \mathbb{J}(S_1^*, I_1^*) &= \begin{pmatrix} -\gamma & -rN-\gamma \\ 0 & rN-\alpha \end{pmatrix} \\ \mathrm{tr}\mathbb{J}(S_1^*, I_1^*) &= rN-\alpha-\gamma \\ \det \mathbb{J}(S_1^*, I_1^*) &= \gamma(\alpha - rN) . \end{aligned}$$

For the second fixed point (assuming $N > \alpha/r$)

$$\begin{aligned} \mathbb{J}(S_2^*, I_2^*) &= \begin{pmatrix} -\gamma\frac{Nr+\gamma}{\alpha+\gamma} & -\alpha-\gamma \\ \gamma\frac{Nr-\alpha}{\alpha+\gamma} & 0 \end{pmatrix} \\ \mathrm{tr}\mathbb{J}(S_2^*, I_2^*) &= -\gamma\frac{Nr+\gamma}{\alpha+\gamma} < 0 \\ \det \mathbb{J}(S_2^*, I_2^*) &= \gamma(Nr-\alpha) > 0 . \end{aligned}$$

The second fixed point has negative trace and positive determinant and must therefore be stable (node or spiral) when it exists ($N > \alpha/r$).

The eigenvalues of the first fixed point can be determined using

$$\lambda_\pm = \frac{1}{2}\left(\mathrm{tr}\mathbb{J} \pm \sqrt{(\mathrm{tr}\mathbb{J})^2 - 4\det\mathbb{J}}\right)$$

to get (for case $N > \alpha/r$)

$$\begin{aligned} \lambda_- &= \frac{1}{2}\left(rN-\alpha-\gamma-\sqrt{r^2N^2+\alpha^2+\gamma^2-2rN\alpha-2rN\gamma+2\alpha\gamma-4\gamma\alpha+4\gamma rN}\right) \\ &= \frac{1}{2}\left(rN-\alpha-\gamma-\sqrt{(rN-\alpha+\gamma)^2}\right) \\ &= \frac{1}{2}\left(rN-\alpha-\gamma-rN+\alpha-\gamma\right) \\ &= -\gamma < 0 \\ \lambda_+ &= rN-\alpha > 0 , \end{aligned}$$

and to get (for case $N < \alpha/r$)

$$\begin{aligned} \lambda_- &= rN-\alpha < 0 \\ \lambda_+ &= -\gamma < 0 . \end{aligned}$$

In conclusion, when $N < \alpha/r$ the system has a single fixed point that is a stable node with $I = R = 0$ and $S = N$, i.e. the influenza disappears and the full population is healthy. In contrast, for population sizes over the threshold, $N > \alpha/r$, the first fixed point is unstable (saddle point) and the population is drawn to the second, positive fixed point. Thus, the influenza stays in the population for long times with a fraction of infected persons that depend on the parameters and the population size N.

3. Reaction kinetics [2 points] Assume that a chemical reaction proceeds as follows

$$\mathrm{S} + \mathrm{E} \underset{\mathrm{k_{-1}}}{\overset{\mathrm{k_1}}{\rightleftharpoons}} \mathrm{SE} \xrightarrow{\mathrm{k_2}} \mathrm{P} + \mathrm{E}\,, \tag{4}$$

where S denotes a substrate, E an enzyme, and P a product. The parameters k_1, k_{-1} and k_2 are rate constants.

a) What is the law of mass action? What does it assume about the underlying chemical reaction?

Solution
The law of mass action states that reaction rates are proportional to the concentrations of the reactants (encounter rates between chemicals as they diffuse around in a solution). The proportionality coefficients are rate constants of the reaction.

The law of mass action assumes that the reaction is slow and that the chemicals are well mixed.

b) Use the law of mass action together with the reaction in Eq. (4) to set up a dynamical system for the concentrations $s = [S]$, $e = [E]$, $c = [SE]$ and $p = [P]$ of the reactants. Assume appropriate initial conditions, that are determined just before the reaction starts.

Solution
Using the law of mass action Eq. (4) gives the system:

$$\dot{s} = -k_1 es + k_{-1}c\,, \qquad s(0) = s_0 \tag{5}$$

$$\dot{e} = -k_1 es + (k_{-1} + k_2)c\,, \qquad e(0) = e_0 \tag{6}$$

$$\dot{c} = k_1 es - (k_{-1} + k_2)c\,, \qquad c(0) = 0 \tag{7}$$

$$\dot{p} = k_2 c\,, \qquad p(0) = 0\,. \tag{8}$$

c) What is the role of the enzyme E? How is this role reflected in the dynamics you found in subtask b)?

Solution
The role of the enzyme is as a catalyst, i.e. it should speed up the reaction without getting consumed (this property typically implies that it is enough with very small enzyme concentrations).

Without the enzyme, there is no reaction path to the product and the substrate remains at constant concentration in Eq. (8). I.e. the enzyme catalyzes the reaction. The enzyme is found in both concentrations c and e. From the dynamics (8), $\frac{\mathrm{d}}{\mathrm{d}t}[c+e] = 0$, implying that the total amount of enzyme is unchanged. For large times, the system (8) approaches a state where $s = c = 0$ and $e = e_0$, i.e. the enzyme is not consumed.

d) Consider the following reaction between two reactants A and X

$$\mathrm{A} + \mathrm{X} \underset{\mathrm{k}_{-1}}{\overset{\mathrm{k}_1}{\rightleftharpoons}} 2\,\mathrm{X}\,.$$

Assume that A is maintained at constant concentration $a = [A] =$ const. Using the law-of mass action, set up a dynamical system for the concentrations $a = [A]$ and $x = [X]$.

Solution

The dynamical system becomes

$$\dot{x} = k_1 a x - k_{-1} x^2\,, \tag{9}$$

i.e. a system of the form of logistic growth.

e) How would the dynamics change in subtask d) if you remove the assumption that a is maintained at constant concentration, i.e. if the concentration a is only influenced by the reaction and not from any external sources?

Solution

If there is no inflow of reactant A, then the concentration a will change with time. Eq. (9) is modified as

$$\begin{aligned} \dot{x} &= k_1 a x - k_{-1} x^2 \\ \dot{a} &= -k_1 a x + k_{-1} x^2\,. \end{aligned}$$

It follows that $a + x = a_0 + x_0$ is constant. Inserting this relation into the equation for x, we have

$$\dot{x} = k_1(a_0 + x_0)x - (k_{-1} + k_1)x^2\,,$$

i.e. x still satisfies a logistic equation with modified growth rate and carrying capacity compared to the case in subtask d).

4. Macroscopic diffusion [2 points] The diffusion equation in one spatial dimension can be written on the form

$$\frac{\partial n}{\partial t} = D\frac{\partial^2 n}{\partial x^2}\,, \tag{10}$$

where n is a concentration, time t has unit T and position x has unit L.

a) What are the dimensional units of the diffusion coefficient

   i) for the diffusion equation in one spatial dimension?
   ii) for the diffusion equation in two spatial dimensions?
   iii) if the concentration n (unit L^{-d}, where d is the spatial dimension) in Eq. (10) was replaced by a probability density Q?

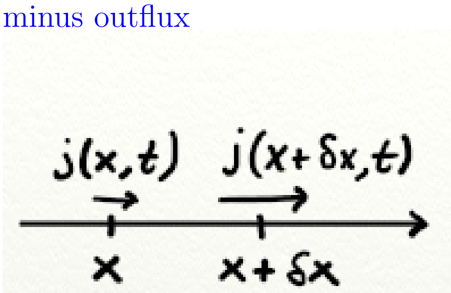
   Solution
   The units of the diffusion coefficient are L^2/T for all three cases.

b) Derive the one-dimensional diffusion equation (10) starting from Fick's law.

   Solution
   Fick's law states that matter is transported from high concentrations to low concentrations according to:

$$j(x,t) = -D\frac{\partial n}{\partial x}\,.$$

   The change of concentration in a small interval δx is equal to influx minus outflux



   is

$$\frac{\partial}{\partial t}\int_x^{x+\delta x} \mathrm{d}x' n(x',t) = j(x,t) - j(x+\delta x,t)\,. \qquad (11)$$

   Divide by δx and let $\delta x \to 0$ to get

$$\frac{\partial n}{\partial t} = -\frac{\partial j}{\partial x} = D\frac{\partial^2 n}{\partial x^2}$$

   i.e. same as the diffusion equation Eq. (10).

c) Explain how the derivation in subtask b) must be modified to result in Fisher's equation instead of the diffusion equation. Write down Fisher's equation.

   Solution
   Fisher's equation assumes a reaction term on the form of a logistic growth. We therefore need to add a source term $f(x,t) = rn(x,t)(1 - n(x,t)/K)$ to the right-hand side of the continuity equation (11), i.e. the concentration can change due to flux over the boundaries or due to local production.

Fisher's equation becomes:

$$\frac{\partial}{\partial t} n(x,t) = rn(x,t)\left(1 - \frac{n(x,t)}{K}\right) + D\nabla^2 n(x,t)\,. \tag{12}$$

d) In one of the problem sets you simulated a one-dimensional version of Fisher's equation. The length unit in your simulations was (at least supposed to be) unity. Under what conditions do you expect this to be a good choice for your simulations?

Solution

In the simulations the spatial domain was divided into discrete patches of unit length, $\Delta x = 1$. Upon changing to dimensionless units $n = n'K$, $t = t'/r$, $x = \sqrt{D/r}x'$, Eq. (12) becomes

$$\frac{\partial}{\partial t'} n'(x,t) = n'(x,t)\,(1 - n'(x,t)) + \frac{\partial^2}{\partial x'^2} n'(x,t)\,.$$

This equation is parameter free. In these units the spatial discretization step $\Delta x'$ in a numerical solution should satisfy $\Delta x' \ll 1$ (in the limit $\Delta x \to 0$ the discretized solution approaches the actual continuous solution of Fisher's equation). Using that $\Delta x' = \sqrt{r/D}\Delta x$, a step length $\Delta x = 1$ is a good choice if $\sqrt{r/D} \ll 1$.

Side remark: in the problem set $r = b - c = 1/2$ and $D = 1$ was used, which is really on the border of what is suitable. Actually, the observed wave speed $c \approx 1.8$ smaller than the theoretical minimal speed $c_{\min} = 2$ in this solution is mainly due to this discretization error.

5. Coupled oscillators [2 points] Consider N coupled oscillators with phases $\theta_1, \theta_2, \ldots \theta_N$ and with the following time evolution

$$\dot{\theta}_i = \omega_i + \frac{K}{N}\sum_{j=1}^{N} \sin(\theta_j - \theta_i)\,. \tag{13}$$

a) When $N = 2$, show that the two oscillators approaches a phase-locked dynamics. What is the relative phase between the oscillators?

Solution

Let $\theta_- \equiv \theta_1 - \theta_2$ and use Eq. (13) to get the dynamics

$$\dot{\theta}_- = \omega_1 - \omega_2 - K\sin\theta_-\,.$$

This dynamics has fixed points where $\theta^*_{-,1} = \arcsin((\omega_1 - \omega_2)/K)$ and $\theta^*_{-,2} = \pi - \arcsin((\omega_1 - \omega_2)/K)$. By plotting the flow, one finds that $\theta^*_{-,1}$ is a globally attracting stable fixed point.

Since the relative angle θ_- is attracted to a single fixed point, the oscillators approaches a phase locked dynamics with relative phase $\theta_- = \arcsin((\omega_1 - \omega_2)/K)$.

b) What is the angular velocity of the phase-locked dynamics you found in subtask a)?

Solution
The angular velocity can be obtained by evaluating $\dot{\theta}_1$ (or alternatively $\dot{\theta}_2$) using the phase-locked solution found in subtask a):

$$\begin{aligned}\dot{\theta}_1 &= \omega_1 - \frac{K}{2}\sin\theta_- \\ &= \omega_1 - \frac{1}{2}(\omega_1 - \omega_2) \\ &= \frac{1}{2}(\omega_1 + \omega_2)\,.\end{aligned}$$

Thus, the compromise frequency is the average frequency between the two oscillators.

Alternative solution Since we know $\dot{\theta}_1 = \dot{\theta}_2$ we can use

$$\omega_{\mathrm{c}} = \frac{1}{2}(\dot{\theta}_1 + \dot{\theta}_2) = \frac{1}{2}(\omega_1 + \omega_2)\,.$$

c) Introduce the order parameter

$$re^{\mathrm{i}\psi} = \frac{1}{N}\sum_{j=1}^{N} e^{\mathrm{i}\theta_j} \tag{14}$$

Show how to rewrite Eq. (13) on the following form

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

Solution
Multiplication of Eq. (14) with $e^{-\mathrm{i}\theta_i}$ and evaluation of the imaginary part gives

$$\begin{aligned}\mathcal{I}m[re^{\mathrm{i}(\psi-\theta_i)}] &= \mathcal{I}m\left[\frac{1}{N}\sum_{j=1}^{N} e^{\mathrm{i}(\theta_j-\theta_i)}\right] \\ \Rightarrow r\sin(\psi - \theta_i) &= \frac{1}{N}\sum_{j=1}^{N}\sin(\theta_j - \theta_i)\,.\end{aligned}$$

Inserting this relation into Eq. (13) gives the sought form:

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

d) Briefly explain what the order parameter quantifies and briefly explain what the purpose is of introducing the order parameter in the Kuramoto model.

Solution

The value of ψ measures the average phase of the oscillators and the value of r measures the phase-coherence, it takes small values if the phases of the oscillators are close to uniformly distributed, and it takes values close to unity if the phases are closeby (synchronisation).

The introduction of the order parameter allows to rewrite Eq. (13) in terms of a dynamics where all interactions between oscillators are encoded in r and ψ (as seen in subtask c). This allows for the application of a mean-field theory in the limit of a large numbers of oscillators, $N \to \infty$, in which the order parameter can be replaced by an average quantity and the dynamics of individual oscillators decouples. The mean-field theory allows for an analytical solution of the Kuramoto model in the limit of large N. The value of the order parameter is determined by the self-consistency condition that it is a constant when evaluated from the solutions to the decoupled equations for θ_i.

6. Difference filters [2 points] A linear filter denotes the procedure of convoluting a time series x_n with a discrete weight function a_n to form a filtered time series y_n:

$$y_n \equiv \sum_{m=-\infty}^{\infty} a_m x_{n-m}\,. \tag{15}$$

a) A first-order difference filter has weights $a_0 = 1$, $a_1 = -1$ and all other $a_n = 0$. Evaluate Eq. (15) using these weights.

Solution

The filter becomes $y_n = a_0 x_{n-0} + a_1 x_{n-1} = x_n - x_{n-1}$.

b) A second-order difference filter is obtained by applying a first-order difference filter two times on a time series. Write down the form of Eq. (15) for a second-order filter and read off the non-zero weights a_n of the second-order difference filter.

Solution

Apply the time series twice:

$$y_n = (x_n - x_{n-1}) - (x_{n-1} - x_{n-2}) = x_n - 2x_{n-1} + x_{n-2}\,.$$

The non-zero weights are $a_0 = 1$, $a_1 = -2$, $a_2 = 1$.

c) Show that the first-order difference filter removes linear trends in a time series up to a constant, by applying it to a time series with a linear trend: $x_n = An + \eta_n$, where η_n denotes fluctuations around the linear trend and A is a constant.

Solution
Apply the first-order difference filter to the time series:

$$y_n = x_n - x_{n-1} = An + \eta_n - (A(n-1) + \eta_{n-1}) = A + \eta_n - \eta_{n-1}\,.$$

Thus, the filtered time series show fluctuations around a constant, the linear trend has been removed.

d) Show that the second-order difference filter may be used to remove a quadratic trend in a time series up to a constant.

Solution
Inspired from subtask c) we assume a time series with a quadratic trend as $x_n = An^2 + \eta_n$. Applying the second-order difference filter we obtain

$$\begin{aligned} y_n &= x_n - 2x_{n-1} + x_{n-2} \\ &= An^2 + \eta_n - 2(A(n-1)^2 + \eta_{n-1}) + A(n-2)^2 + \eta_{n-2} \\ &= 2A + \eta_n - 2\eta_{n-1} + \eta_{n-2}\,. \end{aligned}$$

That is the quadratic trend is removed up to a constant.

e) Show that a p:th order difference filter, obtained by applying a first-order difference filter p times, can be used to eliminate a time series that is on the form of a general polynomial in n of degree $p-1$.

Solution
Assume the time series is $x_n = \sum_{m=0}^{p-1} A_m n^m$, where A_m are general coefficients of the polynomial. Applying a first-order difference filter we obtain

$$\begin{aligned} y_n &= \sum_{m=0}^{p-1} A_m n^m - \sum_{m=0}^{p-1} A_m (n-1)^m \\ &= \sum_{m=0}^{p-1} A_m n^m - \sum_{m=0}^{p-1} A_m \sum_{k=0}^{m} \binom{m}{k} n^{m-k}(-1)^k \\ &= \sum_{m=0}^{p-1} A_m n^m - \sum_{m=0}^{p-1} A_m \underbrace{\binom{m}{0}}_{1} n^m - \sum_{m=0}^{p-1} A_m \sum_{k=1}^{m} \binom{m}{k} n^{m-k}(-1)^k \\ &= -\sum_{m=0}^{p-1} A_m \sum_{k=1}^{m} \binom{m}{k} n^{m-k}(-1)^k\,. \end{aligned}$$

Thus y_n is a polynomial of degree $p-2$, the degree is lowered by one by applying the filter. For each successive application the degree of the polynomial is reduced by one. After $p-1$ applications of the filter, the degree of the polynomial is zero, i.e. the filtered series is constant. Finally, after the p:th application of the filter, the result becomes zero.