

Matematyka stosowana

Modelowanie matematyczne w biologii i medycynie

Urszula Foryś

urszula@mimuw.edu.pl

<http://www.mimuw.edu.pl/~urszula>

Jan Poleszczuk (ilustracje)

j.poleszczuk@mimuw.edu.pl

Uniwersytet Warszawski, 2011



Streszczenie. Wykład dotyczy szeroko pojętego modelowania matematycznego w biologii i medycynie. Jego podstawę stanowią modele ekologiczne, budowane na bazie równań różniczkowych i różnicowych, teorii grafów i teorii gier, poszerzone o modele reakcji odpornościowej i podstawy klasycznej genetyki (teoria Mendla) w kontekście łańcuchów Markowa.

Wersja internetowa wykładu:

<http://mst.mimuw.edu.pl/lecture.php?lecture=mbm>

(może zawierać dodatkowe materiały)



Niniejsze materiały są dostępne na [licencji Creative Commons 3.0 Polska](#):
Uznanie autorstwa — Użycie niekomercyjne — Bez utworów zależnych.

Copyright © U. Foryś, Uniwersytet Warszawski, Wydział Matematyki, Informatyki i Mechaniki, 2011. Niniejszy plik PDF został utworzony 9 czerwca 2011.



Projekt współfinansowany przez Unię Europejską w ramach Europejskiego Funduszu Społecznego.



Skład w systemie L^AT_EX, z wykorzystaniem m.in. pakietów beamer oraz listings. Szablony podręcznika i prezentacji: Piotr Krzyżanowski; koncept: Robert Dąbrowski.

Spis treści

Wstęp — pojęcie modelu matematycznego	5
1. Modelowanie pojedynczej populacji I	7
1.1. Model Malthusa	7
1.2. Procesy rozrodczości i śmiertelności	9
1.3. Migracje	10
2. Modelowanie pojedynczej populacji II	14
2.1. Równanie logistyczne — model Verhulsta	14
2.2. Dyskretne równanie logistyczne	17
3. Modelowanie pojedynczej populacji III	20
3.1. Efekt Alleego	20
3.2. Funkcjonalna odpowiedź Hollinga i funkcja Hilla	22
4. Modele pojedynczej populacji z uwzględnieniem wieku I	25
4.1. Ciąg Fibonacciego	25
4.2. Macierze Lesliego	28
5. Modele pojedynczej populacji z uwzględnieniem wieku II	32
5.1. Modele z opóźnieniem	32
6. Modele oddziaływań między dwiema populacjami I	36
6.1. Model Lotki – Volterra	36
6.2. Konstruktywna krytyka modelu Lotki – Volterra	40
7. Modele oddziaływań między dwiema populacjami II	42
7.1. Model drapieżnik – ofiara z ograniczoną pojemnością środowiska dla ofiar	42
7.2. Model z kryjówkami dla ofiar	47
8. Modele oddziaływań między dwiema populacjami III	49
8.1. Układ konkurujących gatunków	49
8.2. Modelowanie symbiozy	52
9. Doświadczenia Mendla: łańcuchy Markowa w klasycznej genetyce I	57
9.1. Łańcuchy Markowa	58
9.2. Klasyfikacja stanów i łańcuchów	60
10. Doświadczenia Mendla: łańcuchy Markowa w klasycznej genetyce II	63
10.1. Łańcuchy pochłaniające i ciągle krzyżowanie z dominantą	63
10.2. Łańcuchy regularne i ciągle krzyżowanie z hybrydą	66
11. Modele z zależnością przestrzenną: dyfuzja w procesach biologicznych	69
11.1. Równania ewolucyjne: równanie Fishera – Kołmogorowa	71
11.2. Wzory Turinga	76
12. Modelowanie odpowiedzi odpornościowej	80
12.1. Układ odpornościowy człowieka	80
12.2. Proste modele odpowiedzi odpornościowej	82
13. Model Marczuka humoralnej odpowiedzi odpornościowej	87
13.1. Prezentacja modelu Marczuka	87
13.2. Podstawowe własności układu (13.1)	89

14. Łańcuchy pokarmowe	96
14.1. Podstawy teorii grafów	96
14.2. Łańcuchy pokarmowe	98
15. Podstawy teorii gier	103
15.1. Podstawy teorii gier	103
15.1.1. Gra w postaci normalnej	103
15.2. Przykłady gier	105
15.2.1. Gra jastrzęb – gołąb	105
15.2.2. Dylemat więźnia	106
Literatura	108

Wstęp — pojęcie modelu matematycznego

W dzisiejszych czasach coraz szerzej rozumiana i akceptowana jest idea modelowania matematycznego zjawisk przyrodniczych. Pod określeniem model rozumiemy dwuskładnikową strukturę — pierwszy składnik stanowi teoretyczny opis danego zjawiska na podstawie bieżącej wiedzy, często ten opis nazywany modelem heurystycznym, natomiast drugi składnik to struktura matematyczna, w której próbujemy odzwierciedlić model heurystyczny. Budując model heurystyczny musimy zdecydować, jakie procesy wchodzące w skład danego zjawiska mają wpływ na końcowy efekt, który chcemy odzwierciedlić za pomocą modelu, a które procesy można pominąć. To bardzo ważny etap, gdyż pozwala zredukować liczbę zmiennych i parametrów stosowanych potem do budowy równań czy innego typu struktury matematycznej.

Klasycznie najczęściej stosowanym formalizmem matematycznym są równania różniczkowe i różnicowe, ponieważ pierwsze modele w biologii budowane były przy wykorzystaniu ugruntowanych sposobów, z jakich korzystano w modelowaniu w fizyce. Obecnie często buduje się modele stochastyczne czy modele mieszane. Przyjmując dany formalizm matematyczny musimy jasno określić, co stanowi zmienne, a co parametry naszego modelu. Parametry należy wyznaczyć na podstawie eksperymentów, pomiarów czy obserwacji w naturze, natomiast zmienne stanowią niewiadome, które obliczamy/analizujemy ich przebieg na podstawie modelu. Dobrze zbudowany model stanowi przedmiot badań analitycznych i komputerowych, dzięki którym poznajemy własności rozwiązań.

Budując model należy pamiętać o pewnych podstawowych zasadach. Poprawnie zbudowany model powinien mieć rozwiązania, rozwiązania powinny być jednoznaczne, a także stabilne względem warunków początkowych i parametrów. Taka koncepcja poprawnego modelu została zaproponowana przez [Hadamarda](#).

Po zbudowaniu modelu kolej na jego weryfikację. Trzeba zatem zaprojektować odpowiednie eksperymenty — dopóki wyniki eksperymentów nie przeczą wnioskowi płynącemu z modelu, dopóty model możemy uważać za poprawny. Wiąże się to z [koncepcją falsyfikowalności](#), która mówi, że model czy teoria naukowa powinny być tak zbudowane, aby za pomocą eksperymentu można było je obalić. Trzeba mieć świadomość, że nawet bardzo duża liczba eksperymentów potwierdzających nie daje całkowitej gwarancji poprawności modelu, ale wystarczy jeden eksperyment falsyfikujący, aby wykazać jego niepoprawność.

W ramach tego wykładu zaprezentujemy szeroki przegląd modeli i metod matematycznych stosowanych w biologii i medycynie, zaczynając od klasycznych modeli populacyjnych opisywanych równaniami różnicowymi, równaniami różniczkowymi zwyczajnymi, równaniami różniczkowymi z opóźnionym argumentem, równaniami różniczkowymi cząstkowymi, a potem przejdziemy do modeli budowanych w oparciu o teorię grafów i modeli stochastycznych.

Zauważmy tutaj zasadnicze różnice między opisem ciągłym a dyskretnym. W opisie ciągłym zakładamy, że znamy prawa rządzące danym zjawiskiem w dowolnej chwili t , natomiast w opisie dyskretnym interesuje nas tylko to, co dzieje się w wyróżnionych momentach t_n , $n \in \mathbb{N}$, na przykład w takich, w których dokonujemy pomiaru. Z kolei z matematycznego punktu widzenia — w opisie dyskretnym stosujemy ciągi liczbowe i możemy oczekiwać, że przy odpowiednich założeniach kolejne wyrazy ciągu będą liczbami naturalnymi, a co za tym idzie bezpośrednio mogą opisywać np. liczebność populacji, a w opisie ciągłym, rozwiązanie równania różniczko-

wego tylko w bardzo szczególnych przypadkach przyjmuje wartości ze zbioru liczb naturalnych. W związku z tym, jeśli stosujemy opis ciągły, to musimy pamiętać, że „liczebność” populacji oznacza w rzeczywistości jej zagęszczenie, czyli liczbę osobników przypadającą na jednostkę powierzchni siedliska. Podobną interpretację możemy stosować w przypadku opisu dyskretnego i wtedy nie musimy dbać o dobieranie parametrów w taki sposób, aby wyrazy ciągu należały do zbioru liczb naturalnych. W trakcie wykładu będziemy zamiennie posługiwać się pojęciem liczebności populacji i jej zagęszczenia.

W ramach tego wykładu omówimy następujące zagadnienia

- modele pojedynczej populacji
 - model Malthusa w wersji dyskretniej i ciągłej
 - migracje
 - ciąg Fibonacciego
 - modele ze strukturą wieku — macierze Lesliego
 - model logistyczny
 - efekt Allego
 - funkcjonalna odpowiedź Hollinga typu I i II
 - równanie logistyczne z opóźnieniem
 - dyskretne równanie logistyczne: chaos
- modele opisujące układ dwóch populacji w ekosystemie
 - klasyczny model Lotki – Volterry
 - inne modele typu drapieżnik – ofiara
 - konkurencja
 - mutualizm
 - model Kołmogorowa
 - model Maya
 - model Nicholsona – Baileya
- modele w epidemiologii i immunologii
 - modele epidemiologiczne
 - proste modele odpowiedzi odpornościowej
 - model Marczuka
- dyfuzja w procesach biologicznych
 - równanie dyfuzji
 - równanie Fishera
 - model melanogenezy — wzory Turinga
- analiza łańcuchów pokarmowych przy użyciu teorii grafów
- łańcuchy Markowa
 - klasyfikacja stanów i łańcuchów — przykłady biologiczne
 - zastosowanie łańcuchów w klasycznej genetyce — doświadczenia Mendla
- teoria gier w zagadnieniach strategii populacji
 - gra jastrzęb – gołąb
 - równowaga Nasha i strategie ewolucyjnie stabilne

1. Modelowanie pojedynczej populacji I

W pierwszej części wykładu przedstawimy podstawowe modele opisujące ekosystem, w którym występuje tylko jedna populacja i opiszemy dynamikę takiej populacji w zależności od przyjętego modelu heurystycznego i formalizmu matematycznego.

1.1. Model Malthusa

W najprostszym przypadku zakładamy, że opisywana populacja ma w danym siedlisku bardzo dobre warunki rozwoju, wyrażające się głównie w taki sposób, że każdy osobnik ma nieograniczony dostęp do pożywienia i miejsc lęgowych oraz że obserwujemy jedynie proces rozmnażania. Aby opis był możliwie najprostszy, osobniki w danej populacji powinny być jednakowe i podlegać jednakowym prawom. Co więcej, powinny być równomiernie rozłożone w przestrzeni, żeby nie wystąpiła konieczność analizowania zależności od przestrzeni, a tylko od czasu. Przyjmujemy zatem następujący opis heurystyczny dla populacji \mathcal{P}

- populacja jest jednorodna, składa się z genetycznie identycznych osobników rozmnażających się partenogenetycznie;
- osobnik rodzi się w pełni ukształtowany, zdolny do rozrodu i może rozmnażać się w dowolnym wieku;
- momenty rozmnażania są w dowolnym przedziale czasu rozłożone jednostajnie;
- każdy osobnik wydaje na świat potomstwo co τ jednostek czasu, τ jest ustalone i jednakowe dla wszystkich osobników;
- każdorazowo jeden rodzic ma λ osobników potomnych.

Na podstawie powyższych informacji spróbujemy zaproponować równanie opisujące średnią liczebność populacji $N(t)$ w chwili t . Dokładniej, załóżmy że znamy liczebność w pewnej ustalonej chwili t i chcemy obliczyć $N(t + \Delta t)$ po upływie czasu Δt . Zauważmy, że w przedziale czasu $(t, t + \Delta t)$ jest

$$\frac{\Delta t}{\tau}$$

momentów rozmnażania. Każdy rodzic ma w tym przedziale czasu średnio

$$\frac{\Delta t}{\tau} \lambda$$

potomków. Liczba osobników zdolnych do rozmnażania w chwili t wynosi $N(t)$, zatem przyrost liczebności w krótkim przedziale czasu $(t, t + \Delta t)$ możemy przybliżyć równaniem

$$N(t + \Delta t) - N(t) = \frac{\Delta t}{\tau} \lambda N(t). \quad (1.1)$$

Dzieląc stronami przez Δt i przechodząc do granicy otrzymujemy

$$\lim_{\Delta t \rightarrow 0} \frac{N(t + \Delta t) - N(t)}{\Delta t} = \frac{d}{dt} N(t) = \frac{\lambda}{\tau} N(t).$$

Przyjmijmy oznaczenia $\frac{d}{dt} N(t) = \dot{N}(t)$, $\frac{\lambda}{\tau} = r$ i przepiszmy powyższe równanie w postaci

$$\dot{N}(t) = rN(t), \quad (1.2)$$

gdzie r nazwiemy współczynnikiem rozrodczości populacji \mathcal{P} . Zbudowaliśmy w ten sposób model ciągły (tzn. zmienna czasowa t zmienia się w sposób ciągły i model heurystyczny jest opisany za pomocą równania różniczkowego). Jeśli natomiast założymy, że $\Delta t = 1$, gdzie 1 oznacza ustaloną jednostkę czasu i zastosujemy oznaczenie $N(t) = N_t$, to równanie (1.1) możemy zapisać jako

$$N_{t+1} = (1 + r)N_t, \quad (1.3)$$

otrzymujemy więc model dyskretny (zmienna czasowa zmienia się w sposób dyskretny, najczęściej $t \in \mathbb{N}$, a liczebność populacji opisana jest za pomocą równania rekurencyjnego).

Obydwa równania (1.2) i (1.3) nazywamy **modelem Malthusa**, odpowiednio w wersji ciągłej i dyskretniej. **Thomas Malthus**, angielski ekonomista i demograf, pod koniec XVIII wieku zwrócił uwagę na zbyt szybki przyrost liczebności populacji ludzkiej.

W pracy zatytułowanej „**An Essay on the Principle of Population**” sformułował swoje słynne prawo, które mówi, że liczba ludności wzrasta w tempie geometrycznym, natomiast zasoby żywności w tempie arytmetycznym, co w oczywisty sposób musi doprowadzić do katastrofy. Zauważmy, że równanie dyskretnie (1.3) w bezpośredni sposób odzwierciedla fragment prawa Malthusa odnoszący się do liczebności populacji ludzkiej, natomiast równanie (1.2) jest jego ciągłym odpowiednikiem.

Każde z równań (1.2) i (1.3) możemy w prosty sposób rozwiązać, korzystając z ich liniowości. Przypomnijmy, że dla układów liniowych równań różniczkowych poszukujemy rozwiązań bazowych w postaci wykładniczej

$$N(t) = N_0 e^{\lambda t},$$

a dla równań różnicowych — w postaci potęgowej

$$N_t = N_0 \lambda^t,$$

przy czym stałą λ nazywamy wartością własną. Oczywiście w przypadku jednej zmiennej równanie (1.2) rozwiązujemy wprost jako równanie o zmiennych rozdzielonych

$$\int_{N_0}^{N(t)} \frac{dN}{N} = r \int_0^t ds \Rightarrow \ln \frac{N(t)}{N_0} = rt,$$

czyli $N(t) = N_0 e^{rt}$. Natomiast równanie (1.3) opisuje ciąg geometryczny o ilorazie $\frac{N_{t+1}}{N_t}$, którego poszczególne wyrazy obliczamy ze wzoru ogólnego

$$N_t = N_0 (1 + r)^t,$$

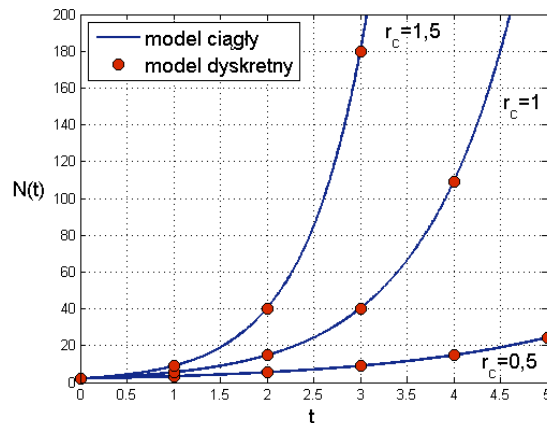
co można łatwo udowodnić stosując zasadę indukcji matematycznej.

Jakie są własności rozwiązań obu wersji modelu Malthusa? Widzimy, że zarówno funkcja $N_0 e^{rt}$ jak i ciąg geometryczny $N_0 (1 + r)^t$, $r > 0$, są rosnące nieograniczenie, por. rys. 1.1. Co więcej, dla ustalonego N_0 , jeśli współczynnik rozrodczości w modelu ciągłym oznaczymy przez r_c , a w dyskretnym — r_d , to możemy dobrać r_d do r_c (albo odwrotnie), tak by rozwiązania dla $t \in \mathbb{N}$ pokrywały się. Mianowicie, chcemy mieć

$$N_0 e^{r_c n} = N_0 (1 + r_d)^n, \quad n \in \mathbb{N},$$

więc $e^{r_c} = 1 + r_d$, czyli $r_d = e^{r_c} - 1$, przy czym oczywiście $r_d > 0 \Leftrightarrow r_c > 0$.

Z dynamiki modeli (1.2) i (1.3) wynika jasno, że mogą one mieć tylko ograniczony zakres stosowalności. Jedynie hipotetyczna populacja może rozrastać się nieograniczenie, natomiast w rzeczywistości zawsze działają różne mechanizmy ograniczające taki przyrost, w szczególności pojemność siedliska, w którym występuje populacja \mathcal{P} jest zawsze ograniczona i zbyt duża liczba osobników w tym siedlisku się nie zmieści. Pokreślić należy, już sam Malthus zwrócił uwagę na fakt, że ilość żywności przyrasta znacznie wolniej niż liczba ludności, co w konsekwencji musi prowadzić do konkurencji o pożywienie. Zawsze także obok procesu rozrodczości występuje proces śmiertelności.



Rysunek 1.1. Porównanie rozwiązań wersji ciągłej (1.2) oraz dyskretnej (1.3) modelu Malthusa dla różnych wartości parametru r_c przy założeniu, że $r_d = e^{r_c} - 1$.

1.2. Procesy rozrodczości i śmiertelności

Ponieważ uświadomiliśmy sobie, że w naszym modelu dynamiki populacji musimy uwzględnić nie tylko rozrodczość, ale także i inne procesy, w szczególności śmiertelność osobników danej populacji, a chcemy to ponownie zrobić w jak najprostszy sposób, to odpowiedni model heurystyczny budujemy analogicznie jak w przypadku procesu rozrodczości, przyjmując takie same założenia o jednorodności populacji. Wobec tego zamiast równania (1.1) dostaniemy w rezultacie

$$N(t + \Delta t) - N(t) = r\Delta t N(t) - s\Delta t N(t),$$

gdzie s nazywamy współczynnikiem śmiertelności. Odzwierciedla on procent osobników umierających w jednostce czasu. Zauważmy, że $\frac{1}{s}$ możemy interpretować jako średnią długość życia pojedynczego osobnika.

Ostatecznie z analitycznego punktu widzenia, po uwzględnieniu procesu śmiertelności w modelu Malthusa dostajemy taką samą strukturę matematyczną jak poprzednio. Oznaczmy przez $r_n = r - s$ współczynnik rozrodczości „netto”, czyli różnicę między współczynnikiem rozrodczości a śmiertelności. Często r_n nazywamy współczynnikiem reprodukcji lub współczynnikiem przyrostu naturalnego. Wtedy mamy

$$\dot{N}(t) = r_n N(t) \Rightarrow N(t) = N_0 e^{r_n t} \quad (1.4)$$

w modelu ciągłym oraz

$$N_{t+1} = (1 + r_n) N_t \Rightarrow N_t = N_0 (1 + r_n)^t \quad (1.5)$$

w modelu dyskretnym.

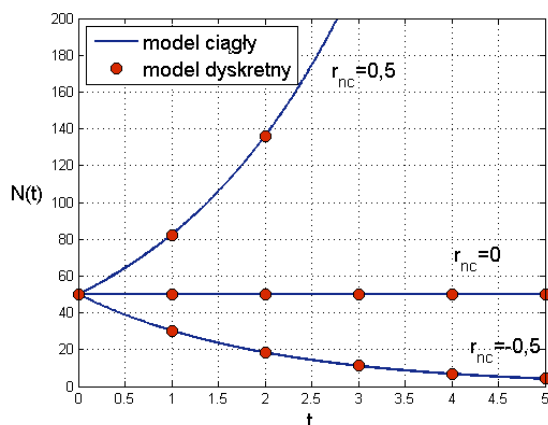
Widzimy, że jeśli przyrost naturalny jest dodatni, $r_n > 0$, to liczebność populacji rośnie, choć oczywiście nieco wolniej niż dla $s = 0$. Natomiast gdy $r_n < 0$, to charakter rozwiązań obu modeli zmienia się. W modelu ciągłym mamy

$$\dot{N}(t) < 0 \quad \text{dla} \quad N(t) > 0,$$

zatem liczebność populacji spada. Rozwiązaniem jest funkcja wykładnicza o ujemnym wykładniku, czyli

$$\lim_{t \rightarrow \infty} N(t) = 0,$$

wobec tego asymptotycznie populacja wymiera. Analogicznie w modelu dyskretnym mamy ciąg geometryczny o ilorazie mniejszym niż 1, czyli także N_t maleje do zera przy $t \rightarrow \infty$, por. rys. 1.2.



Rysunek 1.2. Porównanie rozwiązań wersji ciągłej (1.4) oraz dyskretnej (1.5) modelu Malthusa z uwzględnieniem śmiertelności dla różnych wartości współczynnika r_{nc} przy założeniu, że $r_{nd} = e^{r_{nc}} - 1$, gdzie r_{nc} i r_{nd} to współczynniki rozrodczości „netto” odpowiednio dla modelu ciągłego i dyskretnego.

1.3. Migracje

Kolejnym krokiem przybliżającym nasz model do rzeczywistości może być uwzględnienie emigracji osobników w związku z ograniczoną pojemnością siedliska. W najprostszych modelach zakłada się dwa typy migracji

1. migracja stała w czasie;
2. migracja zależna od zagęszczenia.

Zauważmy, że przy drugim typie migracji w równaniu (1.4) ponownie musimy uwzględnić składnik $\alpha N(t)$, gdzie α oznacza frakcję osobników migrujących przypadających na jednostkę czasu, zatem

$$\dot{N}(t) = r_n N(t) - \alpha N(t)$$

i taki sposób migracji nie wpływa na ogólną postać rozwiązań modelu — rozwiązaniem jest zawsze funkcja wykładnicza, przy czym jeśli emigracja nie jest zbyt duża, czyli $r_n > \alpha$, to populacja w dalszym ciągu rośnie nieograniczenie, jeśli przyrost naturalny jest dodatni, natomiast zbyt duża emigracja $\alpha > r_n$ spowoduje, że siedlisko zostanie opuszczone, ale matematycznie realizuje się to w nieskończonym czasie. Wobec tego dynamika nie różni się od modelu rozrodczości/śmiertelności omawianego poprzednio. Tak samo przedstawia się dynamika wersji dyskretnej.

Zastanówmy się wobec tego jak zmieni się model, jeśli założymy, że emigracja jest stała w czasie i nie zależy od zagęszczenia populacji. Wtedy współczynnik emigracji m odzwierciedla liczbę osobników, które emigrują (zatem ubywa osobników w siedlisku) w jednostce czasu (na jednostkę powierzchni, ponieważ opisujemy zagęszczenie). Zamiast równania (1.4) otrzymamy

$$\dot{N}(t) = r_n N(t) - m, \quad m > 0. \quad (1.6)$$

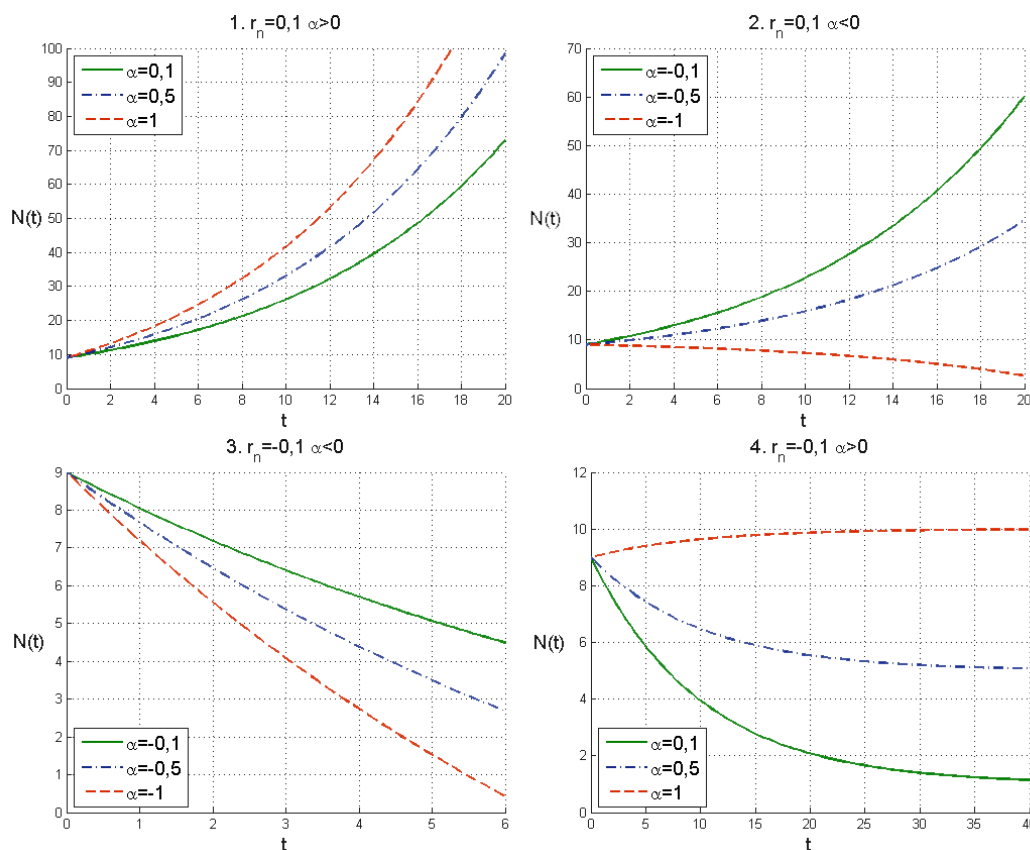
Zwróćmy uwagę, że z matematycznego punktu widzenia możemy rozważać różne układy znaków parametrów r_n i α w równaniu

$$\dot{N}(t) = r_n N(t) + \alpha, \quad (1.7)$$

odzwierciedlające odpowiednio

1. $r_n > 0$ i $\alpha > 0$ — dodatni przyrost naturalny i imigrację;
2. $r_n > 0$ i $\alpha < 0$ — dodatni przyrost naturalny i emigrację;
3. $r_n < 0$ i $\alpha < 0$ — ujemny przyrost naturalny i emigrację;
4. $r_n < 0$ i $\alpha > 0$ — ujemny przyrost naturalny i imigrację.

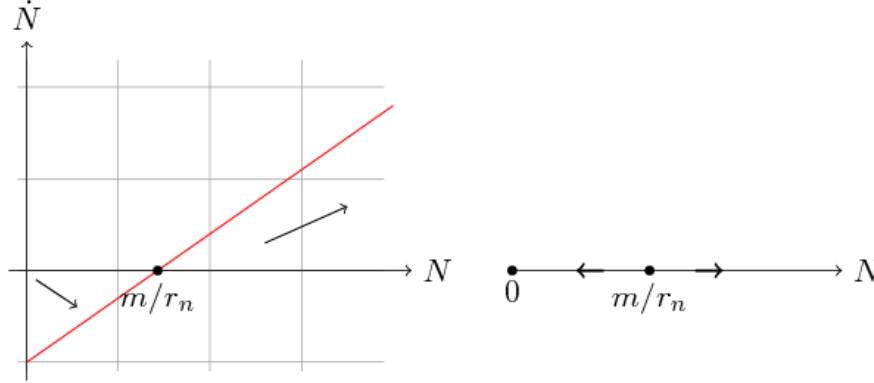
Model (1.7) możemy także interpretować w inny sposób. Współczynnik $\alpha > 0$ oznacza wprowadzanie (introdukcję) nowych osobników do siedliska, natomiast $\alpha < 0$ oznacza odławianie osobników. Wyobraźmy sobie zatem staw z rybami i spróbujmy zinterpretować każdy przypadek. Najłatwiejsza wydaje się interpretacja drugiego przypadku — populacja ryb zwiększa się, więc je odławiamy. Przypadek pierwszy jest z tego punktu widzenia najbardziej problematyczny, ale zakładając niewielką wartość r_n możemy pewien czas możemy nie zauważać istotnego przyrostu, więc introdukcja w celu podtrzymania gatunku może wydawać się konieczna. Podobnie w trzecim przypadku — przy małym ujemnym przyroście naturalnym możemy przez pewien czas nie zauważać zmniejszania się populacji z tego powodu i eksploatować tę populację przez odłów (czasem zdarza się tak, że zauważamy ujemny przyrost naturalny, ale z różnych powodów w dalszym ciągu eksploatujemy populację). Przypadek czwarty oznacza oczywiście pozytywną ingerencję człowieka — ponieważ przyrost naturalny jest ujemny, więc zarybiamy staw, by utrzymać populację w siedlisku.



Rysunek 1.3. Przykładowe rozwiązania równania (1.7) dla różnych układów znaków parametrów r_n oraz α .

Z punktu widzenia interesującego nas teraz zagadnienia ograniczenia pojemności siedliska najważniejszy jest przypadek drugi, zatem przeprowadzimy jego analizę, a pozostałe trzy należy potraktować jako ćwiczenie. Zajmiemy się więc teraz równaniem (1.6) z dodatnimi parametrami

$r_n > 0$ i $m > 0$. Narysujmy najpierw zależność (N, \dot{N}) w przestrzeni $(\mathbb{R}^+)^2$, gdzie \mathbb{R} oznacza liczby rzeczywiste nieujemne, por. rys. 1.4.



Rysunek 1.4. Zależność (N, \dot{N}) dla równania (1.6) w przestrzeni $(\mathbb{R}^+)^2$, gdzie \mathbb{R}^+ oznacza liczby rzeczywiste nieujemne (po lewej) oraz odpowiadający tej zależności portret fazowy (po prawej).

Zauważmy, że pochodna $\dot{N}(t)$ jest dla $N < \frac{m}{r_n}$ ujemna, co oznacza, że liczebność maleje, natomiast dla $N > \frac{m}{r_n}$ mamy $\dot{N}(t) > 0$, czyli liczebność rośnie. Jakie wnioski możemy wysnuć z tej prostej analizy? Musimy zacząć od przypomnienia sobie podstawowych własności rozwiązań równań różniczkowych zwyczajnych (RRZ). W naszym przypadku prawa strona równania (1.6) jest funkcją klasy \mathbf{C}^1 (nawet analityczną), więc rozwiązania tego równania są jednoznaczne. Co więcej, ponieważ równanie (1.6) jest liniowe, więc dla dowolnego $N_0 \geq 0$ (tylko taki warunek początkowy ma sens biologiczny, z analitycznego punktu widzenia możemy rozpatrywać dowolne $N_0 \in \mathbb{R}$) rozwiązanie istnieje dla wszystkich $t \in \mathbb{R}$, choć zwykle interesuje nas przewidywanie dynamiki populacji w przyszłości, więc ograniczamy się do $t \geq 0$. Zauważmy dalej, że $\bar{N} = \frac{m}{r_n}$ jest rozwiązaniem stacjonarnym, czyli jeśli $N_0 = \frac{m}{r_n}$, to $N(t) = \frac{m}{r_n}$ dla dowolnego t , co oznacza, że rozwiązanie nie zmienia się w czasie. Jest to jedyne rozwiązanie stacjonarne. Stąd jeśli $N_0 < \bar{N}$, to liczebność populacji maleje i nie ma żadnej „bariery”, która mogłaby ten spadek zahamować. Z teorii RRZ wynika, że taką barierę może stanowić tylko rozwiązanie stacjonarne. Faktycznie — założmy, że $N(t) \rightarrow N_g$ przy $t \rightarrow \infty$. Wtedy $\dot{N}(t) \rightarrow r_n N_g - m$. Ponieważ $N_g < \bar{N} = \frac{m}{r_n}$, to od pewnego momentu \bar{t} zachodzi $\dot{N}(t) < -a < 0$ z czego wynika, że $N(t) \leq N(\bar{t}) - a(t - \bar{t})$, a to implikuje, że $N(t) \rightarrow -\infty$. Wobec tego nie może zachodzić nierówność $N_g < \bar{N}$. Wnioskujemy więc, że $N(t) \rightarrow -\infty$, zatem istnieje taki moment $t_k > 0$, dla którego $N(t_k) = 0$. Ze względu na liniowość równania (1.6) możemy je łatwo rozwiązać (jak poprzednio jest to równanie o zmiennych rozdzielonych lub też możemy skorzystać z metody uzmienniania stałej)

$$N(t) = \frac{m}{r_n} + \left(N_0 - \frac{m}{r_n} \right) e^{r_n t}.$$

Analityczna postać rozwiązania tylko potwierdza wcześniejsze wnioski płynące z zależności (N, \dot{N}) . Zauważmy dalej, że obliczając drugą pochodną dostajemy

$$\ddot{N}(t) = r_n \dot{N},$$

zatem jeśli funkcja $N(t)$ jest rosnąca, to rośnie w sposób wypukły (jest to w zasadzie wzrost wykładniczy), natomiast gdy jest malejąca — jest także wklęsła, co oczywiście oznacza, że w skończonym czasie t_k przekracza oś poziomą (co już wcześniej zauważyliśmy). Znając postać

rozwiązania możemy wyznaczyć t_k jako taki czas, dla którego liczebność populacji zeruje się. Skoro $N(t_k) = 0$, to

$$0 = \frac{m}{r_n} + \left(N_0 - \frac{m}{r_n}\right) e^{r_n t_k} \implies t_k = \frac{1}{r_n} \ln \frac{m}{m - r_n N_0}, \quad \text{o ile } N_0 < \frac{m}{r_n}.$$

Przykładowe rozwiązania tego modelu zostały przedstawione na prawym górnym wykresie na rys. 1.3. Widzimy, że w zależności od wielkości emigracji populacja albo rośnie nieograniczenie, jak w przypadku bazowego modelu Malthusa, albo wymiera w skończonym czasie, jeśli zbyt dużo osobników emigruje. Pozostałe przypadki migracji analizujemy w analogiczny sposób, pozostawiamy tę analizę i płynące z niej wnioski jako ćwiczenie.

2. Modelowanie pojedynczej populacji II

2.1. Równanie logistyczne — model Verhulsta

Jak sprawdziliśmy, model Malthusa z emigracją nie rozwiązuje problemu ograniczonej pojemności siedliska. Wszystkie modele rozważane do tej pory bazują na modelu Malthusa, przy czym są liniowe. Teraz omówimy model zaproponowany ok. 1840 r. przez belgijskiego matematyka [Pierre'a F. Verhulsta](#) w [16], który jest prawdopodobnie pierwszym znanym w biologii populacyjnej modelem nieliniowym.

Model ten powstał w wyniku dyskusji grupy uczonych nad modelem Malthusa. Większość osób biorąca udział w tej dyskusji zgadzała się co do tego, że istnieją naturalne procesy, które hamują postulowany przez Malthusa geometryczny (nieograniczony) przyrost liczebności populacji. Według Verhulsta przyrost liczebności hamuje np. konkurencja o zasoby siedliska, która wiąże się w oczywisty sposób z ograniczonością tych zasobów. W związku z tym w równaniu (1.2) — które w dalszym ciągu będziemy utożsamiać z równaniem (1.4) dla $r_n > 0$ — należy uwzględnić składnik opisujący konkurencję. W tym kontekście mówimy o konkurencji wewnątrzgatunkowej w odróżnieniu od konkurencji zewnątrzgatunkowej, występującej pomiędzy osobnikami różnych gatunków. Trzeba się więc zastanowić, jak reprezentować tę konkurencję w języku matematyki. Wszystkie modele biologiczne powstające w XIX (czy nawet w pierwszej połowie XX wieku) miały swoje korzenie w ugruntowanych już modelach fizycznych. Zatem nie należy się dziwić, że konkurencję między osobnikami tego samego gatunku Verhulst opisał w taki sam sposób jak losowe zderzenia cząsteczek gazu elementarnego. Liczba takich zderzeń jest proporcjonalna do kwadratu liczby cząstek N^2 , a biorąc pod uwagę, że osobniki mogą konkurować ze sobą tylko wtedy, gdy się bezpośrednio zetkną można przyjąć, że konkurencja jest proporcjonalna do kwadratu liczebności, czyli równanie (1.2) zamienia się na

$$\dot{N}(t) = rN(t) - bN^2(t), \quad (2.1)$$

gdzie b oznacza współczynnik konkurencji wewnątrzgatunkowej i odzwierciedla „szkodliwość” tej konkurencji dla populacji \mathcal{P} .

To samo równanie możemy otrzymać na podstawie innego modelu heurystycznego. Wydaje się oczywiste, że przyjęty w modelu Malthusa (1.4) stały współczynnik reprodukcji $r_n = \text{const}$ stanowi często zbyt duże uproszczenie. Zwykle współczynnik ten zależy od wielu czynników, w tym także bezpośrednio od liczebności populacji. Jeśli liczebność jest niewielka, to osobniki mają dobry dostęp do zasobów siedliska i mogą się bez przeszkód rozmnażać, ale przy wzroście liczebności ten dostęp zmniejsza się. W takiej sytuacji przyrost *per capita* (czyli względny przyrost na jednego osobnika w populacji)

$$\frac{\dot{N}(t)}{N(t)} = \frac{dN}{dt} \frac{1}{N(t)},$$

stały w modelu Malthusa (1.4), staje się funkcją zależną od $N(t)$. Ponieważ według naszych założeń reprodukcja zmniejsza się wraz ze wzrostem liczebności populacji, to najprostsze matematyczne odzwierciedlenie takiej zależności stanowi liniowa funkcja malejąca. Wobec tego

$$\frac{\dot{N}(t)}{N(t)} = h(N(t)), \quad \text{gdzie} \quad h(x) = r \left(1 - \frac{x}{K}\right)$$

i przy $b = \frac{r}{K}$ równanie (2.1) i równanie logistyczne

$$\dot{N}(t) = rN(t) \left(1 - \frac{N(t)}{K}\right) \quad (2.2)$$

pokrywają się. Nazwa „równanie logistyczne” wiąże się z typowym logistycznym kształtem części rozwiązań równania (2.2). Równanie to nazywa się często także równaniem Verhulsta lub równaniem Verhulsta – Pearla, przy czym nazwisko biologa R. Pearla pojawia się w tej nazwie w związku z zastosowaniami równania (2.2) do opisu konkretnych populacji.

Przejdziemy teraz do analizy przebiegu rozwiązań równania (2.2). Jak poprzednio jest to równanie o zmiennych rozdzielonych, które możemy rozwiązać korzystając z rozkładu ułamka $\frac{1}{N(1-N/K)}$ na ułamki proste,

$$\frac{1}{N \left(1 - \frac{N}{K}\right)} = \frac{1}{N} + \frac{\frac{1}{K}}{1 - \frac{N}{K}}.$$

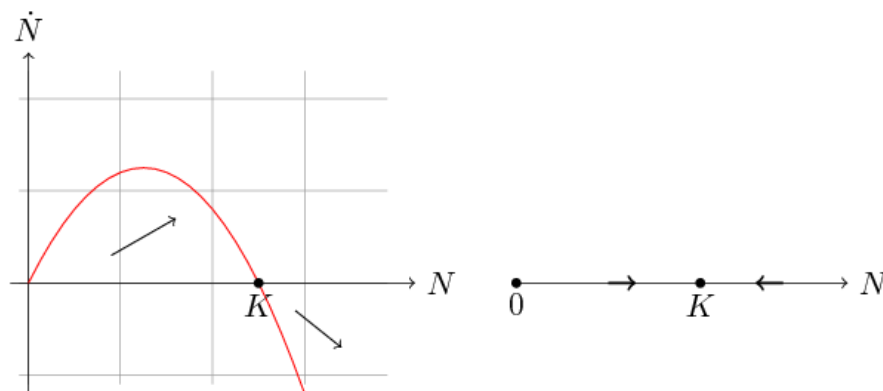
Naszym głównym zadaniem nie jest jednak poszukiwanie analitycznych rozwiązań budowanych modeli, tylko analiza własności i przebiegu tych rozwiązań. Zauważmy, że prawa strona równania (2.2) jest funkcją klasy \mathbf{C}^1 , co implikuje istnienie i jednoznaczność rozwiązań. Łatwo też sprawdzić, że dla $N_0 \geq 0$ mamy $N(t) \geq 0$ dla $t \geq 0$. Jedną z metod tej wykazania tej własności stanowi zapisanie równania (2.2) w równoważnej postaci całkowej

$$\frac{\dot{N}}{N} = r \left(1 - \frac{N}{K}\right) \implies \int_{N_0}^{N(t)} \frac{dN}{N} = \int_0^t r \left(1 - \frac{N(s)}{K}\right) ds \implies N(t) = N_0 e^{r \int_0^t \left(1 - \frac{N(s)}{K}\right) ds},$$

skąd od razu wynika nieujemność. Z kolei nieujemność implikuje oszacowanie

$$\dot{N}(t) \leq rN(t) \implies N(t) \leq N_0 e^{rt},$$

czyli wzrost jest co najwyżej wykładniczy i rozwiązanie istnieje dla wszystkich $t \geq 0$.



Rysunek 2.1. Wykres funkcji $f(N) = rN \left(1 - \frac{N}{K}\right)$ (po lewej) oraz odpowiadający mu portret fazowy równania (2.2) (po prawej).

Zbadamy teraz zależność (N, \dot{N}) , która determinuje monotoniczność rozwiązań. Przyjrzyjmy się wykresowi funkcji $f(N) = rN \left(1 - \frac{N}{K}\right)$, por. rys. 2.1. Jest to parabola o miejscach zerowych $N = 0$ i $N = K$ skierowana w dół. Wynika stąd, że $N \equiv 0$ i $\bar{N} \equiv K$ są rozwiązaniami stacjonarnymi. Jeśli $N(t) \in (0, K)$, to $\dot{N}(t) > 0$, czyli rozwiązania dla warunku początkowego

$N_0 < K$ są rosnące i ograniczone, ponieważ z jednoznaczności rozwiązań wynika, że nie mogą przeciąć stacjonarnego rozwiązania \bar{N} . Zatem jeśli $0 < N_0 < K$, to

$$\lim_{t \rightarrow \infty} N(t) = K.$$

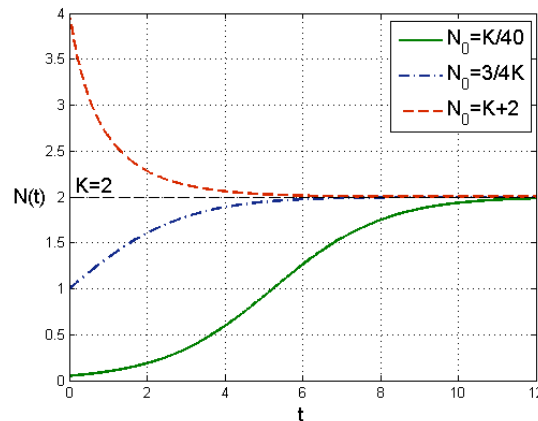
Analogicznie jeśli $N_0 > K$ (rozwijająca się w naturalnych warunkach populacja nie może przekroczyć granicy K , jak pokazaliśmy powyżej, czyli taka sytuacja jest możliwa jedynie w przypadku introdukcji osobników), to $\dot{N}(t) < 0$, więc dla takich wartości początkowych liczebność maleje i ponownie asymptotycznie osiąga wartość K . Jednak aby dokładniej przeanalizować przebieg rozwiązań musimy znać nie tylko wartość pierwszej pochodnej, ale także drugiej, która determinuje wypukłość/wklęsłość funkcji $N(t)$.

Obliczmy zatem

$$\begin{aligned} \ddot{N}(t) &= \frac{d}{dt} \dot{N}(t) = \frac{d}{dt} \left(rN(t) \left(1 - \frac{N(t)}{K} \right) \right) = \\ &= r\dot{N}(t) \left(1 - \frac{N(t)}{K} \right) + rN(t) \left(-\frac{\dot{N}(t)}{K} \right) = r\dot{N}(t) \left(1 - \frac{2N(t)}{K} \right). \end{aligned}$$

Widzimy, że jeśli rozwiązanie w pewnej chwili t_p przyjmuje wartość $N(t_p) = \frac{K}{2}$, to zmienia się charakter przebiegu — funkcja $N(t)$ ma punkt przegięcia. Analizując znaki czynników $\dot{N}(t)$ i $1 - \frac{2N(t)}{K}$ możemy wydzielić kilka obszarów, por. rys. 2.2

- $N_0 \in \left(0, \frac{K}{2}\right)$, to $N(t)$ rośnie do K przy $t \rightarrow \infty$, więc istnieje t_p , takie że $\ddot{N}(t_p) = 0$ i $N(t)$ jest wypukła dla $t < t_p$ (szybki wzrost, porównywalny z wykładniczym, dla małych zagęszczeń), a wklęsła dla $t > t_p$ (następuje spowolnienie wzrostu dla większych zagęszczeń);
- $N_0 \in \left(\frac{K}{2}, K\right)$, wtedy funkcja $N(t)$ rośnie, ale wzrost jest powolny, funkcja N jest wklęsła;
- $N_0 > K$, co (należy ponownie podkreślić) odpowiada introdukcji osobników, wtedy $N(t)$ maleje w sposób wypukły do K , zatem spadek liczebności jest dość szybki.



Rysunek 2.2. Przykładowe rozwiązania równania logistycznego (2.2) dla różnych warunków początkowych.

Widzimy więc, że dla $N_0 \in \left(0, \frac{K}{2}\right)$ rozwiązanie ma kształt wydłużonej litery S, jak krzywa logistyczna, która dobrze odzwierciedla wiele procesów, przebiegających w początkowej fazie szybko i intensywnie, a po pewnym czasie ulegających wysyceniu i stabilizujących się na maksymalnym dla danego procesu poziomie.

2.2. Dyskretne równanie logistyczne

Często równoległe z modelem ciągłym rozważa się model dyskretny. Przedstawimy teraz dyskretną wersję równania (2.2). W tym celu przybliżymy pochodną \dot{N} ilorazem różnicowym

$$\dot{N}(t) \approx \frac{N(t + \Delta t) - N(t)}{\Delta t},$$

z założymy (analogicznie jak w przypadku modelu Malthusa), że Δt jest równa hipotetycznej jednostce czasu $\Delta t = 1$ i zastosujemy notację dyskretną $N(t) = N_t$, $t \in \mathbb{N}$. Otrzymujemy

$$N_{t+1} = (1 + r)N_t - \frac{r}{K}N_t^2 = aN_t \left(1 - \frac{N_t}{K_1}\right), \quad (2.3)$$

gdzie $a = 1 + r$ i $K_1 = \frac{1+r}{r}K$. Zauważmy, że o ile z biologicznego punktu widzenia mamy w równaniu logistycznym dwa parametry: współczynnik rozrodczości i pojemność siedliska, to z matematycznego punktu widzenia jeden z nich możemy z łatwością wyeliminować. Niech $x_t = \frac{N_t}{K_1}$. Dzieląc równanie (2.3) stronami przez K_1 dostajemy

$$x_{t+1} = ax_t(1 - x_t), \quad (2.4)$$

czyli model jednoparametrowy. Zauważmy, że model ten ma sens, o ile kolejne wyrazy ciągu $(x_t)_{t=0}^{\infty}$ pozostają w przedziale $[0, 1]$, ponieważ reprezentują one zagęszczenia populacji w proporcji do pojemności siedliska K_1 . Wyrazy te obliczamy jako kolejne iteracje funkcji $F(x) = ax(1 - x)$, która jest funkcją kwadratową z wartością maksymalną $F\left(\frac{1}{2}\right) = \frac{a}{4}$. Zatem aby wszystkie wyrazy pozostawały w przedziale $[0, 1]$ musi zachodzić $0 < a \leq 4$. Będziemy więc badać dynamikę ciągu $(x_t)_{t=0}^{\infty}$ dla $a \in (0, 4]$ i warunku początkowego $x_0 \in [0, 1]$. Oczywiście jeśli $x_0 = 0$ albo $x_0 = 1$, to $x_t = 0$ dla $t \geq 1$, więc interesujące są wartości $x_0 \in (0, 1)$.

Jeśli ciąg $(x_t)_{t=0}^{\infty}$ ma granicę, to równa się ona któremuś ze stanów stacjonarnych. Wyznamy więc te stany. Niech \bar{x} oznacza stan stacjonarny. Wtedy

$$\bar{x} = a\bar{x}(1 - \bar{x}) \implies \bar{x} = 0 \quad \text{albo} \quad \bar{x}_1 = \frac{a-1}{a}.$$

Widzimy więc, że dodatni stan stacjonarny istnieje, o ile $a > 1$. Lokalną stabilność tych stanów badamy za pomocą wartości własnych przekształcenia $dF(x)|_{\bar{x}}$. W naszym przypadku $dF(x) = a(1 - 2x)$. Jeśli $a \in (0, 1)$, to stan stacjonarny jest wyznaczony jednoznacznie $\bar{x} = 0$ i $dF(0) = a$, przy czym nierówność $0 < a < 1$ implikuje lokalną stabilność. Co więcej widzimy, że $x_{t+1} < x_t$, ponieważ $x_{t+1} = F(x_t) < x_t$ dla dowolnego $x_t \in (0, 1)$ przy założeniu $a \leq 1$. Mamy więc ciąg malejący i ograniczony, czyli zbieżny, zatem stan $\bar{x} = 0$ jest globalnie stabilny w zbiorze $[0, 1]$.

Jeśli $a > 1$, to istnieją dwa stany stacjonarne, przy czym $\bar{x} = 0$ staje się niestabilny, gdyż teraz $dF(0) = a > 1$. Dla dodatniego stanu stacjonarnego $\bar{x}_1 = \frac{a-1}{a}$ mamy $dF(\bar{x}_1) = 2 - a$ oraz

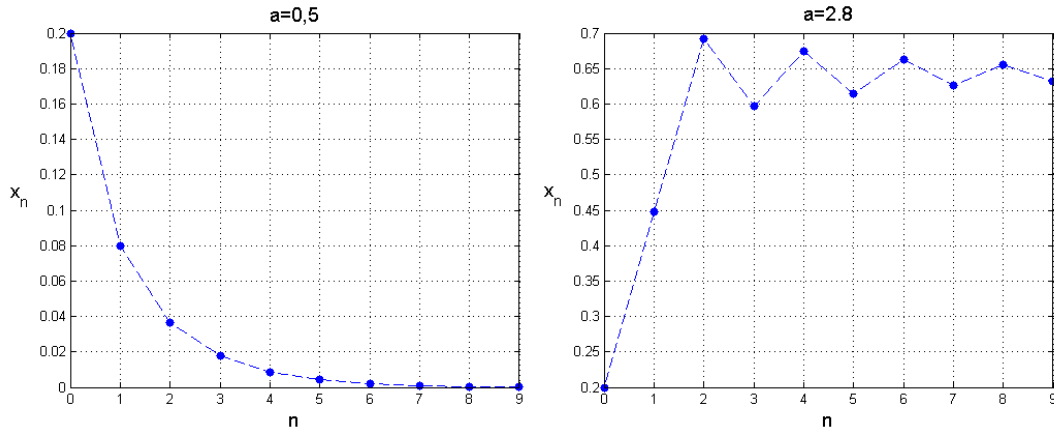
$$|2 - a| < 1 \Leftrightarrow a \in (1, 3).$$

Prostymi metodami, jak w przypadku zerowego stanu stacjonarnego pokazujemy, że dla $a \in (1, 2]$ ciąg $(x_t)_{t=0}^{\infty}$ jest monotoniczny od pewnego miejsca, natomiast dla $a \in (2, 3)$ możemy wyróżnić dwa podciągi monotoniczne od pewnego miejsca. Faktycznie, jeśli $a \in (1, 2]$, to dla $x_t \in (0, \bar{x}_1)$ zachodzi $F(x_t) > x_t$, natomiast dla $x_t \in (\bar{x}_1, 1)$ mamy nierówność przeciwną. Oczywiście wystarczy rozpatrywać $x_t \in \left(0, \frac{1}{2}\right]$, gdyż wykres iterowanej funkcji $F(x)$ jest symetryczny względem prostej $x = \frac{1}{2}$ oraz $F(x) \leq \frac{a}{4} \leq \frac{1}{2}$, więc po pierwszej iteracji wyrazu

$x_0 \in \left(\frac{1}{2}, 1\right)$ dostajemy $x_1 \in \left(0, \frac{1}{2}\right)$. Wystarczy teraz wykazać, że jeśli $x_t < \bar{x}_1$, to $x_{t+1} < \bar{x}_1$. Nierówność $ax(1-x) < \bar{x}_1$ jest równoważna nierówności kwadratowej

$$ax^2 - ax + \frac{a-1}{a} > 0,$$

której wyróżnik $\Delta = (a-2)^2$ i dla $a \in (1, 2]$ mamy dwa pierwiastki \bar{x}_1 oraz $x = \frac{1}{a} \geq \frac{1}{2}$. Wobec tego nierówność ta jest spełniona dla $x \in (0, \bar{x}_1)$. Analogicznie wykazujemy, że $x_t \in \left(\bar{x}_1, \frac{1}{2}\right]$ implikuje $x_{t+1} \in \left(\bar{x}_1, \frac{1}{2}\right]$. Mamy więc ciągi monotoniczne i ograniczone. Stosując tę samą technikę dowodzimy monotoniczności podciągów w przypadku $a \in (2, 3)$.



Rysunek 2.3. Przykładowe rozwiązania równania (2.4) dla $0 < a \leq 1$, czyli gdy $\bar{x} = 0$ jest globalnie asymptotycznie stabilny (lewa strona) oraz dla $a \in (1, 3)$, czyli w przypadku stabilności stanu stacjonarnego $\bar{x}_1 = \frac{a-1}{a}$ (strona prawa).

Dla omówionych wartości parametrów rozwiązania modelu dyskretnego są zbieżne, podobnie jak w przypadku modelu ciągłego, już dla tych wartości widzimy jednak zasadnicze różnice. Po pierwsze — dla małych $a \leq 1$ rozwiązania dążą do 0, czyli populacja wymiera. Nie należy się jednak temu dziwić, gdyż w modelu dyskretnym wartości współczynnika rozrodczości poniżej 1 odpowiadają procesowi śmiertelności, więc w tym przypadku model dyskretny opisuje sytuację, w której populacja wymiera nawet bez konkurencji, zatem jeśli konkurencja występuje, to sytuacja może się tylko pogorszyć. Przy $a \in (1, 2]$ mamy monotoniczną zbieżność do dodatniego stanu stacjonarnego, zatem zachowanie jest analogiczne jak w modelu ciągłym, ale gdy a przekracza 2, pojawiają się gasnące oscylacje — takie zachowanie nie jest możliwe w modelu ciągłym. Co się dzieje, gdy a przekracza 3?

Dla $a > 3$ oba rozwiązania stacjonarne stają się niestabilne. Możemy jednak sprawdzić, że po przekroczeniu tej wartości rozwiązania zyskują charakter stabilnego cyklu granicznego, najpierw o okresie 2, potem o okresie 4 itd. Cykl o okresie 2 znajdujemy stosunkowo łatwo, jako rozwiązanie równania

$$F^2(x) = x, \quad \text{gdzie} \quad F^2(x) = F(F(x)) = a^2x(1-x)(1-ax+ax^2)$$

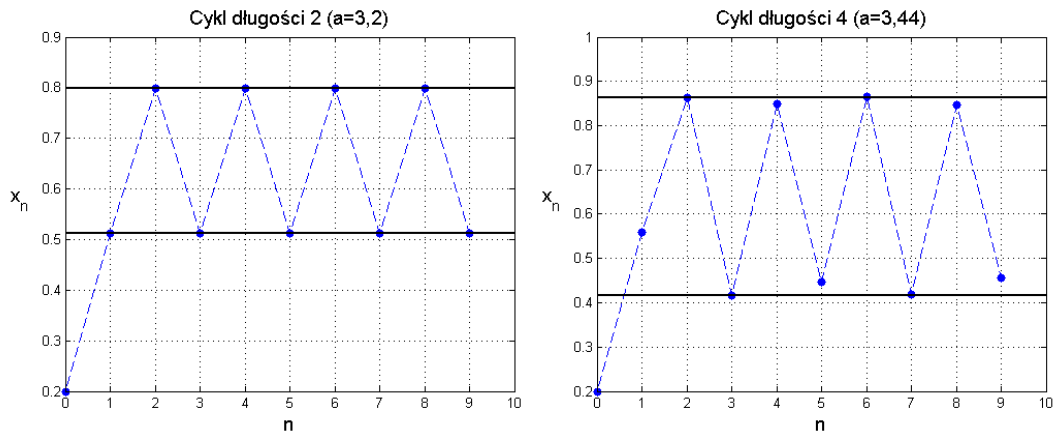
różne od znanych już stanów stacjonarnych. Szukamy więc dodatnich miejsc zerowych wielomianu

$$W(x) = a^3x^3 - 2a^3x^2 + a^2(a+1)x + 1 - a^2.$$

Dzielimy W przez $x - \bar{x}_1$ i obliczamy

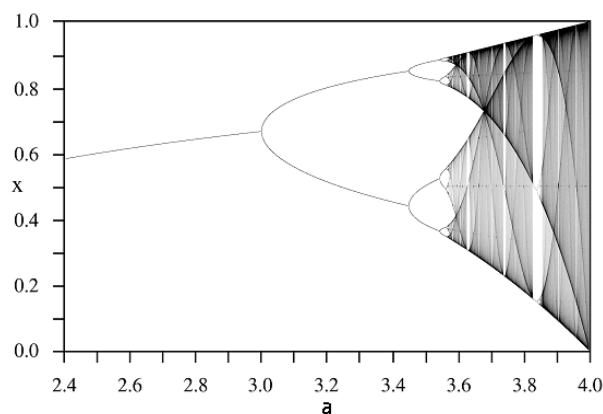
$$c_{1,2} = \frac{a+1 \pm \sqrt{a^2 - 2a - 3}}{2a}, \quad C = \{c_1, c_2\}.$$

Zbiór C jest poszukiwanym cyklem o okresie 2. Jego lokalną stabilność badamy znanymi już metodami. Dla $a \in (3, 1 + \sqrt{6})$ cykl jest stabilny. Gdy a przekracza $1 + \sqrt{6}$ następuje destabilizacja i pojawia się stabilny cykl o okresie 4. Taką zmianę dynamiki rozwiązań nazywamy bifurkacją podwojenia okresu (albo bifurkacją rozwidleniową).



Rysunek 2.4. Przykład cyklicznych rozwiązań równania (2.4) o różnych długościach cyklu.

Wraz z rosnącym a wyczerpują się okresy postaci 2^n . Przy przekroczeniu wartości krytycznej $a^* \approx 3,569$ pojawiają się rozwiązania o innych okresach. Okresy te są związane z [porządkiem Szarkowskiego](#). Na końcu bifurkuje orbita okresowa o okresie 3, która powszechnie kojarzona jest z chaosem. Omawianie pojęcia chaosu wykracza poza ramy tego wykładu. Można jednak udowodnić, że dla $a = 4$ rozwiązania równania (2.4) są chaotyczne. Cykl bifurkacji występujący w tym modelu najlepiej obrazuje diagram bifurkacyjny, zwany w tym przypadku drzewem Feigenbauma, por. rys. 2.5. Drzewo Feigenbauma ma pewne własności obiektów zwanych frak-



Rysunek 2.5. Diagram bifurkacyjny modelu (2.4), zwany drzewem Feigenbauma (*Źródło: wikipedia*).

talami, w szczególności charakteryzuje się samopowtarzalnością — jeśli przybliżymy wycinek wykresu, to zobaczymy fragment ładząco podobny do całości. Podkreślić należy, że tego typu dynamikę generują wszystkie funkcje o własnościach podobnych do funkcji logistycznej, czyli $F : [0, 1] \rightarrow [0, 1]$ unimodalne.

3. Modelowanie pojedynczej populacji III

3.1. Efekt Allego

W większości przypadków, jeśli nie uwzględniamy migracji, to model pojedynczej populacji \mathcal{P} , zarówno w przypadku ciągłym jak i dyskretnym, możemy zapisać w ogólnej postaci

$$\begin{aligned} \dot{N}(t) &= N(t)f(N(t)) \text{ — model ciągły,} \\ N_{t+1} &= N_t f(N_t) \text{ — model dyskretny,} \end{aligned} \tag{3.1}$$

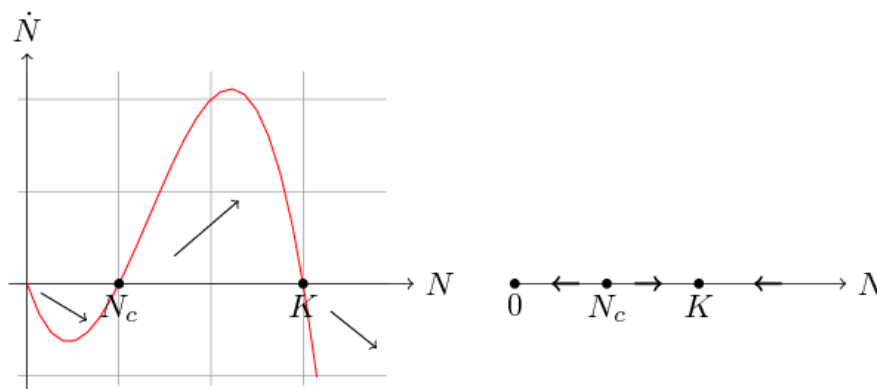
gdzie f odzwierciedla przyrost *per capita*, podobnie jak w przypadku równania logistycznego (2.2). Omówimy teraz postać funkcji f w przypadku występowania w populacji tzw. **efektu Allego**. Populacje, w których obserwujemy taki efekt, zmniejszają swoją liczebność, jeśli spadnie ona poniżej pewnego progu. Typowo wiążemy efekt Allego z drapieżnictwem — na dynamikę danej populacji wpływa także to, że w ekosystemie występuje gatunek drapieżnika żywiący się osobnikami populacji \mathcal{P} , przy czym zakładamy, że drapieżników jest zawsze dużo. Wobec tego jeśli liczebność populacji \mathcal{P} jest niewielka, to drapieżniki zjadają dostępne osobniki i populacja \mathcal{P} wymiera.

Modele z efektem Allego mogą mieć różną postać, w zależności od funkcji drapieżnictwa. Najprostszy model stanowi modyfikację równania logistycznego (2.2)

$$\dot{N}(t) = rN(t) (N(t) - N_c) \left(1 - \frac{N(t)}{K} \right), \quad N_c \in (0, K), \tag{3.2}$$

gdzie r i K interpretujemy analogicznie jak w równaniu logistycznym, natomiast N_c nazywamy w tym kontekście pułapką drapieżnictwa.

Podobnie jak w przypadku równania (2.2), równanie (3.2) możemy rozwiązać analitycznie rozkładając prawą stronę na ułamki proste, co zostawimy jako ćwiczenie. Najważniejsze własności rozwiązań odczytamy z portretu fazowego, por. rys. 3.1.



Rysunek 3.1. Zależność (N, \dot{N}) dla modelu opisanego równaniem (3.2) (po lewej) oraz odpowiadający tej zależności portret fazowy (po prawej).

Widzimy, że prawa strona równania (3.2) jest wielomianem trzeciego stopnia o miejscach zerowych 0 , N_c , K — są to oczywiście stany stacjonarne naszego równania — i ujemnym współczynnikiem przy N^3 . Wobec tego

1. dla $N_0 \in (0, N_c)$ mamy $\dot{N}(t) < 0$, czyli $N(t)$ maleje do 0 ;
2. dla $N_0 \in (N_c, K)$ mamy $\dot{N}(t) > 0$, czyli $N(t)$ rośnie do K ;
3. dla $N_0 > K$ mamy $\dot{N}(t) < 0$, czyli $N(t)$ maleje do K .

Policzmy też drugą pochodną, aby zbadać wklęsłość/wypukłość rozwiązań

$$\ddot{N}(t) = -\frac{r}{K}\dot{N}(t) \left(3N^2(t) - 2N(t)(N_c + K) + N_cK \right).$$

Z kolei druga pochodna jest funkcją kwadratową zmiennej $N(t)$ o wyróżniku

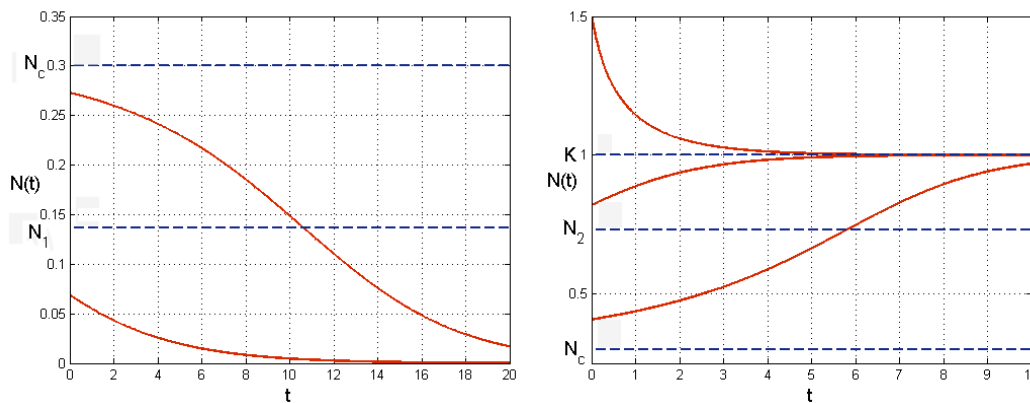
$$\Delta = 4 \left(K^2 - N_cK + N_c^2 \right) > 0$$

i pierwiastkach

$$N_1 = \frac{N_c + K - \sqrt{\Delta/4}}{3} \in (0, N_c), \quad N_2 = \frac{N_c + K + \sqrt{\Delta/4}}{3} \in (N_c, K).$$

Ostatecznie, por. rys. 3.2,

1. dla $N_0 < N_1$ funkcja N maleje i $\ddot{N} > 0$, czyli N jest wypukła;
2. dla $N_0 \in (N_1, N_c)$ funkcja N maleje i pozostaje wklęsła dopóki $N(t) > N_1$, w punkcie \bar{t} , takim że $N(\bar{t}) = N_1$ następuje zmiana charakteru przebiegu funkcji N — staje się ona wypukła;
3. dla $N_0 \in (N_c, N_2)$ przebieg jest symetryczny względem stanu stacjonarnego N_c do opisanego w poprzednim punkcie (podobny do przebiegu krzywych logistycznych);
4. dla $N_0 \in (N_2, K)$ funkcja N rośnie w sposób wklęsły do stanu stacjonarnego K ;
5. dla $N_0 > K$ funkcja N maleje w sposób wypukły.



Rysunek 3.2. Przykładowe wykresy rozwiązań równania (3.2) dla różnych wartości początkowych N_0 .

Widzimy więc, że powyżej progu drapieżnictwa dynamika modelu jest analogiczna jak w przypadku równania logistycznego, por. prawy wykres na rys. 3.2, natomiast poniżej progu drapieżnictwa populacja wymiera, co widzimy na lewym wykresie na rys. 3.2.

3.2. Funkcjonalna odpowiedź Hollinga i funkcja Hilla

W dalszym ciągu będziemy zajmować się sytuacją, w której na opisywaną populację \mathcal{P} ma wpływ drapieżnictwo. Wyprowadzimy wzór funkcji zwanej odpowiedzią funkcjonalną na obecność drapieżnika w oparciu o następujące rozumowanie zwane koncepcją dysku Hollinga [5]. Według tej koncepcji każdy drapieżnik polujący na danej powierzchni zauważa ofiary znajdujące się w obszarze o pewnym określonym promieniu γ i promień ten jest charakterystyczny dla konkretnego gatunku drapieżnika. Spośród dostrzeżonych ofiar drapieżnik może upolować i zjeść pewną część. Zakładamy, że jest ona stała i oznaczamy tę część przez α . Każdy drapieżnik dzieli swój czas na dwa typy aktywności: wyszukiwanie ofiary oraz polowanie i zjedanie ofiary. Szukając ofiar drapieżnik przemieszcza się z pewną ustaloną prędkością v , a polowanie i konsumpcja jednej ofiary zajmują mu średnio T_k jednostek czasu.

Przy tych założeniach oszacujemy, ile średnio ofiar z populacji o zagęszczeniu N zjada w jednostce czasu pojedynczy drapieżnik. Dowolny przedział czasu długości T dzielimy na wymienione wyżej dwa typy aktywności, a za pomocą T_s oznaczmy czas przeznaczony na wyszukiwanie ofiary. W czasie wyszukiwania drapieżnik poruszający się z prędkością v obejmuje swoim zasięgiem powierzchnię $2\gamma v T_s$, gdyż $v T_s$ to droga, którą przebywa drapieżnik, przy czym w każdym kroku rozglądając się widzi ofiary w odległości nie większej niż γ w prawo i lewo. Zatem w przedziale czasu o długości T drapieżnik chwytą i zjada $2\alpha\gamma v T_s N$ ofiar w czasie $2\alpha\gamma v T_s T_k N$. Cały przedział czasu T możemy więc zapisać jako

$$T = T_s + 2\alpha\gamma v T_s T_k N.$$

Jeśli przez f_H oznaczmy funkcję odpowiedzi drapieżnika, czyli

$$f_H = \frac{\text{liczba ofiar zjedzona w czasie } T}{T},$$

to

$$f_H = \frac{2\alpha\gamma v T_s N}{T_s(1 + 2\alpha\gamma v T_k N)}$$

i ostatecznie

$$f_H = \frac{aN}{1 + bN}, \quad (3.3)$$

gdzie $a = 2\alpha\gamma v$ i $b = 2\alpha\gamma v T_k$, a równanie (3.3) nazywamy równaniem dysku Hollinga.

Zauważmy, że

$$\lim_{N \rightarrow \infty} f_H = \frac{1}{T_k},$$

czyli nawet jeśli ofiar jest bardzo dużo, to drapieżnik nie może zjeść więcej niż pewną ich liczbę, określoną przez czas poświęcany na upolowanie i zjedzenie jednej ofiary. Jeśli rozpatrzmy skrajny przypadek, w którym $T_k \rightarrow 0$ przy ustalonej liczebności populacji ofiar, to

$$\lim_{T_k \rightarrow 0} f_H = aN,$$

co oznacza, że w takim przypadku możemy odpowiedź funkcjonalną modelować za pomocą składnika liniowego. Taką postać odpowiedzi funkcjonalnej nazywamy odpowiedzią **Hollinga I typu**, która charakteryzuje się brakiem ograniczenia względem liczebności populacji ofiar, natomiast przy $b \neq 0$ mamy odpowiedź typu II, która jest ograniczona.

Wróćmy teraz do opisu dynamiki populacji \mathcal{P} . Załóżmy, że na jej liczebność wpływa drapieżnictwo, odpowiedź funkcjonalną opisuje funkcja (3.3) i możemy przyjąć stałą liczebność

drapieżników D . Jeśli wewnętrzną dynamikę (bez uwzględnienia drapieżnictwa) populacji \mathcal{P} opisuje równanie logistyczne (2.2), to ostatecznie zmiana liczebności podlega równaniu

$$\dot{N}(t) = rN(t) \left(1 - \frac{N(t)}{K}\right) - \beta \frac{N(t)}{N_{1/2} + N(t)}, \quad (3.4)$$

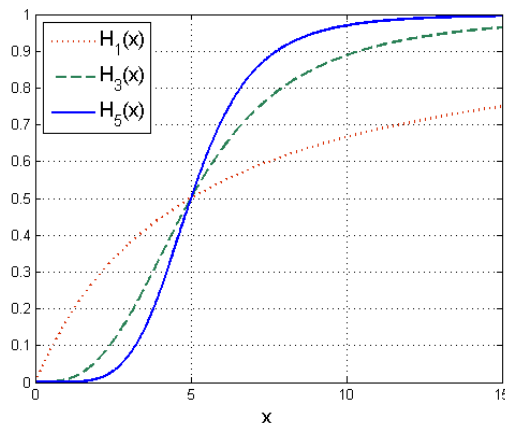
gdzie $\beta = \frac{D}{T_k}$ oznacza maksymalne drapieżnictwo i $N_{1/2} = \frac{1}{aT_k}$ odpowiada takiej liczebności populacji, przy której drapieżnictwo jest równe połowie maksymalnej wielkości β (maksymalną wielkość osiąga się asymptotycznie, przy $N \rightarrow \infty$).

Postać równania (3.4) nie daje gwarancji znalezienia rozwiązania analitycznego, ale metody portretu fazowego i badania wklęsłości/wypukłości rozwiązań omówione w kontekście poprzednich modeli mogą być zastosowane z powodzeniem także w tym przypadku. Przeanalizowanie dynamiki modelu (3.4) pozostawiamy jako ćwiczenie.

Wyprowadzona powyżej postać odpowiedzi funkcjonalnej (3.3) stanowi szczególny przypadek funkcji Hilla

$$H_n(x) = a \frac{x^n}{b^n + x^n}, \quad n > 0, \quad (3.5)$$

przy czym stałą n nazywamy współczynnikiem Hilla i w ogólnym przypadku nie musi to być liczba naturalna.

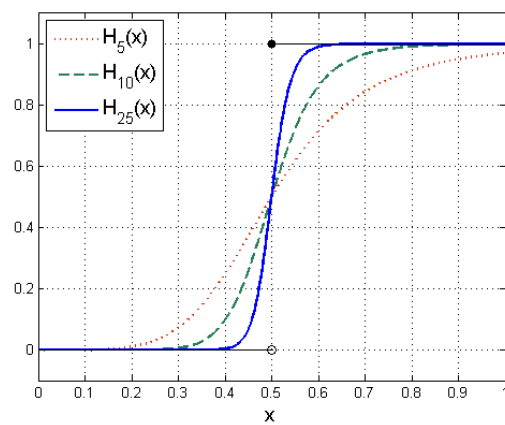


Rysunek 3.3. Porównanie wykresu funkcji Hilla dla różnych wartości parametru n przy ustalonych parametrach a i b .

Widzimy, że dla $n = 1$ funkcja Hilla odpowiada funkcjonalnej odpowiedzi Hollinga typu II. Funkcja Hilla została zaproponowana do opisu reakcji biochemicznych i w tym przypadku odzwierciedla ona szybkość reakcji. W szczególności $n = 2$ opisuje przypadek, gdy w reakcji biorą udział bimery (cząsteczki powstałe z połączenia dwóch pojedynczych monomerów). Zauważmy, że funkcja H_n opisuje reakcje, które ulegają wysyceniu, przy czym w zależności od współczynnika n mają różne tempo przebiegu reakcji dla x blisko 0 i dla dużych x . Im większe n , tym szybsze są zmiany w otoczeniu punktu krytycznego $x_p = \sqrt[n]{\frac{n-1}{n+1}}b$ (będącego punktem przegięcia H_n). Funkcje Hilla możemy także interpretować jako gładkie przybliżenie tzw. funkcji przełączeniowych

$$f_p(x) = \begin{cases} 0 & \text{dla } x \in [0, p), \\ b & \text{dla } x \geq p. \end{cases}$$

Funkcje Hilla są bardzo często wykorzystywane do opisu różnych procesów ulegających wysyceniu — wielkość współczynnika Hilla zależy wtedy od szybkości zmian danego procesu.



Rysunek 3.4. Wykres ilustrujący jak ze wzrostem n funkcja Hilla z odpowiednimi parametrami a i b coraz lepiej przybliża ustaloną funkcję przelączeniową.

4. Modele pojedynczej populacji z uwzględnieniem wieku I

Przejdziemy teraz do opisu sytuacji, gdy populacja nie jest jednorodna — rozróżniamy osobniki w różnym wieku.

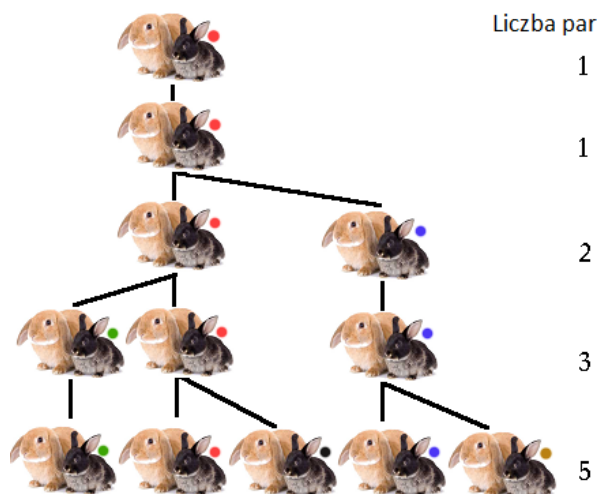
4.1. Ciąg Fibonacciego

W najprostszym przypadku w populacji wyodrębniamy dwie grupy wiekowe:
— osobników niedojrzałych (w wieku przed reprodukcyjnym);
— osobników dojrzałych (w wieku reprodukcyjnym).
Co ciekawe, model opisujący taką sytuację jest najstarszym znanym modelem populacyjnym i znany go pod nazwą **ciągu Fibonacciego**.

W I połowie XIII wieku **Leonardo z Pizy** (Fibonacci, czyli „*filus Bonacci*” — syn Bonacciego) zastosował ten ciąg do opisu następującego zagadnienia populacyjnego.

„Pewien człowiek wziął parę królików i umieścił je w miejscu otoczonym ze wszystkich stron murem. Ile par królików urodzi się z tej pary w ciągu roku, jeśli założymy, że z każdej pary po miesiącu rodzi się nowa para, która staje się płodna po upływie kolejnego miesiąca?”

Liber abaci rozdział III.



Rysunek 4.1. Ilustracja zagadnienia populacyjnego opisanego przez Leonarda z Pizy.

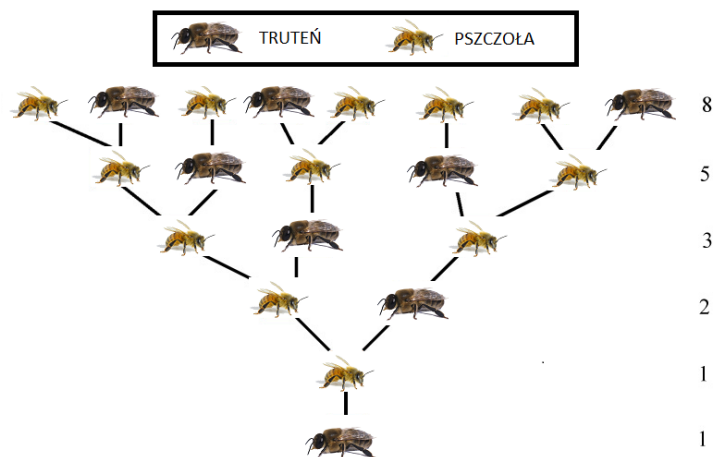
Jako rozwiązanie tego zagadnienia Fibonacci zaproponował ciąg:

$$1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144 \dots,$$

znany dziś właśnie jako ciąg Fibonacciego. Ogólny wyraz tego ciągu opisujemy za pomocą formuły rekurencyjnej

$$F_{n+1} = F_n + F_{n-1}, \quad F_1 = 1, \quad F_2 = 1. \quad (4.1)$$

W *Liber abaci* Fibonacci pominął pierwszy wyraz ciągu — zauważmy, że powinny być dwa wyrazy równe 1, gdyż pierwsza para też staje się płodna dopiero po upływie miesiąca, por. rys. 4.1. Analogiczny schemat opisuje liczbę przodków pojedynczego trutnia w roju pszczół, por. rys. 4.2.



Rysunek 4.2. Ilustracja przedstawiająca schemat opisujący liczbę przodków pojedynczego trutnia w roju pszczół.

Nazwijmy kolejne wyrazy ciągu Fibonacciego liczbami Fibonacciego. Okazuje się, że liczby Fibonacciego opisują różne wielkości przyrodnicze, jak np. liczbę płatków kwiatów (stokrotki stanowią sztandarowy przykład, typowo liczba płatków równa jest 34, 55 albo 89!), liczbę pędów roślin w kolejnych fazach wzrostu, czy też liczbę spiral w różnych konstrukcjach spiralnych (spirale prawoskrętne i lewoskrętne), takich jak kwiatostany słonecznika, owoc ananasa czy szyszki, por. rys. 4.3.



Rysunek 4.3. Zdjęcia przedstawiające wielkości przyrodnicze, które są opisywane przez liczby Fibonacciego. Od lewej: liczba płatków stokrotki; liczba pędów roślin; liczba spiral w kwiatostanie słonecznika.

Równanie (4.1) jest liniowym równaniem dyskretnym, przy czym możemy je traktować jako równanie z opóźnionym argumentem — typowo dyskretny układ dynamiczny bez opóźnienia zapisujemy w postaci

$$X_{n+1} = G(X_n), \quad (4.2)$$

gdzie X_n opisuje liczebność populacji (czy inną wielkość biofizyczną) w chwili n , a G odzwierciedla prawa rządzące dynamiką tej populacji. W ogólnym przypadku $X_n \in \mathbb{R}^n$. Oczywiście postać funkcji G zależy od modelu heurystycznego (np. dla modelu Malthusa (1.3) funkcja G jest liniowa) i w istotny sposób wpływa na przebieg rozwiązań, czyli dynamikę ciągu $(X_n)_{n=1}^{\infty}$. Dys-

kretny układ z opóźnieniem zawsze możemy sprowadzić do układu postaci (4.2) wprowadzając pomocnicze zmienne. Przyjrzyjmy się tej procedurze na przykładzie ciągu Fibonacciego (4.1). Niech Y_n oznacza wektor kolejnych liczebności $Y_n = (X_n, X_{n-1})^T$. Wtedy

$$Y_{n+1} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} Y_n,$$

zatem jedno równanie z opóźnieniem rzędu jednej jednostki czasu zamienia się na układ dwóch równań postaci (4.2). Liczba równań w układzie wyjściowym zależy od wielkości opóźnień w układzie początkowym.

Jak już przypominaliśmy przy rozwiązywaniu równania (1.3), rozwiązań układu (4.2) dla liniowej funkcji G szukamy w postaci $X_n = \lambda^n X_0$, gdzie λ jest wartością własną macierzy przekształcenia G . Podstawiając taką postać rozwiązania do zależności rekurencyjnej (4.1) otrzymujemy

$$\lambda^{n+2} X_0 = \lambda^{n+1} X_0 + \lambda^n X_0,$$

zatem λ spełnia równanie

$$\lambda^2 - \lambda - 1 = 0, \quad (4.3)$$

czyli równanie charakterystyczne dla macierzy $\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$. Rozwiązując równanie (4.3) dostajemy

$$\lambda_1 = \frac{1 + \sqrt{5}}{2}, \quad \lambda_2 = \frac{1 - \sqrt{5}}{2}.$$

Wobec tego wyraz ogólny ciągu Fibonacciego obliczamy jako kombinację liniową

$$F_n = a\lambda_1^n + b\lambda_2^n, \quad (4.4)$$

gdzie a i b musimy wyznaczyć korzystając z warunku początkowego $F_1 = F_2 = 1$. Dostajemy zatem układ równań

$$\begin{aligned} a(1 + \sqrt{5}) + b(1 - \sqrt{5}) &= 2, \\ a(1 + \sqrt{5})^2 + b(1 - \sqrt{5})^2 &= 4, \end{aligned}$$

skąd obliczamy $a = \frac{\sqrt{5}}{5} = -b$ i ostatecznie ogólny wyraz ciągu Fibonacciego

$$F_n = \frac{\sqrt{5}}{5} \cdot \frac{(1 + \sqrt{5})^n - (1 - \sqrt{5})^n}{2^n}. \quad (4.5)$$

Zauważmy, że ciąg $(F_n)_{n=1}^\infty$ jest rosnący i $F_n \rightarrow \infty$ przy $n \rightarrow \infty$. Co więcej, tę samą własność mają wszystkie ciągi określone formułą rekurencyjną (4.1) bez względu na warunek początkowy, o ile tylko $F_1, F_2 > 0$. Jeśli wzór (4.5) przepiszemy w postaci

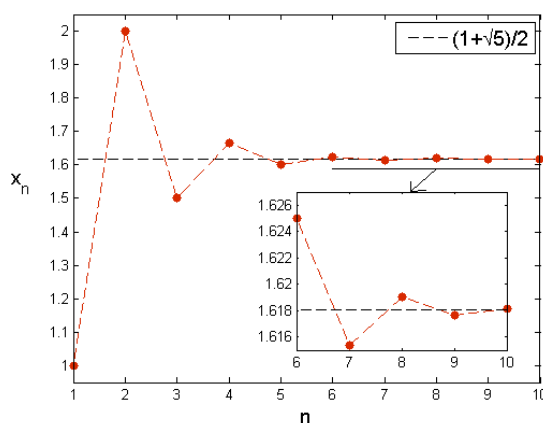
$$F_n = \frac{\sqrt{5}}{5} \left(\frac{1 + \sqrt{5}}{2} \right)^n \left(1 - \left(\frac{1 - \sqrt{5}}{1 + \sqrt{5}} \right)^n \right),$$

to możemy zauważyć, że wraz z rosnącym n dynamika ciągu $(F_n)_{n=1}^\infty$ jest coraz bliższa dynamice ciągu $\left(\frac{\sqrt{5}}{5} \left(\frac{1 + \sqrt{5}}{2} \right)^n \right)_{n=1}^\infty$ determinowanej przez dominującą wartość własną λ_1 . Wnioskujemy więc, że wraz z rosnącym n dynamika ta coraz mniej różni się od dynamiki dyskretnego modelu Malthusa (1.3). Podobne zachowanie zaobserwujemy dla większości wartości początkowych $F_1, F_2 > 0$.

Widzimy więc, że dynamika ciągu Fibonacciego jest raczej „uboga”. Spójrzmy jednak na ciąg $\frac{F_n}{F_{n-1}}$. Zdefiniujmy

$$x_n = \frac{F_n}{F_{n-1}} = \frac{F_{n-1} + F_{n-2}}{F_{n-1}} = 1 + \frac{1}{\frac{F_{n-1}}{F_{n-2}}} = 1 + \frac{1}{x_{n-1}}.$$

Możemy sprawdzić, że ciąg $(x_n)_{n=1}^{\infty}$ oscyluje i dzieli się na dwa podciągi $(x_{2k})_{k=1}^{\infty}$, $(x_{2k+1})_{k=1}^{\infty}$, monotoniczne, zbieżne do tej samej liczby $\frac{1+\sqrt{5}}{2}$, z czego wynika, że cały ciąg zbiega do tej liczby.



Rysunek 4.4. Początkowe wyrazy ciągu $x_n = \frac{F_n}{F_{n-1}}$.

Liczba ta znana jest od starożytności, oznaczamy ją Φ i definiuje ona tzw. **złotą proporcję** lub **złoty podział**. Nazwa ta wiąże się z podziałem odcinka na dwie części, z których dłuższa do krótszej ma się tak jak cały odcinek do dłuższej. W starożytności mówiono także o „boskiej proporcji” i wykorzystywano ją w wielu różnych konstrukcjach, także przy budowie Partenonu. Również w czasach nowożytnych wielu ludzi doszukuje się liczb Fibonacciego czy złotej proporcji w różnych zjawiskach, najczęściej jednak nie ma to prawie żadnych podstaw naukowych.

4.2. Macierze Lesliego

W ogólnym przypadku możemy wyodrębnić więcej grup wiekowych. Niech populacja \mathcal{P} będzie podzielona na k grup wiekowych. W tej sytuacji liczebność całej populacji jest opisana za pomocą wektora $\mathbf{N}_t = (N_t^0, N_t^1, \dots, N_t^k)^T$, gdzie N_t^i oznacza liczebność grupy wiekowej i w chwili t , dla $i = 0$ mamy osobniki nowonarodzone, a dla $i = k$ — osobniki najstarsze, które nie mają szansy dożyć do następnej chwili. Podobnie jak w przypadku modelu Malthusa chcemy na podstawie znajomości wektora liczebności \mathbf{N}_t określić \mathbf{N}_{t+1} . Jak zwykle musimy zacząć od modelu heurystycznego, czyli omówienia procesów, które są według nas istotne przy opisie dynamiki populacji \mathcal{P} . Ponownie podobnie jak w modelu Malthusa zakładamy, że osobniki są jednorodnie w obrębie każdej grupy wiekowej, w każdej grupie wiekowej mamy do czynienia tylko z dwoma procesami — rozrodczości i starzenia się/śmiertelności, przy czym współczynniki rozrodczości i starzenia się są charakterystyczne dla danej grupy wiekowej. Analogicznie jak w przypadku modelu Malthusa w obrębie danej grupy wiekowej oba procesy przebiegają tak samo dla wszystkich osobników i są rozłożone równomiernie w czasie. Zakładamy też, że jednostka czasu jest równa jednostce zmiany wieku (starzenia się), czyli że po upływie np. jednego roku

osobnik starzeje się o rok, co więcej po upływie jednej jednostki czasu przechodzi do następnej grupy wiekowej albo umiera. Dostajemy więc:

$$N_{t+1}^{i+1} = \gamma_i N_t^i, \quad i \in \mathbb{N}, \quad N_{t+1}^0 = \sum_{i=0}^k \alpha_i N_t^i,$$

gdzie γ_i odzwierciedla przeżywalność osobników z grupy wiekowej i (zatem $1 - \gamma_i$ to śmiertelność), a α_i — rozrodczość danej grupy wiekowej. W związku z tym formalnie model Lesliego można przedstawić jako trójkę $(\mathbf{N}_0, \alpha, \eta)$, gdzie wektory $\mathbf{N}_0 = (N_0^0, \dots, N_0^k)^T$, $\alpha = (\alpha_0, \dots, \alpha_k)$, $\eta = (\eta_0, \dots, \eta_{k-1})$ oznaczają odpowiednio liczebności początkowe, współczynniki rozrodzości oraz śmiertelności, $\eta_i = 1 - \gamma_i$.

Zapisując ten model w postaci (4.2) otrzymujemy

$$\mathbf{N}_{t+1} = \mathbf{M}\mathbf{N}_t, \quad (4.6)$$

gdzie

$$\mathbf{M} = \begin{pmatrix} \alpha_0 & \alpha_1 & \alpha_2 & \dots & \alpha_k \\ \gamma_0 & 0 & 0 & \dots & 0 \\ 0 & \gamma_1 & 0 & \dots & 0 \\ \vdots & 0 & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & \gamma_{k-1} & 0 \end{pmatrix},$$

przy czym macierz \mathbf{M} nazywamy **macierzą Lesliego**, a model (4.6) modelem (macierzowym) Lesliego (patrz [8, 9, 2]). Jeśli mamy zadany początkowy rozkład wieku \mathbf{N}_0 , to rozkład w dowolnej chwili $t \in \mathbb{N}$ możemy obliczyć ze wzoru

$$\mathbf{N}_t = \mathbf{M}^t \mathbf{N}_0.$$

Dla dowolnych $t, p \in \mathbb{N}$ zachodzi też $\mathbf{N}_{t+p} = \mathbf{M}^p \mathbf{N}_t$.

Definicja 4.1. Stanem równowagi modelu Lesliego nazwiemy parę $(\bar{\mathbf{N}}, r)$ spełniającą zależność

$$r\bar{\mathbf{N}} = \mathbf{M}\bar{\mathbf{N}},$$

gdzie $\bar{\mathbf{N}}$ jest wektorem liczebności, który nazywamy ustaloną strukturą wieku, a $r \in \mathbb{R}$ oznacza stały współczynnik wzrostu.

W przypadku macierzy Lesliego liniowa niezależność wierszy i kolumn pozwala zapisać \mathbf{M} w postaci

$$\mathbf{M} = \mathbf{V}\Lambda\mathbf{V}^{-1},$$

gdzie Λ to diagonalna macierz wartości własnych, a \mathbf{V} macierz kolumnowych wektorów własnych V_0, \dots, V_k . Wynika stąd, że

$$\mathbf{M}^t = (\mathbf{V}\Lambda\mathbf{V}^{-1})^t = \mathbf{V}\Lambda^t\mathbf{V}^{-1} \implies \mathbf{N}_t = \mathbf{V}\Lambda^t\mathbf{V}^{-1}\mathbf{N}_0.$$

Oznaczmy $W = \mathbf{V}^{-1}\mathbf{N}_0$, $W = (w_0, \dots, w_k)$. Wtedy

$$\mathbf{N}_t = \sum_{i=0}^k V_i \lambda_i^t w_i. \quad (4.7)$$

Własności rozwiązań zależą zarówno od postaci macierzy \mathbf{M} jak i początkowego rozkładu wieku. **Twierdzenie Frobeniusa – Perrona** gwarantuje istnienie dominującej rzeczywistej wartości własnej $\lambda_0 \geq |\lambda_i|$, $i = 1 \dots, k$. Dla tej wartości własnej zarówno prawo jak i lewostronny wektor własny jest rzeczywisty i ma nieujemne współrzędne. Jeśli macierz \mathbf{M} spełnia pewne dodatkowe założenia, np. gdy wskaźniki i , dla których $\alpha_i > 0$ nie mają większego wspólnego dzielnika niż 1, to $\lambda_0 > |\lambda_i|$, $i = 1, \dots, k$ i odpowiadający jej wektor własny $\mathbf{M}V_0 = \lambda_0 V_0$ ma współrzędne nieujemne i dla większości warunków początkowych asymptotyczne zachowanie rozwiązań jest determinowane przez $aV\lambda_0^t$, gdzie $a = \text{const}$ jest stałą zależną od warunku początkowego, ponieważ ze wzoru (4.7) dla dominującej wartości własnej λ_0 wynika rozkład asymptotyczny

$$\frac{1}{\lambda_0^t} \mathbf{N}_t = V_0 w_0 + \sum_{i=1}^k V_i \left(\frac{\lambda_i}{\lambda_0} \right)^t w_i \rightarrow V_0 w_0 \quad \text{przy } t \rightarrow \infty.$$

Wobec tego dla większości warunków początkowych populacja asymptotycznie osiąga rozkład wieku V_0 przy współczynniku rozrodczości λ_0 , gdyż

$$\mathbf{N}_t \approx w_0 V_0 \lambda_0^t \quad \text{przy } t \rightarrow \infty$$

i jeśli

1. $\lambda_0 > 1$, to liczebności wszystkich grup wiekowych rosną nieograniczenie, mówimy że populacja jest rozwojowa;
2. $\lambda_0 < 1$, to populacja jest wymierająca;
3. $\lambda_0 = 1$, to liczebność całej populacji i poszczególnych grup wiekowych stabilizuje się — populacja jest stacjonarna, por rys. 4.5.

Wektor V_0 odzwierciedla rozkład liczebności populacji w stanie równowagi, przy czym jeśli $\|V_0\| = 1$, to współrzędne tego wektora odzwierciedlają procentowy wkład poszczególnych grup wiekowych w populację. W takiej sytuacji dynamika modelu Lesliego jest podobna do dynamiki modelu Malthusa, co więcej, jeśli rozpatrzmy całkowitą liczebność populacji $N_t = \sum_{i=0}^k N_t^i$, to $N_t \sim \lambda_0^t N_0$.

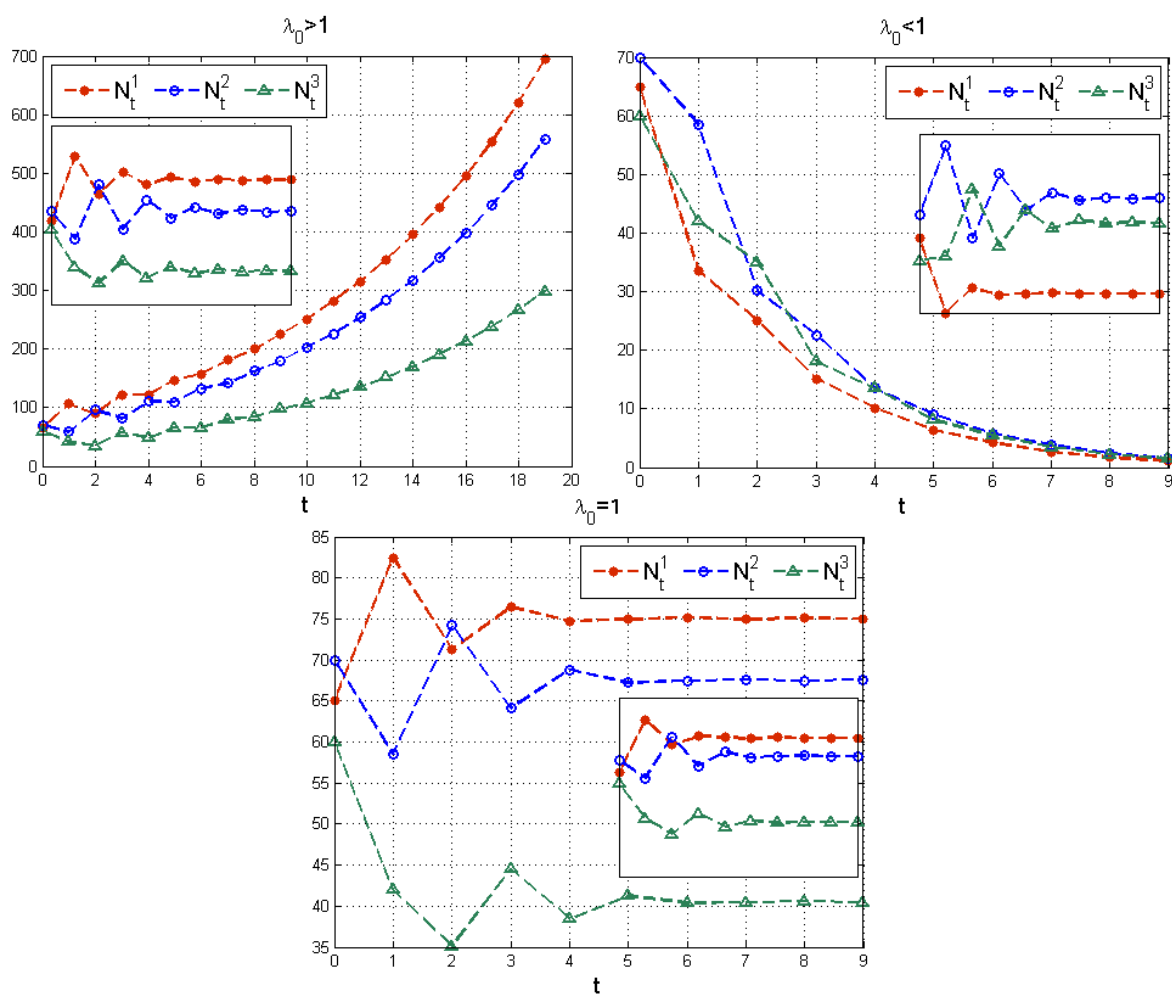
Możliwe jest jednak także inne zachowanie rozwiązań, w szczególności cykliczne zmiany struktury wieku. Wiąże się to między innymi z dynamiką determinowaną przez początkowe rozkłady wieku należące do podprzestrzeni generowanych przez wektory własne odpowiadające wartościom własnym $\lambda_1, \dots, \lambda_k$ w sytuacji, gdy wśród tych wartości własnych nie ma już kolejnej dominującej. Oczywiście zależy to w istotny sposób od postaci macierzy Lesliego. Rozpatrzmy przykład macierzy Lesliego odzwierciedlającej sytuację, gdy rozmnażają się tylko osobniki najstarsze. Rozpatrzmy przykład populacji z wydzielonymi dwoma grupami wiekowymi — rozróżniamy osobniki dojrzałe, zdolne do reprodukcji i osobniki niedojrzałe. Mamy więc

$$\mathbf{M} = \begin{pmatrix} 0 & m \\ s & 0 \end{pmatrix}. \quad (4.8)$$

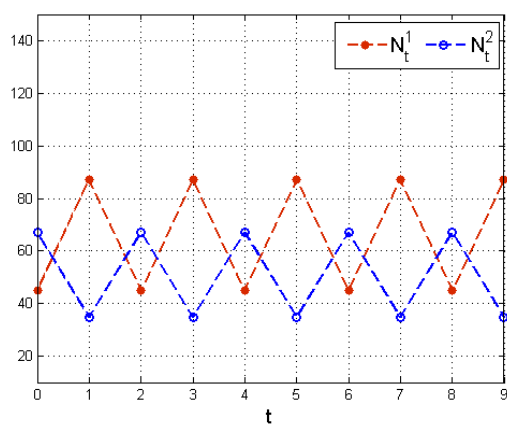
Obliczając kolejne potęgi macierzy (4.6) dostajemy

$$\mathbf{M}^{2t-1} = \begin{pmatrix} 0 & m^t s^{t-1} \\ s^t m^{t-1} & 0 \end{pmatrix} \quad \text{oraz} \quad \mathbf{M}^{2t} = \begin{pmatrix} m^t s^t & 0 \\ 0 & s^t m^t \end{pmatrix},$$

co łatwo wykazać indukcyjnie. Zauważmy, że jeśli $ms = 1$ i $s < 1$, to $\mathbf{N}_{2t} = \mathbf{N}_0$ oraz $\mathbf{N}_{2t-1} = \begin{pmatrix} m N_0^2 \\ s N_0^1 \end{pmatrix}$, czyli mamy zachowanie cykliczne, por. rys. 4.6.



Rysunek 4.5. Przykładowe rozwiązania modelu (4.6) w przypadku trzech grup wiekowych w zależności od dominującej wartości własnej λ_0 . Wstawki pokazują ewolucję procentowego udziału poszczególnych grup wiekowych w całości populacji.



Rysunek 4.6. Cykliczne zachowanie rozwiązań modelu (4.8) gdy $ms = 1$ i $s < 1$.

5. Modele pojedynczej populacji z uwzględnieniem wieku II

5.1. Modele z opóźnieniem

Innym sposobem wprowadzenia pewnej zależności od wieku jest modelowanie przy użyciu równań z opóźnieniem. Wróćmy do opisu dynamiki populacji za pomocą wzoru (2.1) i założmy, że przyrost *per capita* zależy nie od liczebności populacji w bieżącej chwili t , ale od stanu w pewnej chwili w przeszłości $t - \tau$. Kiedy tak się będzie działo? Wyobraźmy sobie populację roślinożerców, które zjadają rośliny będące w pewnym konkretnym wieku τ i jest to jednocześnie wiek, w którym te rośliny rozsiewają nasiona. Jeśli roślina zostanie zjedzona, to nie rozsiewe nasion, a wtedy w przyszłości osobniki opisywanej populacji nie mają co jeść. Ilość zjedzonych roślin zależy od stanu populacji w bieżącej chwili, zatem ilość jedzenia w przyszłości, czyli przyrost *per capita*, zależy od tego stanu. Przy takich założeniach równanie na przyrost *per capita* przyjmuje postać

$$\frac{\dot{N}(t)}{N(t)} = f(N(t - \tau)).$$

W szczególności równanie logistyczne z opóźnieniem zaproponowane przez G. E. Hutchinsona w 1948 r. zapiszemy jako

$$\dot{N}(t) = rN(t) \left(1 - \frac{N(t - \tau)}{K} \right). \quad (5.1)$$

Oczywiste jest, że przy opisanym powyżej modelu heurystycznym powinniśmy rozważać nie dokładnie jedno ustalone opóźnienie τ , ale pewien rozkład opóźnienia, gdyż nie ma w naturze takich roślin, które rozsiewałyby nasiona dokładnie w danym wieku, ale jak zwykle staraliśmy się zbudować jak najprostsz model, zatem przyjęliśmy uproszczenie polegające na ustaleniu opóźnienia τ .

Chcemy zbadać zależność rozwiązań równania (5.1) od wielkości opóźnienia $\tau > 0$. Zajmiemy się najpierw omówieniem podstawowych własności, takich jak istnienie, jednoznaczność i nieujemność rozwiązań dla nieujemnego warunku początkowego. Zauważmy, że aby rozwiązać równanie z opóźnieniem τ nie wystarczy, że określimy początkową liczebność populacji $N(0) = N_0$, ale musimy zadać funkcję początkową określoną na przedziale długości opóźnienia, czyli $N_0 : [-\tau, 0] \rightarrow \mathbb{R}^+$. Typowo w teorii równań różniczkowych z opóźnionym argumentem zakładamy, że funkcja początkowa jest ciągła, ale nie zawsze jest to założenie konieczne. W szczególności — znając funkcję początkową N_0 możemy rozwiązać równanie (5.1) metodą kroków, o ile tylko funkcja początkowa jest całkowalna. Dokładniej, niech $t \in [0, \tau]$. Wtedy

$$\frac{\dot{N}(t)}{N(t)} = r \left(1 - \frac{N_0(t - \tau)}{K} \right) \implies \ln \frac{N(t)}{N_0(0)} = rt - \frac{r}{K} \int_{-\tau}^{t-\tau} N_0(s) ds,$$

czyli

$$N(t) = N_0(0) \exp \left(rt - \frac{r}{K} \int_{-\tau}^{t-\tau} N_0(s) ds \right) \quad \text{dla } t \in [0, \tau].$$

Oznaczmy otrzymane rozwiązanie przez $N_1(t) = N_0(0) \exp\left(rt - \frac{r}{K} \int_{-\tau}^{t-\tau} N_0(s) ds\right)$ dla $t \in [0, \tau]$. Zauważmy, że jest ono dobrze określone dla ciągłej funkcji początkowej, co więcej w tym przypadku wystarczy, żeby funkcja N_0 była całkowalna. Mamy też jednoznaczność, a nieujemność wyniku z nierówności $N_0(0) \geq 0$, przy czym $N_0(0) = 0$ implikuje $N(t) \equiv 0$ (stany stacjonarne równań z opóźnieniem są oczywiście takie same jak dla analogicznego równania bez opóźnienia, skoro nie zależą one od czasu). Teraz zastosujemy metodę indukcji matematycznej. Załóżmy, że znamy rozwiązanie $N_k(t)$ na przedziale $[(k-1)\tau, k\tau]$ i znajdziemy rozwiązanie na kolejnym przedziale

$$N_{k+1}(t) = N_k(k\tau) \exp\left(r(t - k\tau) - \frac{r}{K} \int_{(k-1)\tau}^{t-\tau} N_k(s) ds\right) \quad \text{dla } t \in [k\tau, (k+1)\tau].$$

Wobec tego metoda indukcji matematycznej gwarantuje, że rozwiązanie istnieje dla dowolnego $t \geq 0$ i ma pożądane własności.

Zbadamy teraz własności asymptotyczne rozwiązań, w szczególności stabilność lokalną rozwiązań stacjonarnych. Metoda badania stabilności jest analogiczna, jak stosowana w przypadku równań bez opóźnienia. Przeprowadzamy najpierw linearyzację wokół stanu stacjonarnego. Niech \bar{N} będzie rozwiązaniem stacjonarnym, czyli $\bar{N} = 0$ albo $\bar{N} = K$. Wprowadzamy nową zmienną $x(t)$, która oznacza odchylenie od stanu stacjonarnego, $N(t) = \bar{N} + x(t)$, przy czym zakładamy, że $|x(t)| < \varepsilon$ i pomijamy wyrazy rzędu ε^2 . Mamy

$$\dot{x}(t) = r(\bar{N} + x(t)) \left(1 - \frac{\bar{N} + x(t-\tau)}{K}\right) \approx rx(t) \left(1 - \frac{\bar{N}}{K}\right) - r\bar{N} \frac{x(t-\tau)}{K}$$

po pominięciu składnika $-rx(t) \frac{x(t-\tau)}{K}$ i zauważeniu, że $r\bar{N} \left(1 - \frac{\bar{N}}{K}\right) = 0$ dla obu stanów stacjonarnych.

Dla stanu stacjonarnego $\bar{N} = 0$ równanie zlinearyzowane ma postać

$$\dot{x}(t) = rx(t).$$

Widzimy więc, że odchylenie od stanu stacjonarnego $x(t)$ rośnie, zatem $\bar{N} = 0$ jest niestabilne.

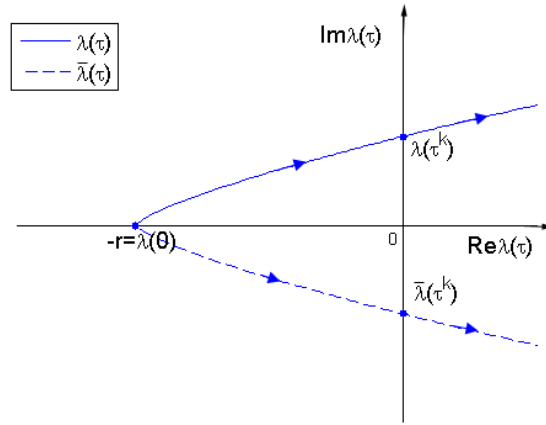
Z kolei dla dodatniego stanu stacjonarnego

$$\dot{x}(t) = -rx(t-\tau).$$

Jak zbadać stabilność stanu stacjonarnego $\bar{x} = 0$ powyższego równania? Tak jak w przypadku równań bez opóźnienia szukamy rozwiązań w postaci wykładniczej $x(t) = x_0 e^{\lambda t}$. Jeśli wszystkie wartości własne λ mają części rzeczywiste ujemne, to $x(t) \rightarrow 0$ przy $t \rightarrow +\infty$ dla dostatecznie małych x_0 . Stąd odchylenie maleje do 0, zatem stan $\bar{N} = K$ równania (5.1) jest lokalnie asymptotycznie stabilny. Jeśli natomiast istnieje wartość własna o części rzeczywistej dodatniej, to $\bar{N} = K$ jest niestabilny. Okazuje się, że dla równań z opóźnieniem równanie charakterystyczne, w tym przypadku

$$\lambda = -re^{-\lambda\tau},$$

ma nieskończenie wiele rozwiązań, które zależą w sposób ciągły od parametrów, w szczególności od opóźnienia. Skoro dla $\tau = 0$ mamy $\lambda = -r < 0$, to dla małych opóźnień stan $\bar{N} = K$ pozostaje stabilny. Zastanówmy się kiedy może nastąpić destabilizacja. Skoro niestabilność wiąże się z pojawieniem się wartości własnej o dodatniej części rzeczywistej, to dla pewnej krytycznej wartości $\tau^k > 0$ musimy mieć $\lambda = \pm i\omega$, $\omega > 0$, i wartości własne przechodzą z lewej półpłaszczyzny zespolonej na prawą, więc $\left. \frac{d\Re(\lambda)}{d\tau} \right|_{\tau=\tau^k} > 0$, gdzie $\Re(\lambda)$ oznacza część rzeczywistą wartości własnej, por. rys. 5.1.



Rysunek 5.1. Poglądowa ilustracja destabilizacji stanu stacjonarnego poprzez przejście wartości własnych z lewej półpłaszczyzny zespolonej na prawą.

Jeśli $\lambda = \pm i\omega$, to

$$\pm i\omega = -re^{\mp i\omega\tau^k} \Rightarrow |i\omega| = |re^{\mp i\omega\tau^k}|,$$

ale $|e^{\mp i\omega\tau^k}| = 1$, więc $\omega_k = r$ (w ogólnym przypadku łatwiej rozpatrywać tę równość po podniesieniu do kwadratu). Znając krytyczną wartość własną obliczamy τ^k

$$\begin{cases} 0 = -r \cos(\omega_k \tau^k), \\ \omega_k = r \sin(\omega_k \tau^k), \end{cases}$$

czyli $\cos(\omega_k \tau^k) = 0$ i $\sin(\omega_k \tau^k) = 1$, wobec tego $\omega_k \tau^k = \frac{\pi}{2} + 2n\pi$, $n \in \mathbb{N}$. Mamy więc ciąg krytycznych wartości własnych $(\tau_n^k)_{n=0}^{\infty}$. Okazuje się, że znak $\left. \frac{d\Re(\lambda)}{d\tau} \right|_{\tau=\tau^k}$ możemy sprawdzić korzystając z już przeprowadzonych obliczeń. W ogólnym przypadku dla układu równań z pojedynczym opóźnieniem τ równanie charakterystyczne ma postać

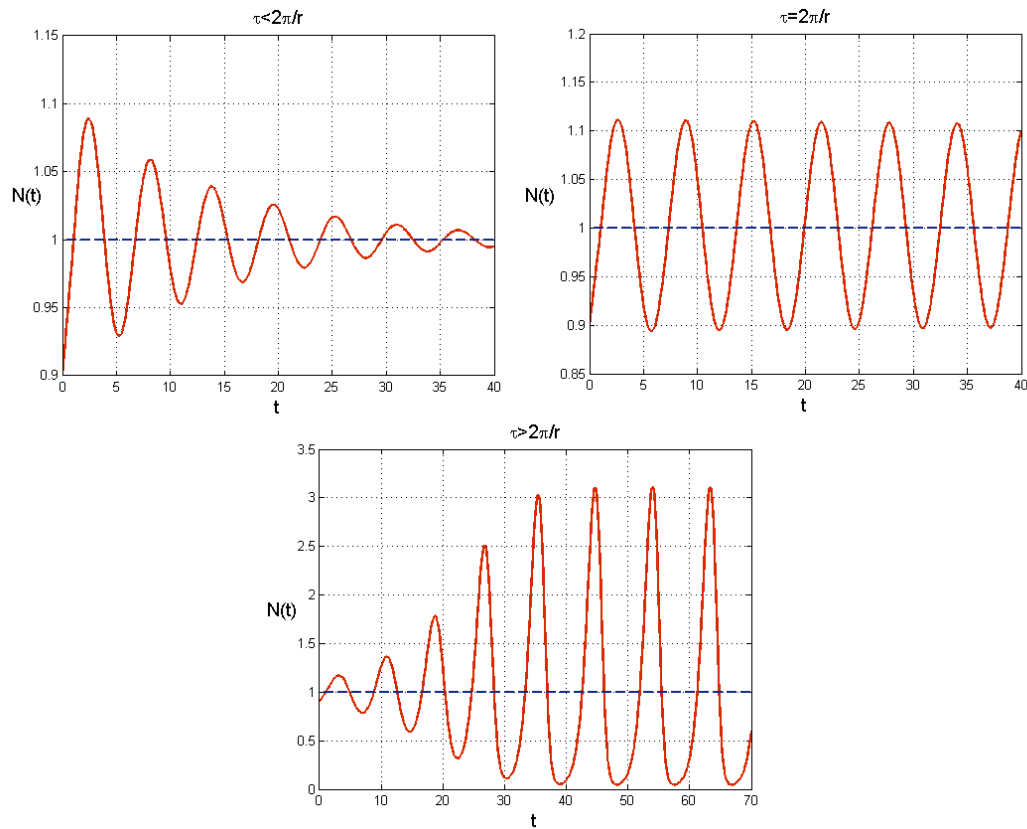
$$P(\lambda) + Q(\lambda)e^{-i\lambda\tau} = 0$$

i dla czysto urojonych wartości własnych $\lambda = i\omega$, $\omega > 0$, definiujemy funkcję pomocniczą

$$F(\omega) = \|P(i\omega)\|^2 - \|Q(i\omega)\|^2,$$

której miejsca zerowe wyznaczają czysto urojone wartości własne. U nas $F(\omega) = \omega^2 - r^2$. Podstawiamy $y = \omega^2$ i rozpatrujemy $\tilde{F}(y) = y - r^2$. Pochodna tej funkcji w punkcie $y = \omega_k^2$ ma taki sam znak jak $\left. \frac{d\Re(\lambda)}{d\tau} \right|_{\tau=\tau^k}$. W naszym przypadku $\tilde{F}'(y) = 1 > 0$, zatem zawsze wartości własne przechodzą z lewej półpłaszczyzny na prawą. Wobec tego dla pierwszej wartości krytycznej $\tau_0^k = \frac{\pi}{2r}$ następuje destabilizacja i rozwiązanie $\bar{N} = K$ pozostaje niestabilne dla wszystkich $\tau > \tau_0^k$. Ten mechanizm destabilizacji nazywamy bifurkacją Hopfa. W jej wyniku pojawiają się nietrywialne rozwiązania okresowe o okresie $\frac{2\pi}{\omega_k} = \frac{2\pi}{r}$, co widzimy na wykresach na rys. 5.2.

Podsumowując tę tematykę należy stwierdzić, że wprowadzenie do opisu heurystycznego zależności od wieku prowadzi najczęściej do dynamiki oscylacyjnej, która jest zwykle obserwowana w przypadku populacji występujących w naturze. Widzimy też, że opis dynamiki za pomocą równań z opóźnieniem może przypominać zachowanie rozwiązań modeli dyskretnych, gdzie też obserwujemy oscylacje. Co więcej, jeśli prawa strona równania z opóźnieniem reprezentuje np. funkcję Hilla, to dla odpowiednio dużych wartości współczynnika Hilla występują



Rysunek 5.2. Wykresy przedstawiające rozwiązania równania (5.1) dla różnych wartości opóźnienia τ . Odpowiednio od lewej: $\tau < \tau_0^k$; $\tau = \tau_0^k$; $\tau > \tau_0^k$.

zachowania chaotyczne, znów analogicznie jak w modelach dyskretnych. Możemy przypuszczać, że podobieństwa te wiążą się z podobną strukturą obu typów modeli — w modelach dyskretnych tak jak w równaniach z opóźnieniem dynamika w chwili bieżącej t zależy od stanu układu z chwili poprzedniej $t - 1$.

6. Modele oddziaływań między dwiema populacjami I

Przejdziemy teraz do omówienia zagadnień związanych z modelowaniem ekosystemu, w którym występują dwie populacje. W zależności od oddziaływań między tymi populacjami, rozróżniamy trzy typy ekosystemów

1. **układ drapieżnik – ofiara**, w którym jeden z gatunków ekosystemu jest pożywieniem dla drugiego;
2. **konkurencja**, w którym gatunki rywalizują o zasoby środowiska (pożywienie, miejsca lęgowe itp.);
3. **symbioza** — zjawisko współżycia przynajmniej dwóch gatunków, które wpływają pozytywnie na siebie (**mutualizm**) lub wyraźne korzyści odnosi jedynie jedna ze stron, nie szkodząc drugiej (**komensalizm**).

W tym rozdziale zajmiemy się opisem średnich zagęszczeń obu populacji, nie uwzględnimy przestrzennego rozmieszczenia osobników w ekosystemie, co oczywiście kryje dodatkowe założenie, że osobniki są rozmieszczone jednorodnie i jest ich dostatecznie dużo, aby można było mówić o średnim zagęszczeniu.

6.1. Model Lotki – Volterry

Wprowadzimy najpierw pierwszy historycznie model opisujący oddziaływania dwóch populacji w ekosystemie. Dotyczy on układu drapieżnik – ofiara i został zaproponowany równolegle jako model populacyjny przez **Vito Volterre** [17] oraz jako model łańcucha reakcji biochemicznych przez **Alfreda Lotkę** [10]. Volterra zaproponował ten model w celu wyjaśnienia pewnego, zdawałoby się paradoksalnego zjawiska, które zostało zaobserwowane po I wojnie światowej. Po ustaniu działań wojennych, kiedy ludzie na powrót zajęli się uprawianiem swoich zawodów, rybacy odkryli, że populacja ryb drapieżnych w Adriatyku zwiększyła się. Uznano to za paradoks, gdyż zdawałoby się, że wszystkie gatunki powinny ucierpieć w wyniku działań wojennych. Volterra na bazie swojego modelu wykazał, że wzrost liczebności drapieżników był całkiem naturalny, ponieważ w czasie wojny zaprzestano połowów i dzięki temu populacja drapieżników mogła wrócić do stanu naturalnego. Co więcej, model ten odzwierciedla znane w ekologii prawo zachowania średnich, które mówi, że w naturalnych siedliskach zmiany liczebności populacji w czasie zachodzą tak, że zachowana zostaje liczebność średnia.

Zacniemy od przedstawienia modelu heurystycznego, na bazie którego wyprowadzamy równania **modelu Lotki – Volterry**. Zakładamy, że w ekosystemie występują dwa gatunki \mathcal{P}_d i \mathcal{P}_o , przy czym osobniki drugiego gatunku stanowią pożywienie osobników pierwszego gatunku, czyli drapieżników. Jeśli nie ma drapieżników, to gatunek \mathcal{P}_o ma bardzo dobre warunki i może się rozwijać, podlegając prawu Malthusa ze współczynnikiem rozrodczości $r > 0$. Natomiast drapieżniki, jeśli nie ma ofiar, to nie mają co jeść, więc giną z głodu. Jeśli w środowisku są osobniki obu gatunków, to drapieżniki polują na ofiary. Zakładamy, że polowanie jest możliwe tylko w przypadku bezpośredniego kontaktu, osobniki poruszają się losowo, zatem liczba kontaktów jest proporcjonalna do liczebności obu gatunków. Zauważmy, że w takiej sytuacji dla pojedynczego

drapieżnika mamy odpowiedź funkcjonalną Hollinga typu I. W związku z polowaniem ubywa osobników gatunku \mathcal{P}_o , proporcjonalnie do liczby spotkań, a współczynnik proporcjonalności odzwierciedla skuteczność drapieżnika. Po upolowaniu ofiary drapieżnik zyskuje energię, której część przeznaczana na rozmnażanie.

Niech $P(t)$ oznacza zagęszczenie drapieżników, a $V(t)$ — zagęszczenie ofiar. Na podstawie powyższego modelu heurystycznego zapiszemy układ równań różniczkowych dynamiki obu populacji

$$\begin{aligned}\dot{V} &= rV - aVP, \\ \dot{P} &= -sP + abVP,\end{aligned}\tag{6.1}$$

gdzie dla uproszczenia zapisu pomijamy zmienną niezależną t . Poszczególne składniki układu (6.1) mają następującą interpretację

- rV i $-sP$ opisują wewnętrzną dynamikę odpowiednio gatunku \mathcal{P}_o i \mathcal{P}_d , r jest współczynnikiem rozrodczości ofiar, a s — współczynnikiem śmiertelności drapieżników;
- aVP odzwierciedla liczbę losowych spotkań osobników obu gatunków, a jest współczynnikiem skuteczności polowania, składnik ten interpretujemy także jako biomasę upolowanych ofiar;
- współczynnik b odzwierciedla część upolowanej biomasy, którą gatunek \mathcal{P}_d przeznaczają na reprodukcję.

Przejdziemy teraz do analizy zachowania rozwiązań układu (6.1). Ponieważ układ (6.1) jest autonomiczny, więc bez straty ogólności przyjmujemy, że $t_0 = 0$ i $V(0) = V_0$, $P(0) = P_0$. Sformułujemy najpierw następujące stwierdzenie.

Stwierdzenie 6.1. *Rozwiązania układu (6.1) dla nieujemnego warunku początkowego (V_0, P_0) , $V_0, P_0 \geq 0$, istnieją dla wszystkich $t > 0$, są jednoznaczne i nieujemne.*

Dowód. Ponieważ prawa strona układu (6.1) jest klasy \mathbf{C}^1 (nawet analityczna) w \mathbb{R}^2 , to lokalne rozwiązania istnieją dla dowolnego warunku początkowego i są one jednoznaczne. Niech $[0, t^*)$ będzie przedziałem określoności rozwiązania dla warunku początkowego (V_0, P_0) . Pokażemy, że rozwiązanie to jest nieujemne. Niech $V_0 = 0$. Wtedy $\dot{V}(0) = 0$ i zmienna V nie zmienia się. Mamy więc $V(t) \equiv 0$ i wtedy $\dot{P} = -sP$, czyli $P(t) = P_0 e^{-st}$. Analogicznie dla $P_0 = 0$ mamy $P(t) \equiv 0$ i $V(t) = V_0 e^{rt}$. Niech teraz $V_0, P_0 > 0$. Załóżmy, że istnieje taka chwila $\bar{t} > 0$, że zmienna V staje się ujemna. Wtedy $V(\bar{t}) = 0$, co jest sprzeczne z jednoznacznością rozwiązań, czyli $V(t) > 0$ dla wszystkich $t > 0$, dla których rozwiązanie istnieje. Podobnie $P(t) > 0$ dla $t > 0$.

Pozostaje jeszcze udowodnić istnienie rozwiązań dla wszystkich $t > 0$. Załóżmy, że rozwiązanie istnieje na pewnym skończonym przedziale $[0, t^*)$. Wiemy już, że rozwiązanie jest nieujemne, więc

$$\dot{V} \leq rV.$$

Po pomnożeniu powyższej nierówności przez e^{-rt} dostajemy

$$\frac{dV}{dt} e^{-rt} - rV e^{-rt} \leq 0 \implies \frac{d}{dt} (V e^{-rt}) \leq 0,$$

czyli $V(t)e^{-rt} \leq V_0$ dla $t \geq 0$, więc $V(t) \leq V_0 e^{rt}$. Stąd dla dowolnego skończonego t^* zachodzi $V(t^*) \leq V_{\max} = V_0 e^{rt^*}$ oraz $\dot{V}(t^*) \leq rV_{\max}$. Teraz możemy oszacować pochodną zmiennej P . Dla $t \leq t^*$ otrzymujemy $\dot{P} \leq abV_{\max}P$, czyli $P(t) \leq P_0 e^{abV_{\max}t} \leq P_{\max} = P_0 e^{abV_{\max}t^*}$ oraz $\dot{P} \leq abV_{\max}P_{\max}$. Szacując pochodne z dołu dostajemy $\dot{V} \geq -aV_{\max}P_{\max}$ oraz $\dot{P} \geq -sP_{\max}$. Mamy więc funkcje ograniczone o ograniczonych pochodnych, z czego wynika, że istnieją granice $\lim_{t \rightarrow t^*} V(t) = V_g$ oraz $\lim_{t \rightarrow t^*} P(t) = P_g$. Gdyby któraś z tych granic nie istniała, to możemy wybrać dwa podciągi zbieżne do różnych granic. Niech np. $\lim_{n_1 \rightarrow \infty} V(t_{n_1}) = V_{g1}$ oraz

$\lim_{n_2 \rightarrow \infty} V(t_{n_2}) = V_{g_2}$ dla $t_{n_1}, t_{n_2} \rightarrow t^*$ i $V_{g_1} \neq V_{g_2}$. Wtedy z twierdzenia o wartości średniej $\dot{V}(t_{n_\xi}) = \frac{V(t_{n_1}) - V(t_{n_2})}{t_{n_1} - t_{n_2}} \rightarrow \infty$ przy $n_\xi \rightarrow \infty$, co jest sprzeczne z ograniczonością pochodnej. Biorąc warunek początkowy $t_0 = t^*$, $V(t_0) = V_g$, $P(t_0) = P_g$ i korzystając z twierdzenia o istnieniu i jednoznaczności rozwiązań przedłużamy rozwiązanie dla $t > t^*$. Wobec dowolności t^* rozwiązanie istnieje dla wszystkich $t \geq 0$. \square

Zauważmy, że gładkość prawej strony i liniowe oszacowanie pochodnej zawsze gwarantuje przedłużalność rozwiązań.

Przystąpimy teraz do analizy asymptotyki rozwiązań układu (6.1). Wyznamy najpierw stany stacjonarne (\bar{V}, \bar{P}) . Spełniają one układ równań

$$\begin{aligned} 0 &= \bar{V}(r - a\bar{P}), \\ 0 &= \bar{P}(-s + ab\bar{V}), \end{aligned}$$

zatem albo $(\bar{V}, \bar{P}) = (0, 0)$, albo $(\bar{V}, \bar{P}) = (\frac{s}{ab}, \frac{r}{a})$. Lokalną stabilność spróbujemy zbadać stosując metodę linearyzacji. Macierz Jacobiego układu (6.1) ma postać

$$\begin{pmatrix} r - aP & -aV \\ abP & -s + abV \end{pmatrix}.$$

Dla zerowego stanu stacjonarnego otrzymujemy

$$\begin{pmatrix} r & 0 \\ 0 & -s \end{pmatrix},$$

widzimy więc, że jest to punkt siodłowy, gdyż wartości własne $\lambda_1 = r$, $\lambda_2 = -s$ mają różne znaki. Z kolei dla dodatniego stanu stacjonarnego

$$\begin{pmatrix} 0 & -a\bar{V} \\ ab\bar{P} & 0 \end{pmatrix},$$

czyli wartości własne macierzy Jacobiego $\lambda_{1,2} = \pm i\sqrt{rs}$ są czysto urojone, zatem nie jest spełnione założenie twierdzenia o linearyzacji dotyczące hiperboliczności układu [14] i nie możemy w tym przypadku skorzystać z tej metody.

Spróbujemy naszkicować portret fazowy. Zaczniemy od wyznaczenia izoklin zerowych, czyli krzywych, na których $\dot{V} = 0$ (izoklina zerowa zmiennej V) albo $\dot{P} = 0$ (izoklina zerowa zmiennej P). Mamy

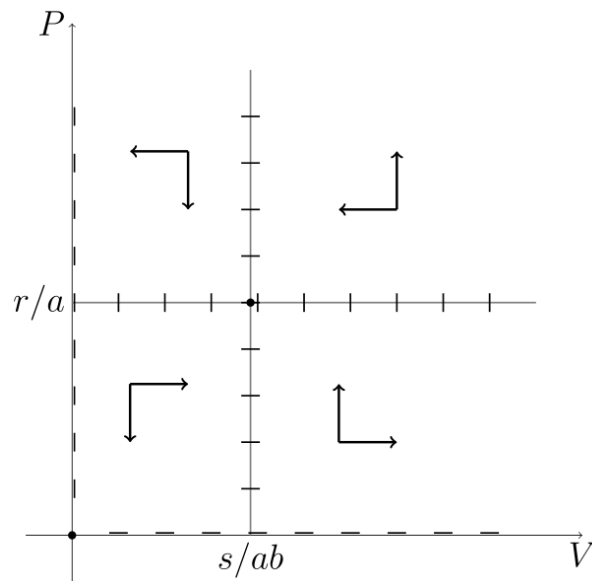
$$\dot{V} = 0 \Rightarrow V = 0 \quad \text{albo} \quad P = \frac{r}{a}$$

oraz

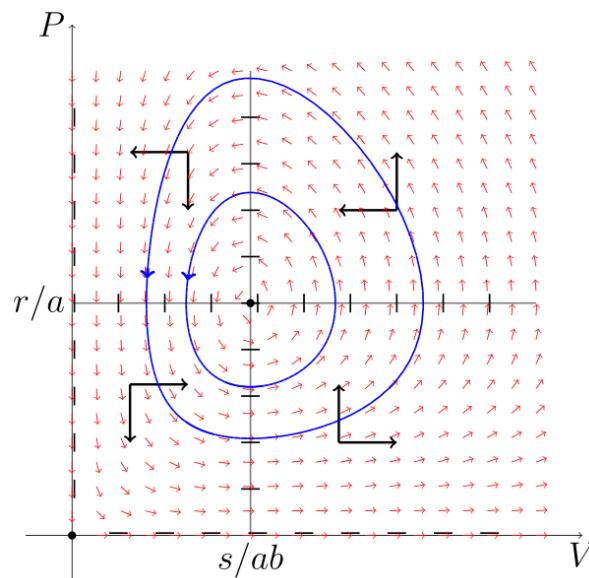
$$\dot{P} = 0 \Rightarrow P = 0 \quad \text{albo} \quad V = \frac{s}{ab}.$$

Izokliny zerowe dzielą przestrzeń fazową $(\mathbb{R}^+)^2$ na cztery podprzestrzenie, w których poszczególne zmienne mają ustalony kierunek przebiegu, por. rys. 6.1. Widzimy, że otrzymane w ten sposób pole wektorowe pokazuje, że rozwiązania okrążają dodatni stan stacjonarny (\bar{V}, \bar{P}) .

Można pokazać, że wszystkie rozwiązania są okresowe, co wynika z własności całki pierwszej [14] układu (6.1). Znalezienie całki pierwszej i zbadanie jej własności pozostawiamy jako ćwiczenie.



Rysunek 6.1. Szkic portretu fazowego układu (6.1) z zaznaczonymi izoklinami zerowymi i kierunkami przebiegu poszczególnych zmiennych.



Rysunek 6.2. Portret fazowy układu (6.1) wraz z przykładowymi krzywymi fazowymi.

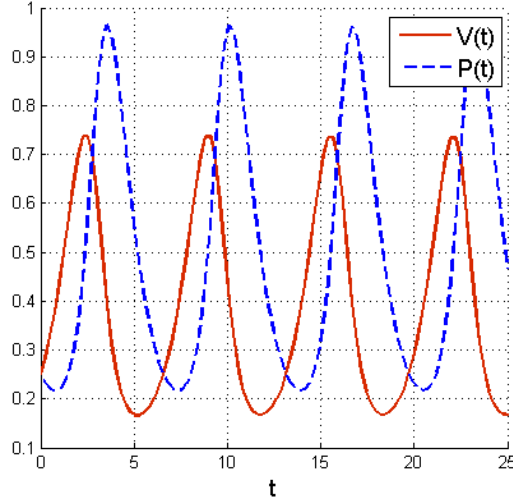
Na koniec zauważmy, że wszystkie rozwiązania oscylują wokół stanu stacjonarnego (\bar{V}, \bar{P}) , por. rys. 6.2. Pokażemy, że każde rozwiązanie ma wartość średnią równą współrzędnym tego stanu. Niech T oznacza okres rozwiązania $(V(t), P(t))$.

$$\frac{dV}{V} = r - aP \implies \ln \frac{V(T)}{V_0} = rT - a \int_0^T P(s) ds,$$

a ponieważ rozwiązanie jest okresowe, to $V(T) = V_0$, czyli

$$\int_0^T P(s) ds = \frac{rT}{a} \implies P_{\text{sr}} = \frac{1}{T} \int_0^T P(s) ds = \frac{r}{a}.$$

Analogicznie pokazujemy, że $V_{\text{sr}} = \frac{s}{ab}$. Rozwiązania oscylują wokół stanu stacjonarnego, przy czym oscylacje populacji drapieżników i ofiar są przesunięte w fazie, por. rys. 6.3.



Rysunek 6.3. Przebieg przykładowych rozwiązań układu (6.1).

Na tej podstawie wyjaśnimy paradoks związany z populacją ryb drapieżnych w Adriatyku. Zauważmy, że jeśli odławiamy ryby, to zakładając jednakową intensywność odłowu dla obu populacji \mathcal{P}_o i \mathcal{P}_d , otrzymamy następującą modyfikację układu (6.1)

$$\begin{aligned}\dot{V} &= (r - \gamma)V - aVP, \\ \dot{P} &= -(s + \gamma)P + abVP,\end{aligned}\tag{6.2}$$

gdzie γ oznacza współczynnik odłowu. Zauważmy dalej, że aby populacja ofiar nie wyginęła, musimy dokonywać odłowów w sposób sensowny, tak by $\gamma < r$. Jeśli łowimy zbyt dużo, $\gamma > r$, to $\dot{V} < 0$ dla wszystkich t i populacja ginie. Natomiast gdy $\gamma < r$, to układ (6.2) ma dokładnie taką samą postać jak (6.1), więc odławiane populacje mają średnie zagęszczenia równe

$$V_{\text{sr}}^{\text{odławiane}} = \frac{s + \gamma}{ab} > V_{\text{sr}}, \quad P_{\text{sr}}^{\text{odławiane}} = \frac{r - \gamma}{a} < P_{\text{sr}},$$

zatem odławianie powoduje zmniejszenie średniej liczebności populacji drapieżników i zwiększenie średniej liczebności populacji ofiar. W związku z tym w czasie gdy nie było połowów, populacje wróciły do stanu naturalnego, więc liczebność drapieżników wzrosła.

6.2. Konstruktywna krytyka modelu Lotki – Volterry

Chociaż model Lotki – Volterry odzwierciedla obserwowane w naturze oscylacje w układach drapieżnik – ofiara, a także w sposób analityczny dowodzi ekologicznego prawa zachowania średnich w tych układach, to jednak nie jest pozbawiony wad. Przede wszystkim z matematycznego punktu widzenia jego główną wadę stanowi brak **stabilności strukturalnej**. Pojęcie stabilności strukturalnej wykracza poza ramy tego wykładu — w skrócie stabilność strukturalna oznacza, że niewielka zmiana prawej strony układu nie prowadzi do radykalnych zmian w dynamice rozwiązań. Widzimy, że w przypadku układu (6.1) nawet bardzo mała zmiana prawej strony może

skutkować zmianą typu stanu stacjonarnego — stan typu centrum bardzo łatwo zaburzyć. Dodatkowo model ma też pewną własność, która podlega krytyce z ekologicznego punktu widzenia. Rozwiązania układu (6.1) oscylują w taki sposób, że wzrost populacji drapieżników poprzedza wzrost populacji ofiar, natomiast w większości rzeczywistych układów drapieżnik – ofiara takie oscylacje są przesunięte w stosunku do tych rozwiązań; najpierw obserwujemy wzrost populacji ofiar, a potem następujący po nim wzrost populacji drapieżników. Należy też zauważyć, że w modelu heurystycznym pominięto wiele istotnych dla układów drapieżnik – ofiara czynników. W związku z tym model Lotki – Volterry był na wiele różnych sposobów modyfikowany i poniżej omówimy pokrótce kilka z tych modyfikacji.

7. Modele oddziaływań między dwiema populacjami II

7.1. Model drapieżnik – ofiara z ograniczoną pojemnością środowiska dla ofiar

Jak już wspomnieliśmy, w modelu Lotki – Volterry zostały pominięte różne istotne czynniki, gdyż w zamyśle Volterry taki model powinien być możliwie najprostszy. W szczególności wewnętrzna dynamika populacji ofiar została zbudowana w oparciu o model Malthusa zakładający nieograniczoną pojemność środowiska dla tego gatunku, co przy niewielkiej liczebności drapieżników prowadzi do gwałtownego wzrostu populacji ofiar i przyczynia się do okresowości rozwiązań. Załóżmy więc, że wewnętrzna dynamika gatunku \mathcal{P}_o rządzi się innymi prawami, w szczególności liczebność tej populacji jest ograniczona przez pojemność środowiska. Najprostszym znanym nam modelem opisującym taką dynamikę jest równanie logistyczne, zbudujmy zatem model drapieżnik – ofiara w oparciu o to równanie. Wobec tego, przy oznaczeniach wprowadzonych w wyjściowym modelu Lotki – Volterry, model z ograniczoną pojemnością środowiska dla ofiar możemy opisać następującym układem równań

$$\begin{aligned}\dot{V} &= rV\left(1 - \frac{V}{K}\right) - aVP, \\ \dot{P} &= -sP + abVP,\end{aligned}\tag{7.1}$$

gdzie K oznacza pojemność środowiska, analogicznie jak w modelu logistycznym. Dynamika rozwiązań układu (7.1) zależy w istotny sposób od tego parametru. Zauważmy, że podstawowe własności takie jak istnienie, jednoznaczność, nieujemność i przedłużalność rozwiązań możemy zbadać w analogiczny sposób jak dla układu (6.1). Wnioskujemy zatem, że dla dowolnego nieujemnego warunku początkowego (V_0, P_0) istnieje jednoznaczne, nieujemne rozwiązanie układu (7.1) określone dla wszystkich $t \geq 0$. Dość prosto możemy wykazać też ograniczoną rozwiązań. Sformułujmy odpowiednie stwierdzenie

Stwierdzenie 7.1. *Rozwiązania układu (7.1) dla nieujemnego warunku początkowego (V_0, P_0) , $V_0, P_0 \geq 0$, pozostają ograniczone. Co więcej, $V(t) \leq \max\{V_0, K\}$ dla $t \geq 0$.*

Dowód. Z nieujemności rozwiązań układu (7.1) wynika następujące szacowanie $\dot{V} \leq rV\left(1 - \frac{V}{K}\right)$. Rzeczywiście, jeśli $V_0 = 0$, to $V(t) \equiv 0$, jak dla równania logistycznego, a jeśli $P_0 = 0$, to $P(t) \equiv 0$ i V spełnia równanie logistyczne, stąd $V(t) \leq \max(V_0, K)$. Dla $V_0, P_0 > 0$ mamy $V(t), P(t) > 0$ dla dowolnego $t > 0$. Rozważmy równanie logistyczne $\dot{x} = rx\left(1 - \frac{x}{K}\right)$ z warunkiem początkowym $x_0 = V_0 > 0$ oraz różnicę $y(t) = x(t) - V(t)$. Zauważmy, że $\dot{y}(0) = \dot{x}(0) - \dot{V}(0) = aV_0P_0 > 0$. Zatem na pewnym odcinku $(0, t^*)$ zachodzi $x(t) > V(t)$. Jeśli istnieje $t > 0$, takie że $V(t) > x(t)$, to istnieje pierwsza taka chwila \bar{t} , że $V(\bar{t}) = x(\bar{t})$ oraz $\dot{y}(\bar{t}) \leq 0$, ale $\dot{y}(\bar{t}) = aV(\bar{t})x(\bar{t}) > 0$, co jest sprzeczne z definicją punktu \bar{t} . Wynika stąd, że zmienna V nie przekracza rozwiązań równania logistycznego z warunkiem początkowym V_0 . Zatem $V(t) \leq V_{\max} = \max\{V_0, K\}$. Załóżmy teraz, że zmienna P rośnie nieograniczenie. Mamy wtedy asymptotycznie $\dot{V}(t) \sim C_1 - C_2P(t)$, gdzie stałe $C_1, C_2 > 0$ zależą od parametrów i warunku początkowego. Stąd $\dot{V}(t) \rightarrow -\infty$ przy $t \rightarrow +\infty$, co przeczy ograniczoną V . Zatem zmienna P także pozostaje ograniczona. \square

Zajmiemy się teraz analizą portretu fazowego. Izokliny zerowe są następującymi prostymi — dla zmiennej V : $V = 0$ lub $P = \frac{r}{a} \left(1 - \frac{V}{K}\right)$;
— dla zmiennej P : $P = 0$ lub $V = \frac{s}{ab}$.

W zależności od parametru K — nietrywialne izokliny mogą się przeciąć w analizowanej przestrzeni fazowej, gdy K jest relatywnie duże, albo nie, gdy K jest małe. Zatem liczba stanów stacjonarnych zależy od K : zawsze są 2 stany $(0, 0)$ i $(K, 0)$, natomiast dla $K > \frac{s}{ab}$ istnieje dodatni stan stacjonarny (\bar{V}, \bar{P}) , gdzie $\bar{V} = \frac{s}{ab}$ i $\bar{P} = \frac{r}{a} \left(1 - \frac{s}{abK}\right)$. Wyznamy wartości własne dla poszczególnych stanów stacjonarnych. Macierz Jacobiego układu (7.1) ma postać

$$MJ(V, P) = \begin{pmatrix} r \left(1 - 2\frac{V}{K}\right) - aP & -aV \\ abP & -s + abV \end{pmatrix}.$$

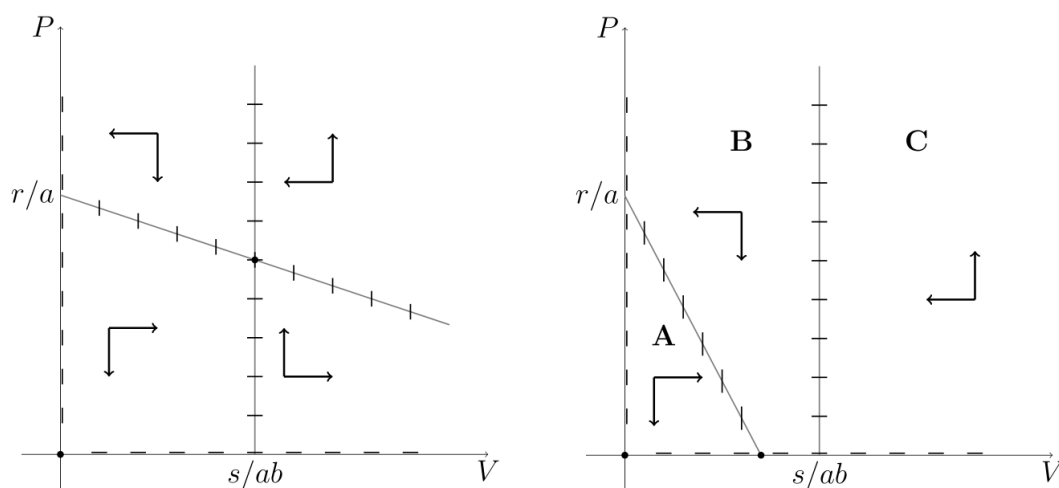
Stąd wyznaczamy wartości własne

- dla $(0, 0)$ mamy $\lambda_1 = r > 0$, $\lambda_2 = -s < 0$;
- dla $(K, 0)$: $\lambda_1 = -r < 0$, $\lambda_2 = abK - s$;
- dla (\bar{V}, \bar{P}) wartości własne są rozwiązaniem równania charakterystycznego

$$\lambda^2 + \frac{r\bar{V}}{K}\lambda + a^2b\bar{V}\bar{P} = 0,$$

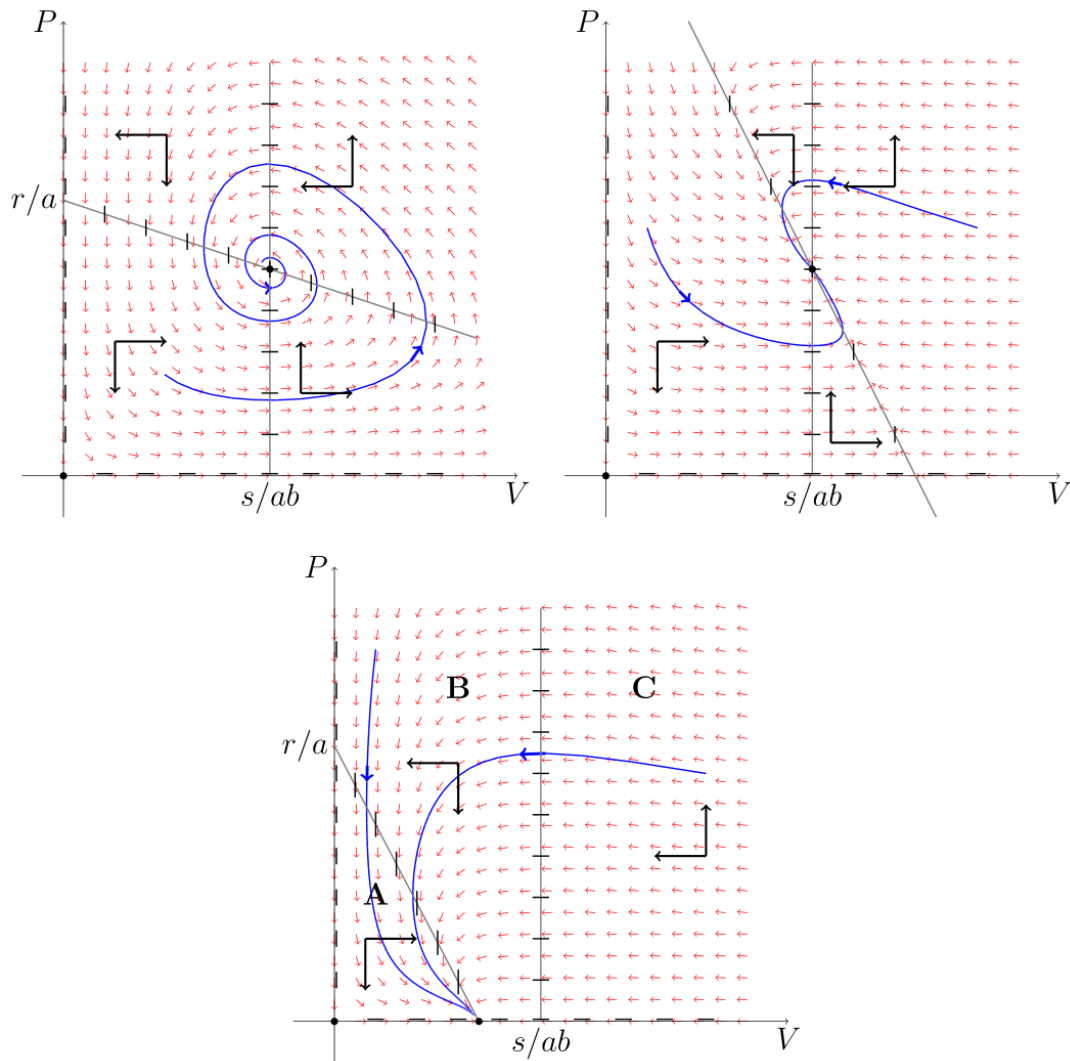
więc w zależności od parametrów albo λ_1 i λ_2 są rzeczywiste ujemne, albo są zespolone i mają ujemne części rzeczywiste.

Wobec tego stan stacjonarny $(0, 0)$ jest siodłem bez względu na parametry układu, z kolei $(K, 0)$ jest albo węzłem stabilnym, o ile $K < \frac{s}{ab}$, czyli dodatni stan stacjonarny nie istnieje, albo siodłem, gdy $K > \frac{s}{ab}$, a wtedy istnieje (\bar{V}, \bar{P}) , który jest albo węzłem, albo ogniskiem stabilnym. Typ punktu krytycznego (\bar{V}, \bar{P}) zależy także od K , jeśli K jest duże, tak że $4a^2b^2K^2 - 4asbK - rs > 0$, to obserwujemy ognisko i gasnące oscylacje, w przeciwnym przypadku — węzeł i zbieżność do stanu stacjonarnego jest od pewnego momentu monotoniczna.



Rysunek 7.1. Szkice portretu fazowego układu (7.1) z zaznaczonymi izoklinami zerowymi i kierunkami przebiegu poszczególnych zmiennych w przypadku gdy istnieje dodatni punkt stacjonarny (po lewej) oraz w przypadku gdy dodatni punkt stacjonarny nie istnieje (po prawej).

Przystępując do narysowania portretu fazowego zauważmy, że trywialne izokliny zerowe stanowią trajektorie układu. Rzeczywiście: jeśli $(0, P_0)$, $P_0 > 0$, jest warunkiem początkowym,



Rysunek 7.2. Portret fazowy układu (7.1) wraz z przykładowymi krzywymi fazowymi w przypadku gdy istnieje dodatni punkt stacjonarny (wykresy u góry) oraz w przypadku gdy dodatni punkt stacjonarny nie istnieje (wykres u dołu).

to oczywiście $\dot{V} = 0$, czyli zmienna V nie zmienia się, $V(t) = 0$ dla $t > 0$, natomiast dla zmiennej P mamy równanie $\dot{P} = -sP$, czyli $P(t) = P_0 \exp(-st)$ dla $t > 0$. Podobnie dla warunku początkowego $(V_0, 0)$, $V_0 > 0$, zachodzi $\dot{P} = 0$, czyli $P(t) = 0$ dla $t > 0$ oraz $\dot{V} = rV \left(1 - \frac{V}{K}\right)$, zatem $V(t)$ spełnia równanie logistyczne, co zauważyliśmy już wcześniej w dowodzie stwierdzenia 7.1.

Jeśli izokliny nie przecinają się wewnątrz przestrzeni fazowej, to przestrzeń tę możemy podzielić na trzy obszary, por. rys. 7.1 i rys. 7.2: **A** pod izokliną dla V , **B** nad tą izokliną i na lewo od izokliny dla P oraz **C** — na prawo od izokliny dla P . Zauważmy, że rozwiązania dla warunku początkowego z **C** muszą przejść do **B**. Gdyby tak się nie stało, to mielibyśmy obie zmienne monotoniczne, przy czym zmienna V pozostawałaby malejąca i ograniczona, a zmienna P rosłaby nieograniczenie. W takim przypadku z pierwszego równania asymptotycznie dostajemy $\dot{V} \sim c_1 - c_2P$, gdzie c_1 i c_2 są dodatnimi stałymi, czyli $\dot{V} \rightarrow -\infty$, co przeczy ograniczoności V . Zatem rozwiązanie przechodzi do **B** — tu obie zmienne są malejące, zatem albo pozostają w

B i wobec tego mają granicę, czyli zbiegają do stanu stacjonarnego $(K, 0)$ na brzegu obszaru, albo przechodzą do A . W A obie zmienne także są monotoniczne, tyle że V rośnie. Przebieg pola wektorowego pokazuje, że rozwiązania nie mogą wyjść z obszaru A , więc także są zbieżne i zbiegają do $(K, 0)$.

W drugim przypadku, gdy istnieje dodatni stan stacjonarny, przebieg pola wektorowego wskazuje, podobnie jak dla modelu Lotki – Volterry, że rozwiązania okrążają punkt krytyczny (\bar{V}, \bar{P}) w przestrzeni fazowej, por. rys. 7.1 i rys. 7.2. Wartości własne dla tego punktu implikują, że jest on lokalnie stabilny, ale na takiej podstawie nie możemy wykluczyć istnienia orbit zamkniętych, czyli cykli granicznych. Można to próbować wykazać na różne sposoby, w tym przypadku dzięki prostocie modelu działa kilka standardowych metod. Najprostszy sposób w przypadku układów dwuwymiarowych polega na zastosowaniu kryterium Dulaca – Bendixsona. Dla danego układu o prawej stronie opisanej funkcją $G(V, P) = (G_1(V, P), G_2(V, P))^T$ rozważamy pewną funkcję $B : (\mathbb{R}^+)^2 \rightarrow \mathbb{R}$ klasy \mathbf{C}^1 , przy czym

$$\operatorname{div}BG = \frac{\partial BG_1}{\partial V} + \frac{\partial BG_2}{\partial P}$$

nie zmienia znaku w $(\mathbb{R}^+)^2$. Wtedy w $(\mathbb{R}^+)^2$ nie ma cykli, w szczególności nie ma cykli granicznych. Typowo stosowana funkcja B ma postać $B(x, y) = \frac{1}{xy}$. W przypadku układu (7.1) dostajemy

$$\operatorname{div}BG(V, P) = \frac{\partial}{\partial V} \left(\frac{r}{P} \left(1 - \frac{V}{K} \right) - a \right) + \frac{\partial}{\partial P} \left(-\frac{s}{V} + ab \right) = -\frac{r}{PK} < 0,$$

z czego wnioskujemy globalną stabilność stanu stacjonarnego (\bar{V}, \bar{P}) , gdyż jeśli nie ma cyklu granicznego, to z twierdzenia Poincarégo – Bendixsona wszystkie rozwiązania zbiegają do stanu stacjonarnego, w tym przypadku do jedyne go stanu stabilnego (\bar{V}, \bar{P}) .

Bardziej uniwersalna metoda (działająca także w przypadku modeli o większej liczbie zmiennych) polega na znalezieniu funkcjonału Lapunowa. W tym przypadku można zastosować funkcjonał Lapunowa zaproponowany dla modelu Lotki – Volterry. W celu uproszczenia obliczeń przeprowadźmy najpierw zamianę zmiennych $x = V - \bar{V}$, $y = P - \bar{P}$, która przesuwa stan stacjonarny do punktu $(0, 0)$. Zauważmy, że po zamianie zmiennych powstaje nowa przestrzeń fazowa $\mathcal{DF} = (-\bar{V}, +\infty) \times (-\bar{P}, +\infty)$, a układ (7.1) przyjmuje postać

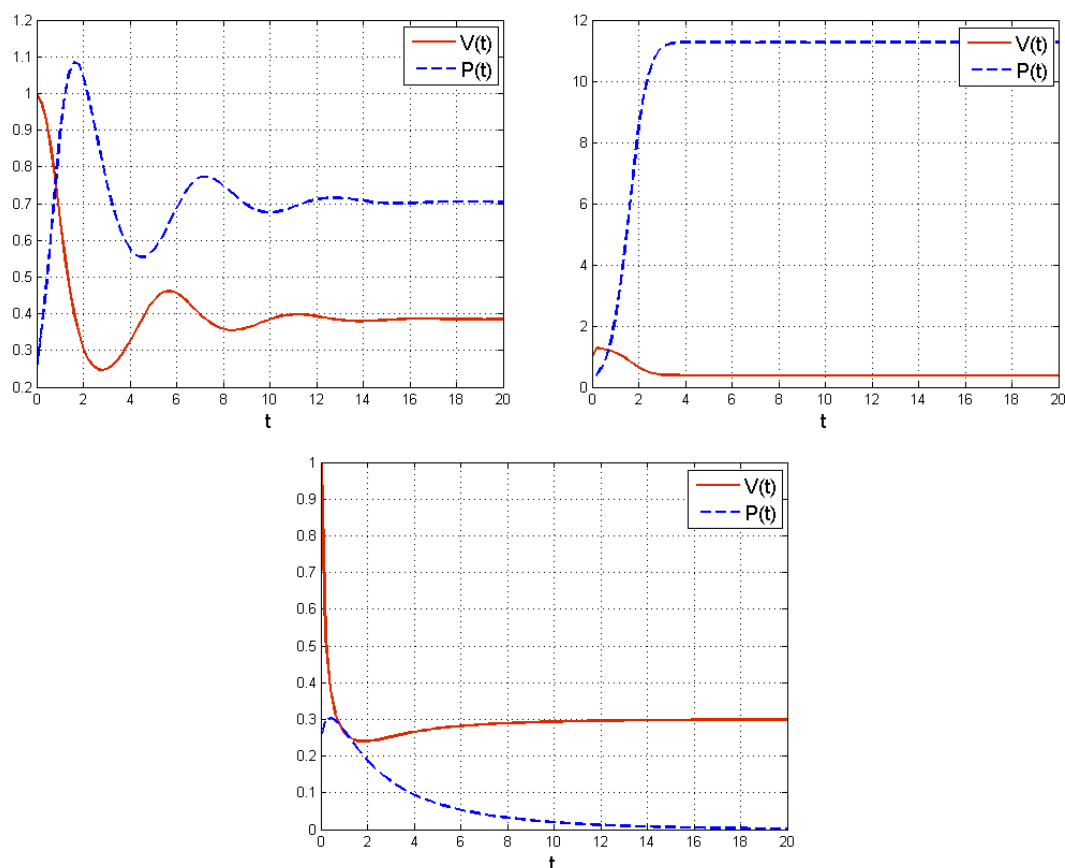
$$\begin{aligned} \dot{x} &= r(x + \bar{V}) \left(1 - \frac{x + \bar{V}}{K} \right) - a(x + \bar{V})(y + \bar{P}) = -(x + \bar{V}) \left(\frac{rx}{K} + ay \right), \\ \dot{y} &= -s(y + \bar{P}) + ab(x + \bar{V})(y + \bar{P}) = abx(y + \bar{P}). \end{aligned} \quad (7.2)$$

Funkcjonał Lapunowa dla układu (7.2) definiujemy jako

$$L(x, y) = \alpha_1 \left(x - \bar{V} \ln \frac{x + \bar{V}}{\bar{V}} \right) + \alpha_2 \left(y - \bar{P} \ln \frac{y + \bar{P}}{\bar{P}} \right).$$

Należy teraz sprawdzić odpowiednie własności funkcjonału L : w szczególności mamy $L(x, y) \geq 0$ w przestrzeni \mathcal{DF} oraz $L(x, y) = 0 \iff x = 0$ i $y = 0$ (czyli $V = \bar{V}$ i $P = \bar{P}$). Policzmy pochodną L wzdłuż trajektorii układu (7.2)

$$\begin{aligned} \dot{L}(x(t), y(t)) &= \alpha_1 \frac{x}{x + \bar{V}} \left(-(x + \bar{V}) \left(\frac{rx}{K} + ay \right) \right) + \alpha_2 \frac{y}{y + \bar{P}} (abx(y + \bar{P})) = \\ &= -\alpha_1 \frac{rx^2}{K} \leq 0 \end{aligned}$$



Rysunek 7.3. Przebieg przykładowych rozwiązań układu (7.1). Wykresy znajdujące się na górze przedstawiają rozwiązania układu (7.1) w przypadku gdy istnieje dodatni punkt stacjonarny i jest on odpowiednio ogniskiem stabilnym (po lewej) lub węzłem stabilnym (po prawej). Wykres na dole przedstawia rozwiązanie gdy nie istnieje dodatni punkt stacjonarny.

dla $\alpha_1 = b$ i $\alpha_2 = 1$. Wynika stąd globalna stabilność dodatniego stanu stacjonarnego. Aby wykazać jego globalną asymptotyczną stabilność wystarczy udowodnić, że $L(V(t), P(t))$ jest ściśle malejąca, co pozostawiamy jako ćwiczenie.

Zauważmy, że tak zdefiniowany układ ma stany stacjonarne asymptotycznie stabilne, przy czym dla $K < \frac{s}{ab}$ wszystkie rozwiązania (oprócz rozwiązań na brzegu przestrzeni fazowej) zbiegają do stanu stacjonarnego $(K, 0)$, natomiast dla $K > \frac{s}{ab}$ mamy zbieżność do dodatniego stanu stacjonarnego (\bar{V}, \bar{P}) , a przy $K = \frac{s}{ab}$ następuje bifurkacja siodło – węzeł. Taki typ bifurkacji oznacza, że zmienia się charakter punktu krytycznego z siodła na węzeł, przy czym dla bifurkacyjnej wartości $K = \frac{s}{ab}$ parametru K ze stanu stacjonarnego $(K, 0)$ bifurkuje dodatni stan stacjonarny i dla takiego K rozwiązania te sklejają się w jedno i choć nie przeprowadziliśmy analizy portretu fazowego dla tego przypadku, to doświadczenia związane z badaniem przebiegu orbit w pozostałych przypadkach pozwalają nam wnioskować, że także i dla bifurkacyjnej wartości K mamy globalną stabilność stanu stacjonarnego $(K, 0)$. Zauważmy dalej, że niewielka zmiana prawej strony układu nie prowadzi w przypadku tego modelu do drastycznych zmian zachowania rozwiązań, zatem taki układ nie ma już niepożądanego własności niestabilności strukturalnej charakteryzującej układ (6.1).

7.2. Model z kryjówkami dla ofiar

Inny sposób na uzyskanie stabilności strukturalnej stanowi przyjęcie założenia, że część ofiar jest stale niedostępna dla drapieżników — w środowisku występuje pewna liczba kryjówek, w których może się schować określona liczba ofiar. Niech liczba ofiar, która może się schować przed drapieżnikiem wynosi K . Jeśli zatem w chwili początkowej $t = 0$ liczebność całkowita ofiar $V(t)$ nie przekracza K , to wszystkie ofiary chowają się przed drapieżnikami i drapieżniki nie są w stanie nic upolować. W takiej sytuacji liczebność ofiar rośnie i w pewnym momencie przekracza K — dopiero wtedy możemy mówić o właściwym układzie drapieżnik – ofiara i zastosować odpowiedni model. W modelu tym zakładamy, że liczba ofiar dostępnych dla drapieżników wynosi $V - K$, $V > K$, zatem tylko takie ofiary mogą zostać upolowane i wpłynąć na rozwój populacji drapieżników. Wyjściowy model Lotki – Volterry przyjmuje przy takim założeniu postać

$$\begin{aligned}\dot{V} &= rV - a(V - K)P, \\ \dot{P} &= -sP + ab(V - K)P,\end{aligned}\tag{7.3}$$

przy czym parametry r , a , s , b mają takie same znaczenie jak dla układu (6.1), natomiast K oznacza tu liczbę kryjówek (w odróżnieniu od typowego oznaczenia pojemności środowiska).

Lokalne istnienie i jednoznaczność rozwiązań układu (7.3) ponownie wynikają z wielomianowej postaci prawej strony układu. Ponieważ model ma sens dla $V \geq K$, to ograniczymy nasze rozważania do przestrzeni fazowej $\mathcal{DF} = [K, +\infty) \times \mathbb{R}^+$. Zauważmy, że

Stwierdzenie 7.2. *Zbiór \mathcal{DF} jest przestrzenią niezmienniczą dla układu (7.3).*

Dowód. Niech $(V_0, P_0) \in \mathcal{DF}$ będzie warunkiem początkowym. Jeśli $P_0 = 0$, to $\dot{P} \equiv 0$ i rozwiązanie pozostaje na osi poziomej, co więcej, żadne rozwiązanie z wnętrza przestrzeni fazowej \mathcal{DF} nie osiągnie tego brzegu ze względu na jednoznaczność rozwiązań. Niech $V_0 = K$. Wtedy $\dot{V}(0) = rK > 0$, zatem rozwiązanie rośnie w pewnym przedziale czasu. Jeśli istnieje chwila $\tilde{t} > 0$, w której rozwiązanie osiąga punkt na brzegu, $V(\tilde{t}) = K$, to oczywiście $\dot{V}(\tilde{t}) = rK > 0$, zatem takie rozwiązanie nie może przekroczyć tego brzegu z prawej strony na lewą. Stąd $V(t) > K$ dla wszystkich $t > 0$. \square

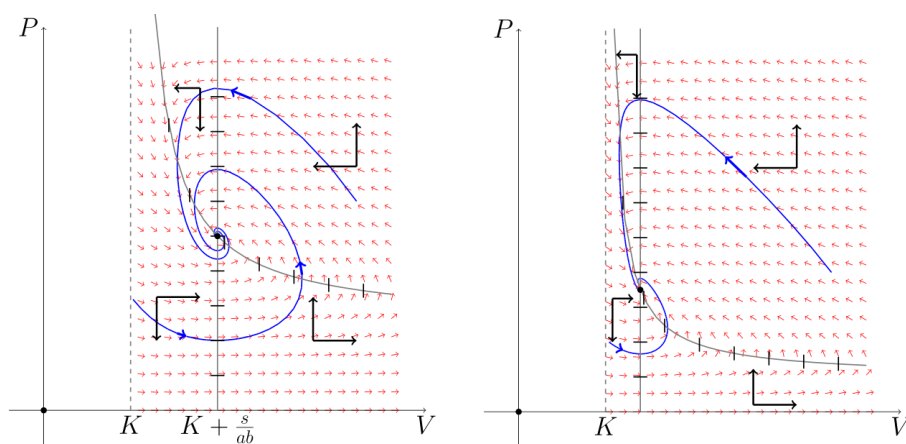
Z niezmienniczości zbioru \mathcal{DF} wynika od razu globalne istnienie rozwiązań: dla zmiennej V dostajemy oszacowanie liniowe $\dot{V} \leq rV$, czyli $V(t) \leq V_0 e^{rt}$ i dalej na każdym ograniczonym przedziale $[0, t^*)$ szacujemy $\dot{P} \leq AP$, gdzie $A = abV_0 e^{rt^*} - s$, podobnie jak w przypadku układu (6.1).

Asymptotykę układu (7.3) badamy znanymi już metodami. Przeprowadzenie analizy portretu fazowego z zastosowaniem np. kryterium Dulaca – Bedixsona pozwala wysnuć następujący wniosek

Wniosek 7.1. *Wszystkie rozwiązania układu (7.3) z warunkiem początkowym (V_0, P_0) , takim że $V_0 \geq K$, $P_0 > 0$, zbiegają asymptotycznie do dodatniego stanu stacjonarnego (\bar{V}, \bar{P}) , $\bar{V} = K + \frac{s}{ab}$, $\bar{P} = \frac{rb}{s}\bar{V}$.*

Zauważmy też, że podobnie jak dla układu (7.1), dodatni stan stacjonarny w zależności od parametrów (np. liczby ofiar mogących schować się w kryjówkach) może być albo ogniskiem, albo węzłem stabilnym, por. rys. 7.4.

Podsumowując — zarówno w przypadku modelu z kryjówkami dla ofiar jak i z pojemnością środowiska dostaliśmy globalną stabilność jednego ze stanów stacjonarnych, przy czym w przypadku modelu z kryjówkami jest to dodatni stan stacjonarny, natomiast w przypadku modelu z ograniczoną pojemnością środowiska może to być albo dodatni stan stacjonarny, albo stan



Rysunek 7.4. Portret fazowy układu (7.3) w przypadku gdy dodatni stan stacjonarny jest stabilnym ogniskiem (po lewej) lub węzłem (po prawej) oraz przykładowe krzywe fazowe w obu przypadkach.

brzegowy opisujący wyginięcie (ekstynkcję) populacji drapieżników. Jeśli stabilny stan stacjonarny jest węzłem, to rozwiązania od pewnego momentu \bar{t} pozostają monotoniczne, natomiast jeśli jest ogniskiem, to obserwujemy oscylacje i — ze względu na globalną stabilność — są to oscylacje gasnące.

8. Modele oddziaływań między dwiema populacjami III

8.1. Układ konkurujących gatunków

Korzystając z doświadczeń związanych z omówionymi już modelami oddziaływań typu drapieżnik – ofiara przejdziemy teraz do opisu oddziaływań innego typu. Jak poprzednio zakładamy, że w środowisku występują dwa gatunki \mathcal{P}_1 i \mathcal{P}_2 , ale tym razem osobniki tych gatunków konkurują ze sobą o zasoby środowiska. Zwykle budując taki model przyjmujemy, że jest to konkurencja o pożywienie, ale może ona dotyczyć także dostępnej przestrzeni życiowej, a często mamy do czynienia z obydwoma rodzajami konkurencji. Ponieważ gatunki konkurują ze sobą, to nie widać powodu do wyróżniania któregośkolwiek z nich, zatem opis powinien być symetryczny — jeśli zamienimy te gatunki i nazwiemy gatunek \mathcal{P}_i ($i = 1, 2$) gatunkiem \mathcal{P}_j ($j = 2, 1$) i na odwrót, to układ powinien pozostać taki sam.

Skoro mamy skorzystać z nabytych doświadczeń, zbudujemy model w oparciu o wewnętrzną dynamikę logistyczną dla każdego z gatunków \mathcal{P}_i . Wobec tego, jeśli w środowisku występuje tylko jeden z tych gatunków, to jego zagęszczenie $N_i(t)$ w chwili t opisuje równanie

$$\dot{N}_i = r_i N_i \left(1 - \frac{N_i}{K_i} \right),$$

gdzie r_i oznacza współczynnik rozrodczości netto dla gatunku \mathcal{P}_i , zaś K_i — pojemność środowiska dla tego gatunku. Oczywiście założenie ograniczonej pojemności środowiska implikuje też występowanie konkurencji między osobnikami tego samego gatunku. Tę konkurencję nazywamy w tym kontekście konkurencją wewnątrzgatunkową lub krócej — konkurencją wewnętrzną. Oprócz tego mamy też do czynienia z konkurencją zewnątrzgatunkową, albo konkurencją zewnętrzną, którą opisujemy podobnie do konkurencji wewnętrznej. Zakładamy więc, że osobniki konkurują ze sobą w trakcie spotkań między dwoma osobnikami różnego gatunku i liczba tych spotkań jest proporcjonalna do liczebności każdego z gatunków \mathcal{P}_i , $i = 1, 2$. Wobec tego składnik konkurencji zewnętrznej w obu równaniach układu jest proporcjonalny do $N_1(t)N_2(t)$.

Ostatecznie otrzymujemy dwóch układ równań

$$\begin{aligned} \dot{N}_1 &= r_1 N_1 \left(1 - \frac{N_1}{K_1} - a_{12} \frac{N_2}{K_2} \right), \\ \dot{N}_2 &= r_2 N_2 \left(1 - \frac{N_2}{K_2} - a_{21} \frac{N_1}{K_1} \right), \end{aligned} \tag{8.1}$$

gdzie a_{ij} są współczynnikami konkurencji zewnętrznej, przy czym dla uproszczenia obliczeń związanych z układem (8.1) współczynnik ten odnosi się do liczebności populacji \mathcal{P}_j w stosunku do jej pojemności środowiska. Zauważmy, że dzięki takiemu zapisowi możemy łatwo przeprowadzić ubezwymiarowanie układu (8.1) wprowadzając zmienne $x_i = \frac{N_i}{K_i}$, $i = 1, 2$. Przypomnijmy, że taką zamianę zmiennych wprowadziliśmy w przypadku dyskretnego równania logistycznego, co umożliwiło ograniczenie rozważań do odcinka $[0, 1]$. W nowych zmiennych układ (8.1) przyjmuje postać

$$\begin{aligned} \dot{x}_1 &= r_1 x_1 (1 - x_1 - a_{12} x_2), \\ \dot{x}_2 &= r_2 x_2 (1 - x_2 - a_{21} x_1), \end{aligned} \tag{8.2}$$

przy czym x_i odzwierciedla procentową eksploatację środowiska przez gatunek i .

Przeprowadzimy teraz analizę dynamiki modelu konkurujących gatunków opisanego przez układ (8.2) podobnie jak dla układu (6.1). Zaczniemy od zbadania podstawowych własności takich jak istnienie, jednoznaczność, nieujemność rozwiązań dla nieujemnych warunków początkowych.

Stwierdzenie 8.1. *Niech $x_i(0) = x_i^0 \geq 0$, $i = 1, 2$, będzie warunkiem początkowym dla układu (8.2). Dla danego warunku początkowego istnieje jednoznaczne, nieujemne rozwiązanie określone dla wszystkich $t \geq 0$. Co więcej, rozwiązanie jest ograniczone, $x_i(t) \leq \max\{x_i^0, 1\}$, $i = 1, 2$.*

Dowód. Istnienie i jednoznaczność łatwo wykazać korzystając z tego, że prawa strona układu jest funkcją klasy \mathbf{C}^1 . Nieujemność rozwiązań wynika np. z postaci logarytmicznej

$$\frac{\dot{x}_i}{x_i} = r_i(1 - x_i - a_{ij}x_j) \implies x_i(t) = x_i^0 \exp\left(r_i \int_0^t (1 - x_i(s) - a_{ij}x_j(s)) ds\right) \geq 0,$$

natomiast przedłużalność wnioskujemy bezpośrednio z nieujemności, gdyż

$$\dot{x}_i \leq r_i x_i,$$

a jeśli prawa strona układu ma liniowe oszacowanie, to rozwiązania istnieją dla wszystkich $t \geq 0$. Co więcej, ograniczoność rozwiązań także wynika z ich nieujemności, gdyż zamiast powyższego liniowego oszacowania możemy wziąć

$$\dot{x}_i \leq r_i x_i (1 - x_i)$$

i korzystając z nierówności różniczkowych oszacować z góry rozwiązania układu (8.1) przez rozwiązania równania logistycznego z pojemnością środowiska $K = 1$, analogicznie jak w przypadku układu (7.1). \square

Po zbadaniu podstawowych własności możemy przejść do przeanalizowania portretu fazowego. Izokliny zerowe układu (8.2) wyznaczamy jako

$$- \dot{x}_1 = 0 \iff x_1 = 0 \text{ lub } x_2 = \frac{1}{a_{12}}(1 - x_1);$$

$$- \dot{x}_2 = 0 \iff x_2 = 0 \text{ lub } x_1 = 1 - a_{21}x_2.$$

Wzajemne położenie izoklin nietrywialnych zależy od wielkości współczynników a_{ij} . Mamy trzy istotnie różne przypadki generyczne

1. jeden ze współczynników jest większy, a drugi mniejszy niż 1 — założymy, że $a_{12} > 1$ i $a_{21} < 1$, drugi układ parametrów implikuje taką samą dynamikę z dokładnością do zamiany miejscami gatunków \mathcal{P}_1 i \mathcal{P}_2 ;
2. $a_{12}, a_{21} > 1$;
3. $a_{12}, a_{21} < 1$.

Oprócz wymienionych powyżej przypadków generycznych mamy też przypadki niegeneryczne, gdy co najmniej jeden ze współczynników konkurencji jest równy 1. Analizę tych przypadków pomijamy. Od strony biologicznej nieprawdopodobne wydaje się, żeby wartość jakiegokolwiek parametru utrzymywała się na stałym poziomie — zwykle występują drobne wahania, a co za tym idzie przypadki niegeneryczne są nieistotne z biologicznego punktu widzenia.

Zbadajmy najpierw stany stacjonarne i ich stabilność. Mamy zawsze trzy stany stacjonarne na brzegu przestrzeni fazowej $(\mathbb{R}^+)^2$: $(0, 0)$, $(1, 0)$ i $(0, 1)$, które odzwierciedlają wymarcie co najmniej jednego z gatunków. Jeśli nietrywialne izokliny przecinają się, to mamy dodatni stan stacjonarny

$$(\bar{x}_1, \bar{x}_2) = \left(\frac{1-a_{12}}{1-a_{12}a_{21}}, \frac{1-a_{21}}{1-a_{12}a_{21}} \right).$$

Widzimy, że dodatni stan stacjonarny istnieje, jeśli albo oba współczynniki a_{12} , a_{21} są poniżej 1, albo oba są powyżej. W celu zbadania stabilności wyznaczamy macierz Jacobiego

$$MJ(x_1, x_2) = \begin{pmatrix} r_1(1 - 2x_1 - a_{12}x_2) & -r_1a_{12}x_1 \\ -r_2a_{21}x_2 & r_2(1 - 2x_2 - a_{21}x_1) \end{pmatrix}$$

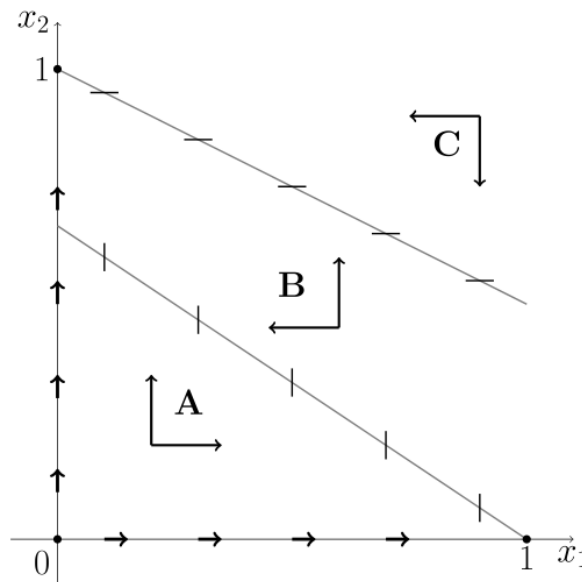
i obliczamy wartości własne

- dla $(0, 0)$ mamy $\lambda_i = r_i > 0$, $i = 1, 2$;
- dla $(1, 0)$: $\lambda_1 = -r_1 < 0$, $\lambda_2 = r_2(1 - a_{21})$ i dla $(0, 1)$ — symetrycznie;
- dla (\bar{x}_1, \bar{x}_2) wartości własne są rozwiązaniami równania charakterystycznego

$$\lambda^2 + \lambda(r_1\bar{x}_1 + r_2\bar{x}_2) + r_1r_2\bar{x}_1\bar{x}_2(1 - a_{12}a_{21}).$$

Wynika stąd, że $(0, 0)$ jest zawsze węzłem niestabilnym, stabilność $(1, 0)$ zależy od parametru a_{21} — jeśli $a_{21} > 1$, to mamy węzeł stabilny, a jeśli $a_{21} < 1$, to siodło (analogicznie dla $(0, 1)$ w zależności od a_{12}). Natomiast stabilność dodatniego stanu stacjonarnego zależy od iloczynu $a_{12}a_{21}$ — jeśli jest on większy niż 1, to (\bar{x}_1, \bar{x}_2) jest siodłem, a dla nierówności przeciwnej — stabilnym węzłem. We wszystkich tych przypadkach następuje bifurkacja siodło – węzeł dla krytycznej wartości parametrów (czyli gdy są równe 1).

Analizę zachowania rozwiązań w przestrzeni fazowej $(\mathbb{R}^+)^2$ zaczniemy od przypadku 1, tj. $a_{21} < 1 < a_{12}$. Przy takim układzie parametrów nietrywialne izokliny dla zmiennych x_1 i x_2 nie przecinają się w omawianej przestrzeni fazowej — nie istnieje dodatni stan stacjonarny. Wiemy już, że $(1, 0)$ jest siodłem, a $(0, 1)$ węzłem stabilnym. Izokliny dzielą przestrzeń fazową na 3 obszary: obszar A pod izokliną dla x_1 , obszar B pomiędzy izoklinami oraz C — powyżej izokliny dla x_2 , por. rys. 8.1. Zauważmy, że izokliny trywialne $x_1 = 0$ oraz $x_2 = 0$ stanowią



Rysunek 8.1. Szkic portretu fazowego układu (8.2) w przypadku gdy $a_{22} < 1 < a_{12}$.

rozwiązania układu. Rzeczywiście, jeśli $x_1^0 = 0$, to $\dot{x}_1 \equiv 0$ i stąd $x_1(t) = 0$ dla $t \geq 0$. Wtedy $\dot{x}_2 = r_2x_2(1 - x_2)$ i x_2 spełnia równanie logistyczne, zatem na osi pionowej rozwiązanie zbiega do punktu $(0, 1)$. Dokładnie takie samo rozumowanie pokazuje, że na osi poziomej rozwiązanie zbiega do $(1, 0)$. Ponieważ ten ostatni punkt jest siodłem, więc wyznaczyliśmy w ten sposób

rozmaitość stabilną. Z przebiegu pola wektorowego w otoczeniu punktu $(1, 0)$ wnioskujemy, że interesujący nas fragment rozmaitości niestabilnej przebiega w obszarze B . W obszarze A obie zmienne rosną, nie mogą pozostać w tym obszarze, bo musiałyby być zbieżne do punktu o dodatnich współrzędnych na izoklinie dla zmiennej x_1 — ale taki punkt musiałby wtedy być stanem stacjonarnym, a nie ma dodatniego stanu stacjonarnego w tym przypadku. Wobec tego rozwiązanie przechodzi do obszaru B . Z kolei jeśli punkt początkowy należy do C , to obie zmienne maleją i analogiczne rozumowanie pokazuje, że albo rozwiązanie zostaje w C i wtedy zbiega do stanu stacjonarnego $(0, 1)$, albo przechodzi do obszaru B . Przebieg pola wektorowego pokazuje, że rozwiązanie nie może wyjść z obszaru B , a ponieważ w tym obszarze jest także monotoniczne — x_1 maleje a x_2 rośnie — to zbiega do stanu stacjonarnego, czyli do $(0, 1)$. Rozmaitość niestabilna stanowi separatoryse, która oddziela rozwiązania zaczynające się w A (dokładniej, jeśli przeprowadzimy analizę przebiegu takich rozwiązań dla $t < 0$, to zauważymy, że zbiegają one do $(0, 0)$ przy $t \rightarrow -\infty$) i w związku z tym mające taką własność, że x_2 stale rośnie (pod izokliną dla zmiennej x_2), natomiast x_1 najpierw rośnie, osiąga wartość maksymalną przekraczając izoklinę, a następnie maleje — musi więc maleć do 0, od rozwiązań zaczynających się w C i przechodzących do B — takie rozwiązanie ma malejącą pierwszą współrzędną, natomiast druga współrzędna osiąga minimum na izoklinie, a następnie rośnie do 1. Ostatecznie wszystkie rozwiązania oprócz $x_2^0 = 0$ zbiegają asymptotycznie do stanu stacjonarnego $(0, 1)$ i są od pewnego miejsca monotoniczne. Wnioskujemy, że w tym przypadku gatunek \mathcal{P}_1 ginie, natomiast \mathcal{P}_2 stabilizuje się na poziomie pojemności środowiska.

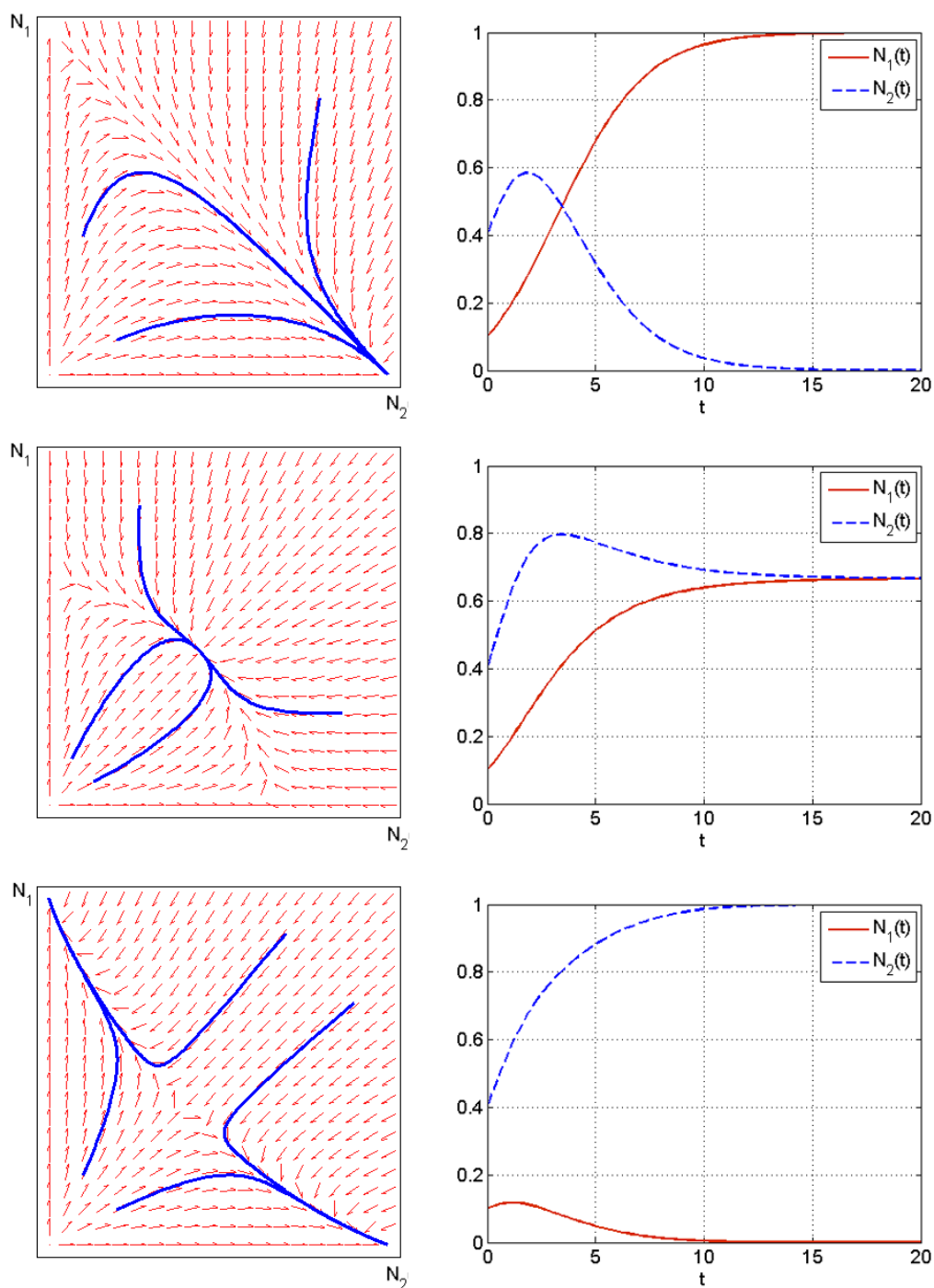
Wspomnieliśmy już, że jeśli układ parametrów a_{12} i a_{21} zmienia się, tj. gdy $a_{12} < 1 < a_{21}$ to rozwiązania zachowują się symetrycznie — dostajemy więc zbieżność wszystkich dodatnich rozwiązań do $(1, 0)$. Opisany przebieg rozwiązań wiąże się z tym, że konkurencja zewnętrzna okazuje się znacznie bardziej niebezpieczna ($a_{12} > 1$ lub $a_{21} > 1$) dla jednego z gatunków, a mało wpływa na drugi. Wtedy ten, dla którego jest niebezpieczna — ginie.

Jeśli oba parametry a_{12} , $a_{21} > 1$ lub a_{12} , $a_{21} < 1$, to nietrywialne izokliny przecinają się tworząc dodatni stan stacjonarny. Jeśli jednak oba parametry są powyżej 1, stan ten jest siodłem, natomiast w przeciwnym przypadku — węzłem stabilnym. W przestrzeni fazowej $(\mathbb{R}^+)^2$ możemy teraz wydzielić 4 obszary i analogicznie jak w przypadku 1., analizując przebieg pola wektorowego wnioskujemy, że w przypadku 3. dodatni stan stacjonarny jest globalnie stabilny we wnętrzu przestrzeni fazowej, natomiast w przypadku 2. rozmaitość stabilna tworzy separatoryse rozdzielającą baseny przyciągania stanów $(1, 0)$ i $(0, 1)$ — w zależności od warunku początkowego jeden gatunek przeżywa, a drugi ginie.

Podsumowując przeprowadzoną analizę zauważymy, że jeśli tylko $a_{12}a_{21} > 1$, co oznacza, że co najmniej jeden ze współczynników a_{12} , a_{21} przewyższa 1, to dynamika modelu odzwierciedla znaną zasadę ekologiczną mówiącą o konkurencyjnym wykluczaniu gatunków — jeśli nisze ekologiczne dwóch konkurujących gatunków zbyt się pokrywają, to jeden z gatunków wypiera drugi. Tylko przy bardzo szczególnych układach parametrów możliwe jest współistnienie takich gatunków w tym samym siedlisku.

8.2. Modelowanie symbiozy

Kolejnym typem oddziaływań występujących między dwoma gatunkami \mathcal{P}_1 , \mathcal{P}_2 jest symbioza. Analogicznie jak w przypadku modelu konkurujących gatunków budując model heurystyczny zakładamy, że wewnętrzną dynamikę każdego z gatunków można opisać za pomocą modelu logistycznego, ze względu na ograniczoną pojemność środowiska, a co za tym idzie — występowanie konkurencji wewnątrzgatunkowej. Oddziaływania międzygatunkowe pojawiają się oczywiście wtedy, gdy spotykają się osobniki różnych gatunków i — ponownie podobnie do poprzednich



Rysunek 8.2. Pole wektorowe układu (8.1) (po lewej) wraz z przykładowymi rozwiązaniami (po prawej) w przypadku gdy $a_{21} < 1 < a_{12}$ (u góry), $a_{12}, a_{21} < 1$ (po środku) lub $a_{12}, a_{21} > 1$ (u dołu).

przypadków — zakłada się, że wpływ tych oddziaływań na dynamikę każdego z gatunków zależy od liczby spotkań między osobnikami gatunków \mathcal{P}_1 oraz \mathcal{P}_2 .

Niech $N_i(t)$ oznaczają liczbę osobników (zagęszczenia) gatunków \mathcal{P}_i , $i = 1, 2$. Układ symbiotyczny możemy opisać za pomocą następujących równań

$$\begin{aligned}\dot{N}_1 &= r_1 N_1 \left(1 - \frac{N_1}{K_1} + b_{12} N_2\right), \\ \dot{N}_2 &= r_2 N_2 \left(1 - \frac{N_2}{K_2} + b_{21} N_1\right),\end{aligned}\tag{8.3}$$

gdzie tradycyjnie r_i oznacza współczynnik rozrodczości gatunku i , pojemności środowiska dla danego gatunku oznaczamy K_i , natomiast b_{ij} opisuje siłę wpływu oddziaływań symbiotycznych dla gatunku i w symbiozie z gatunkiem j .

Zauważmy, że lokalne istnienie, jednoznaczność i nieujemność rozwiązań układu (8.3) dla nieujemnego warunku początkowego wykazujemy dość łatwo — stosując podobne argumenty jak w przypadku układu (8.1). Natomiast określoność rozwiązań dla wszystkich $t > 0$ staje się tu własnością pożądaną, ale nie zawsze występującą. Rozpatrzmy następującą sytuację. Niech oba gatunki charakteryzują się tymi samymi parametrami, $r_1 = r_2$, $K_1 = K_2$ i $b_{12} = b_{21}$. Załóżmy też, że na początku liczebność obu gatunków jest jednakowa. Wobec tego, ze względu na całkowitą symetrię — liczebność tych gatunków jest jednakowa dla dowolnej chwili $t > 0$, dla której istnieje rozwiązanie układu (8.3). Skoro $N_1(t) = N_2(t)$, to

$$\dot{N}_1 = r_1 N_1 \left(1 - \frac{N_1}{K_1} + b_{12} N_1\right)$$

i jeśli $b_{12} > \frac{1}{K_1}$, to N_1 stale rośnie. Co więcej, $\dot{N}_1 > \gamma N_1^2$, gdzie $\gamma = b_{12} - \frac{1}{K_1} > 0$. Wiadomo natomiast, że rozwiązania równania $\dot{x} = \gamma x^2$ dla dodatniego warunku początkowego $x_0 > 0$ są określone w skończonym przedziale czasu $[0, \bar{t})$, ponieważ dla $\bar{t} = \frac{1}{\gamma x_0}$ zachodzi $\lim_{t \rightarrow \bar{t}^-} x(t) = +\infty$. Rzeczywiście, rozwiązując to równanie metodą rozdzielania zmiennych dostajemy $x(t) = \frac{x_0}{1 - \gamma x_0 t}$ i dla $t = \bar{t}$ mianownik rozwiązania zeruje się. Tego typu zachowania możemy się spodziewać za każdym razem, gdy w modelu występuje składnik kwadratowy z dodatnim współczynnikiem.

Przeprowadzimy teraz analizę portretu fazowego układu (8.3) w zależności od parametrów. Izokliny zerowe zadane są za pomocą następujących prostych

- dla zmiennej N_1 : $N_1 = 0$ lub $N_2 = \frac{1}{b_{12}} \left(\frac{N_1}{K_1} - 1\right)$;
- dla zmiennej N_2 : $N_2 = 0$ lub $N_2 = K_2 (1 + b_{21} N_1)$.

Widzimy więc, że niezerowe izokliny są prostymi o dodatnim współczynniku kierunkowym, zatem w zależności od parametrów mogą się przecinać lub nie w pierwszej ćwiartce $(\mathbb{R}^+)^2$, czyli naszej przestrzeni fazowej. Wobec tego układ ma albo 3 stany stacjonarne $(0, 0)$, $(K_1, 0)$ i $(0, K_2)$, jeśli izokliny się nie przecinają, albo 4 takie stany, jeśli się przecinają i wtedy mamy dodatni stan stacjonarny

$$\bar{N}_1 = \frac{1 + b_{12} K_2}{1 - b_{12} b_{21} K_1 K_2} K_1, \quad \bar{N}_2 = \frac{1 + b_{21} K_1}{1 - b_{12} b_{21} K_1 K_2} K_2.$$

Zauważmy, że stan ten istnieje przy założeniu $b_{12} b_{21} K_1 K_2 < 1$, co w praktyce oznacza, że przy zadanych pojemnościach środowiska dla rozważanych gatunków zysk płynący z oddziaływań symbiotycznych nie może być zbyt duży, aby gatunki mogły pozostawać w pewnej równowadze, opisanej przez stan stacjonarny. W takim układzie mamy $\bar{N}_i > K_i$, $i = 1, 2$, czyli współrzędne dodatniego stanu stacjonarnego przewyższają pojemności środowiska.

Wyznamy teraz wartości własne odpowiadające poszczególnym stanom stacjonarnym. Macierz Jacobiego układu (8.3) jest równa

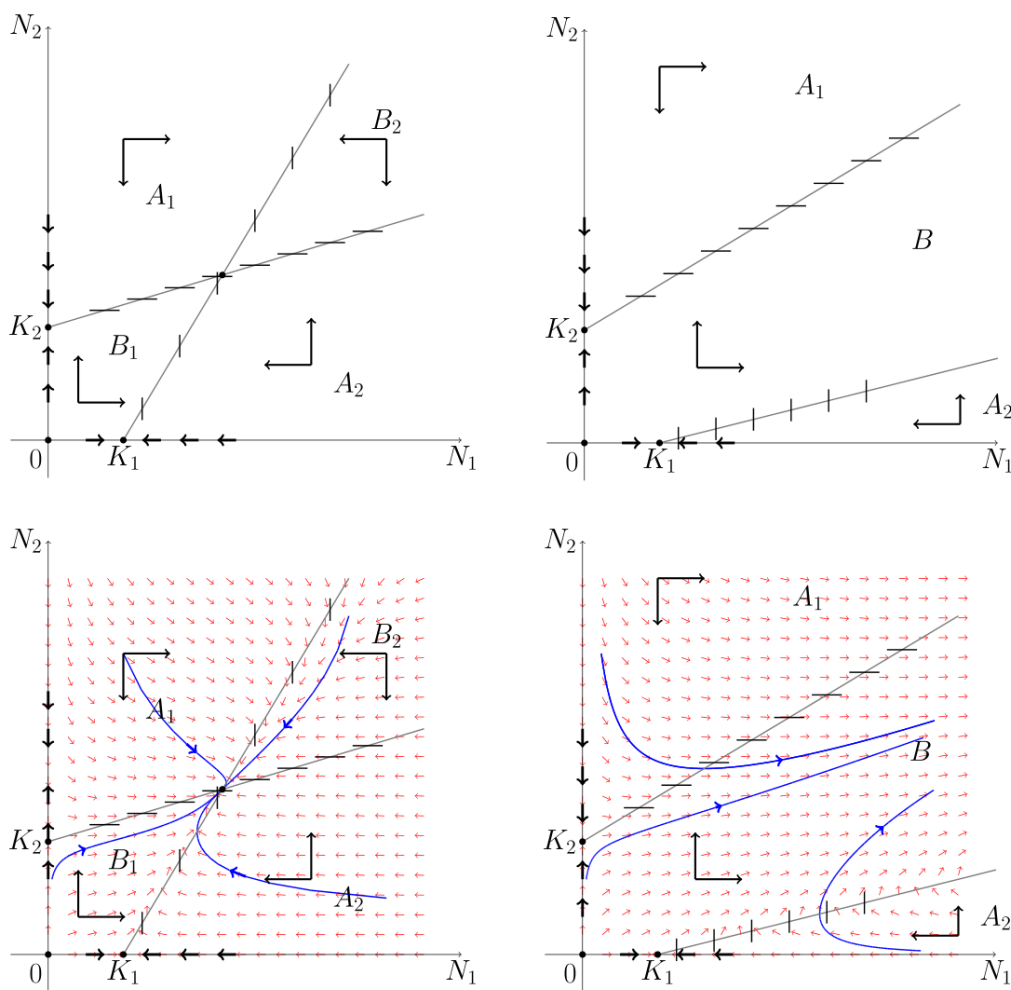
$$MJ(N_1, N_2) = \begin{pmatrix} r_1 \left(1 - 2\frac{N_1}{K_1} + b_{12} N_2\right) & r_1 b_{12} N_1 \\ r_2 b_{21} N_2 & r_2 \left(1 - 2\frac{N_2}{K_2} + b_{21} N_1\right) \end{pmatrix}\tag{8.4}$$

Podstawiając do wzoru (8.4) współrzędne poszczególnych stanów stacjonarnych otrzymujemy wartości własne

- dla $(0, 0)$ mamy $\lambda_i = r_i > 0$, $i = 1, 2$;
- dla $(K_1, 0)$: $\lambda_1 = -r_1 < 0$, $\lambda_2 = r_2(1 + b_{21}K_1) > 0$ i symetrycznie dla $(0, K_2)$: $\lambda_1 = r_1(1 + b_{12}K_2) > 0$, $\lambda_2 = -r_2 < 0$;
- dla (\bar{N}_1, \bar{N}_2) : λ_i , $i = 1, 2$ są rozwiązaniami równania charakterystycznego

$$\lambda^2 + \left(r_1 \frac{\bar{N}_1}{K_1} + r_2 \frac{\bar{N}_2}{K_2} \right) \lambda + r_1 r_2 \frac{\bar{N}_1 \bar{N}_2}{K_1 K_2} (1 - b_{12} b_{21} K_1 K_2) \implies \lambda_1, \lambda_2 < 0.$$

Wobec tego $(0, 0)$ jest węzłem niestabilnym, $(K_1, 0)$ i $(0, K_2)$ są siodłami, natomiast dodatni stan stacjonarny, jeśli istnieje, jest węzłem stabilnym.



Rysunek 8.3. Szkic portretów fazowych układu (8.3) z zaznaczonymi izoklinami zerowymi i kierunkami przebiegu poszczególnych zmiennych (na górze) oraz portrety fazowe z naniesionymi przykładowymi krzywymi fazowymi (na dole) w przypadku gdy istnieje dodatni stan stacjonarny (po lewej) i gdy nie istnieje (po prawej).

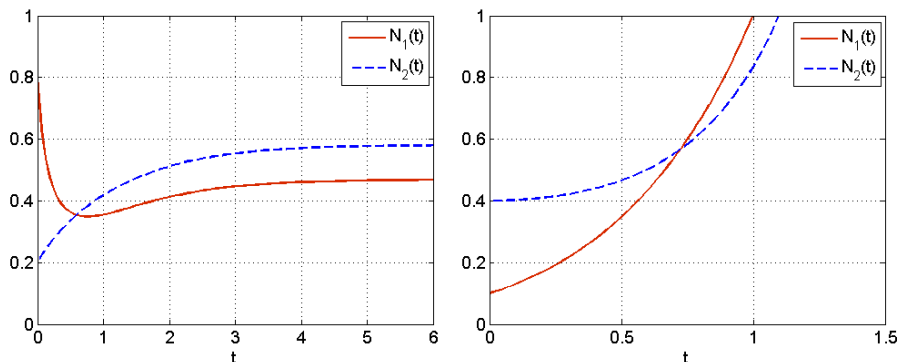
Na rysunku 8.3 widzimy przebieg pola wektorowego (\dot{N}_1, \dot{N}_2) w obu przypadkach. Jeśli izokliny nie przecinają się w analizowanej przestrzeni fazowej, to łatwo możemy wywnioskować, że obie współrzędne rozwiązania są nieograniczone i albo są stale rosnące, albo na początku maleją

osiągając w pewnej chwili t_m minimum, a potem rosną dla $t > t_m$. Rzeczywiście, przestrzeń fazową możemy podzielić na trzy obszary, oznaczone na portrecie fazowym A_1 (ponad izokliną zerową dla N_2), A_2 (pod izokliną zerową dla N_1) i B (pomiędzy izoklinami). W obszarze A_1 zmienna N_2 maleje, a N_1 rośnie. Wewnątrz tego obszaru nie ma stanów stacjonarnych, a stan stacjonarny na brzegu jest siodłem, przy czym jego rozmaitość stabilna pokrywa się z osią N_2 , zatem rozwiązanie nie może zbiegać do tego stanu. Wnioskujemy więc, że rozwiązanie opuszcza ten obszar i przechodzi do obszaru B . Symetrycznie przebiegają orbity w obszarze A_2 , zatem także wchodzi do B . Z kolei w B obie zmienne są rosnące, przy czym rosną nieograniczenie — gdyby wzrost którejś z nich był ograniczony, to musiałaby być zbieżna, co oznacza, że pochodna też byłaby ograniczona, a nawet dążyłaby do 0, co jest niemożliwe ze względu na przebieg pola wektorowego.

Wróćmy teraz do problemu określoności rozwiązań dla wszystkich $t > 0$. Zauważmy, że gdyby np. zmienna $N_1(t)$ rosła do $+\infty$ w skończonym czasie, to oznaczałoby, że pewnej chwili musiałaby przeciąć izoklinę zerową zmiennej N_2 , co znów nie zgadza się z przebiegiem pola. Zatem obie zmienne są określone dla wszystkich $t > 0$ i $\lim_{t \rightarrow \infty} N_i(t) = +\infty$, $i = 1, 2$, oczywiście poza przypadkiem omówionym powyżej, gdy poszczególne współczynniki dla obu gatunków pokrywają się.

Jeśli izokliny zerowe przecinają się w pierwszej ćwiartce układu współrzędnych, to przestrzeń fazowa dzieli się na 4 obszary, A_1 — ponad obydwojema izoklinami, A_2 — pod nimi, B_1 i B_2 — pomiędzy nimi. W obszarze B_1 obie zmienne rosną, natomiast w B_2 — obie maleją. Zauważmy, że każdy prostokąt o jednym wierzchołku w początku układu współrzędnych i dwóch bokach wzdłuż osi oraz naprzeciwległym wierzchołku w obszarze B_2 jest zbiorem niezmienniczym, gdyż pole wektorowe wchodzi do wewnątrz. Wobec tego każde rozwiązanie pozostaje ograniczone i analizując pole wektorowe dochodzimy do wniosku, że każde rozwiązanie układu (8.3) z dodatnim warunkiem początkowym zbiega do punktu krytycznego (\bar{N}_1, \bar{N}_2) .

Zauważmy, że symbioza zawsze działa korzystnie na liczebność gatunków, które w taki sposób na siebie oddziałują, ale tylko w przypadku drugim, gdy efekt tych oddziaływań nie jest nadmierny, gatunki mogą współistnieć w środowisku. Natomiast jeśli oddziaływania symbiotyczne przekraczają pewien próg, to prowadzi do nieograniczonego wzrostu obu populacji, co oczywiście w warunkach rzeczywistych nie może mieć miejsca i musi zakończyć się katastrofą, por. rys. 8.4.



Rysunek 8.4. Przebieg przykładowych rozwiązań układu (8.3) w przypadku gdy istnieje dodatni stan stacjonarny (po lewej) i gdy ten stan nie istnieje (po prawej).

9. Doświadczenia Mendla: łańcuchy Markowa w klasycznej genetyce I

W kolejnych wykładach przejdziemy do omówienia podstaw genetyki klasycznej, w szczególności praw dziedziczenia postulowanych przez Mendla na podstawie wyników przeprowadzonych przez niego doświadczeń. **Gregor Mendel**, żyjący w XIX wieku czeski zakonnik, przez wiele lat zajmował się pracą w ogrodzie. Hodował różne rośliny, w szczególności groszek w dwóch odmianach — o strąkach zielonych i żółtych. Mendel zauważył, że można wyróżnić groszki „czysto żółte”, „czysto zielone” i „mieszane”. Jeśli krzyżuje się dwie rośliny, jedną „czysto żółtą”, a drugą „czysto zieloną”, to w następnym pokoleniu otrzymuje się rośliny o strąkach zielonych, natomiast przy dalszym krzyżowaniu ze sobą tak otrzymanych roślin, w kolejnym pokoleniu $3/4$ osobników ma strąki zielone, a $1/4$ żółte. Wyniki swoich badań opublikował w 1866 roku i na ich podstawie wysnuł teorię, którą obecnie nazywamy całkowitą dominacją genów.

W opisie doświadczeń Mendla będziemy używać stosowanych obecnie w genetyce pojęć, w szczególności jednostki kodujące dane cechy nazywać będziemy genami. Cecha badana przez Mendla, czyli kolor strąków groszku, kodowana jest przez parę genów, **dziedziczonych** po jednym od każdego z rodziców. Gen występuje w dwóch odmianach, które nazywamy **allelami**: **dominującej**, którą oznaczymy D i **recesywnej** R . Mamy więc trzy **genotypy**: DD , DR , RR , ponieważ nie rozróżniamy kolejności występowania genów, zatem DR i RD dają ten sam genotyp. Jeśli w genotypie pojawia się choć jeden gen dominujący, to odpowiadająca mu cecha uwidacznia się w danym osobniku, czyli taki groszek ma strąki zielone. Rozróżniamy więc dwa **fenotypy** — genotypom DD i DR odpowiada fenotyp zielony, natomiast RR ma fenotyp żółty. W zależności od genotypu, stosujemy też następujące nazewnictwo:

- DD to osobnik (czysto) dominujący;
- DR to hybryda;
- RR to osobnik (czysto) recesywny.

W teorii Mendla zakłada się, że osobniki łączą się losowo i potomek dziedziczy losowo po jednym genie od każdego z rodziców, przy czym wybór genów od rodziców jest niezależny. Sprawdźmy więc, jaki genotyp ma potomstwo ustalonej pary rodziców:

- $DD + DD \rightarrow DD$, jeśli krzyżujemy dwa osobniki dominujące, to z prawdopodobieństwem 1 potomek dziedziczy od każdego z rodziców gen dominujący D , zatem każdy potomek ma ten sam genotyp DD ;
- $RR + RR \rightarrow RR$, podobnie jeśli krzyżujemy dwa osobniki recesywne, to z prawdopodobieństwem 1 potomek dziedziczy od każdego z rodziców gen dominujący R , zatem każdy potomek ma ten sam genotyp RR ;
- $DD + RR \rightarrow DR$, jeśli krzyżujemy osobnika dominującego z recesywnym, to z prawdopodobieństwem 1 potomek dziedziczy od jednego z rodziców gen D , a od drugiego gen R , zatem każdy potomek jest hybrydą;
- $DD + DR \rightarrow \frac{1}{2}DD + \frac{1}{2}DR$, przy krzyżowaniu osobnika dominującego z hybrydą od osobnika dominującego potomek dziedziczy gen D z prawdopodobieństwem 1, natomiast od hybrydy — albo gen D z prawdopodobieństwem $1/2$, albo gen R , także z prawdopodobieństwem $1/2$, zatem potomek jest z prawdopodobieństwem $1/2$ albo hybrydą, albo dominantą;

- $RR + DR \longrightarrow \frac{1}{2}RR + \frac{1}{2}DR$, podobnie przy krzyżowaniu osobnika recesywnego z hybrydą potomek jest z prawdopodobieństwem $1/2$ albo recesywny, albo hybrydyczny;
- $DR + DR \longrightarrow \frac{1}{4}DD + \frac{1}{4}RR + \frac{1}{2}DR$, jeśli natomiast krzyżujemy dwie hybrydy, to z prawdopodobieństwem $1/2$ osobnik dziedziczy albo D , albo R od każdego z rodziców, stąd mamy rozkład genotypów potomka $1/4DD$, $1/4RR$ oraz $1/2DR$ (pamiętając, że nie rozróżniamy układu DR i RD).

Przedstawiony powyżej rozkład genotypów pokazuje, skąd w doświadczeniu Mendla wzięły się proporcje $3/4$ strąków zielonych i $1/4$ strąków żółtych. Jeśli bowiem Mendel skrzyżował w pierwszym pokoleniu osobnika czysto dominującego z czysto recesywnym, to każdy osobnik potomny był hybrydą. Następnie krzyżując hybrydy, jeśli tylko potomków w drugiej generacji było dostatecznie dużo, otrzymywał rozkład genotypów odpowiadający proporcjom: po $1/4$ osobników dominujących i recesywnych oraz $1/2$ hybryd, co uzewnętrzniało się jako $1/4$ fenotypu żółtego i $3/4$ fenotypu zielonego.

Tego typu doświadczenia można oczywiście powtarzać, ale jeśli znamy genotyp rodziców, to wiemy też, jaki powinien być wynik doświadczenia. Znacznie ciekawsze doświadczenie może polegać na tym, że na podstawie otrzymanych wyników eksperymentów chcemy wnioskować na temat genotypu rodziców. Opiszemy trzy takie doświadczenia, które nazwiemy ciągłym krzyżowaniem z osobnikiem dominującym, recesywnym i hybrydą. Każde z tych doświadczeń przeprowadzamy w taki sam sposób: ustalamy jednego z rodziców (wybieramy osobnika o danym genotypie) i krzyżujemy go z osobnikiem o genotypie nieznanym. Zakładamy, że mamy dostatecznie dużo potomstwa, tak by rozkład genotypów odpowiadał prawdopodobieństwom obliczonym na podstawie prawa Mendla, wybieramy losowo jednego potomka i ponownie krzyżujemy go z tym samym ustalonym osobnikiem. Doświadczenie powtarzamy wielokrotnie. Otrzymujemy więc ciąg doświadczeń, w których wynik kolejnego doświadczenia zależy od wyniku poprzedniego, mamy więc **łańcuch Markowa**. W celu zinterpretowania wyników eksperymentów przypomnimy podstawowe pojęcia i twierdzenia z teorii łańcuchów Markowa [7, 1].

9.1. Łańcuchy Markowa

Zajmiemy się teraz opisem ciągu zależnych doświadczeń losowych. Niech \mathcal{D} oznacza pewne doświadczenie, którego zbiór możliwych wyników $\mathcal{W}_1, \mathcal{W}_2, \dots, \mathcal{W}_n$, $n \in \mathbb{N}$, znamy. W trakcie powtarzania tego doświadczenia w próbie z numerem $t \in \mathbb{N}$ dostajemy wynik o numerze $X_t \in \{1, 2, \dots, n\}$, przy czym prawdopodobieństwa różnych wartości \mathcal{W}_{X_t} w ogólnym przypadku mogą zależeć od wyników wszystkich poprzednich doświadczeń. Taki obiekt nazwiemy skończonym procesem stochastycznym. Z kolei łańcuch Markowa to pewien szczególny rodzaj procesu stochastycznego, w którym wynik kolejnej próby zależy tylko od wyniku próby poprzedniej.

Definicja 9.1. Niech $\{X_t\}_{t=0}^{\infty}$ oznacza ciąg zmiennych losowych o wartościach całkowitych. Jeśli w próbie z numerem t zrealizowało się zdarzenie \mathcal{W}_j , to przyjmujemy $X_t = j$. Ciąg zmiennych losowych $\{X_t\}_{t=0}^{\infty}$ nazwiemy łańcuchem Markowa, jeśli

$$P(X_t = j | X_0 = k_0, X_1 = k_1, \dots, X_{t-1} = i) = P(X_t = j | X_{t-1} = i),$$

przy czym jeśli $P(X_t = j | X_{t-1} = i)$ nie zależy od numeru próby t , to taki łańcuch nazywamy jednorodnym. Wyniki poszczególnych prób nazwiemy stanami łańcucha.

Podamy teraz kilka prostych przykładów:

- **Rzut monetą** to jeden z najprostszych możliwych procesów stochastycznych. Mamy dwa możliwe wyniki każdej próby orzeł \mathcal{O} i reszka \mathcal{R} , przy czym zawsze prawdopodobieństwo otrzymania każdego z tych dwóch wyników wynosi $1/2$, niezależnie od tego, jakie wyniki otrzymywaliśmy w przeszłości. Wobec tego rzut monetą jest przykładem jednorodnego łańcucha Markowa.

— **Rosyjska ruletka I** to wersja gry, w której używa się sześciostrzałowego rewolweru z jedną kulą. Po zakręceniu bębniem oddaje się strzał i w pierwszej próbie wynik z prawdopodobieństwem $1/6$ jest martwy \mathcal{M} , a z prawdopodobieństwem $5/6$ — żywy \mathcal{Z} . Wynik kolejnej próby zależy od poprzedniego. Jeśli osoba, do której się strzela, została zastrzelona w poprzedniej próbie, to oczywiście w każdej następnej jest martwa z prawdopodobieństwem 1 , natomiast jeśli jest żywa, to w drugiej próbie będzie martwa z prawdopodobieństwem $4/5$, w trzeciej z prawdopodobieństwem $3/4$, następnie $2/3$, $1/2$ i ostatecznie w szóstej próbie wynik jest zawsze \mathcal{M} z prawdopodobieństwem 1 . Widzimy więc, że taka wersja rosyjskiej ruletki stanowi łańcuch Markowa, ale nie jest to łańcuch jednorodny.

— **Rosyjska ruletka II** to inna wersja, w której za każdym razem używa się sześciostrzałowego rewolweru z jedną kulą, ale przed każdą próbą ponownie kręci się bębniem. Zatem przy każdej próbie otrzymywane wyniki są jednakowe: \mathcal{M} z prawdopodobieństwem $1/6$ oraz \mathcal{Z} z prawdopodobieństwem $5/6$. Mamy więc jednorodny łańcuch Markowa.

W dalszej części wykładu zajmiemy się tylko łańcuchami jednorodnymi, wobec tego nie będziemy już powtarzać określenia „jednorodny łańcuch Markowa”, tylko pisząc „łańcuch Markowa” będziemy mieli na myśli łańcuch jednorodny. Taki łańcuch Markowa możemy w jednoznaczny sposób zdefiniować podając jego stany, czyli możliwe wyniki kolejnych eksperymentów, oraz macierz przejścia $\mathbf{P} = (p_{ij})_{i,j=1}^n$, czyli prawdopodobieństwa p_{ij} przejścia ze stanu \mathcal{W}_i w próbie $t - 1$ do stanu \mathcal{W}_j w kolejnej próbie t . Macierz przejścia ma oczywiście następujące własności:

$$p_{ij} \geq 0, \quad \sum_{j=1}^n p_{ij} = 1, \quad i = 1, \dots, n,$$

ponieważ dla dowolnych $i, j \in \{1, \dots, n\}$ liczba p_{ij} jest prawdopodobieństwem, a jeśli łańcuch znajduje się w pewnym stanie \mathcal{W}_i w chwili $t - 1$, to musi przejść do jakiegoś stanu \mathcal{W}_j w chwili t . Macierze o tych własnościach nazywamy **macierzami stochastycznymi**. Pełny opis łańcucha Markowa możemy także przedstawić za pomocą grafu skierowanego ważonego.

Definicja 9.2. Grafem skierowanym nazywamy parę (\mathbb{W}, S) , gdzie \mathbb{W} jest zbiorem elementów nazywanych wierzchołkami grafu, natomiast S stanowi zbiór uporządkowanych par wierzchołków, które nazywamy krawędziami grafu.

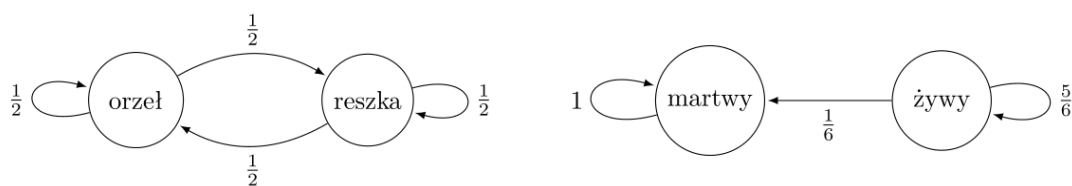
W przypadku grafu ważonego każdej istniejącej krawędzi prowadzącej od wierzchołka i do j przypisujemy pewną wagę w_{ij} . Graf opisujący łańcuch Markowa budujemy w następujący sposób

- wierzchołek z numerem i , $i \in \{1, \dots, n\}$ odpowiada wynikowi \mathcal{W}_i ;
- od wierzchołka z numerem i do wierzchołka z numerem j przeprowadzamy krawędź (i, j) , o ile $p_{ij} > 0$, a jeśli $p_{ij} = 0$, to takiej krawędzi nie ma;
- zbudowanej krawędzi (i, j) nadajemy wagę $w_{ij} = p_{ij}$.

Zaprezentujemy teraz macierze przejścia i grafy przejścia, por. rys. 9.1, odpowiadające wymienionym w przykładach łańcuchom: rzutowi monetą i rosyjskiej ruletce w wersji II.

$$\mathbf{P}_{\text{rm}} = \begin{pmatrix} 0,5 & 0,5 \\ 0,5 & 0,5 \end{pmatrix}, \quad \mathbf{P}_{\text{rr}} = \begin{pmatrix} 1 & 0 \\ 1/6 & 5/6 \end{pmatrix}.$$

Macierz przejścia \mathbf{P} odzwierciedla dynamikę łańcucha w jednym kroku, natomiast my jak zwykle chcemy zbadać, co będzie się działo asymptotycznie, czyli czego możemy się spodziewać, jeśli będziemy nasze doświadczenia powtarzać dostatecznie długo. Do tego celu musimy obliczyć prawdopodobieństwa przejścia wyższego rzędu, czyli $p_{ij}^{(t)}$, które opisuje przejście ze stanu \mathcal{W}_i do



Rysunek 9.1. Macierze przejścia i grafy przejścia odpowiadające rzutowi monetą (z lewej) i rosyjskiej ruletce w wersji II (z prawej).

stanu \mathcal{W}_j w t krokach, czyli po wykonaniu t prób. Mamy więc $p_{ij}^{(t)} = P(X_t = j | X_0 = i)$. Zgodnie ze wzorem na prawdopodobieństwo całkowite dla $t > 1$ otrzymujemy

$$p_{ij}^{(t)} = P(X_t = j | X_0 = i) = \sum_{s=1}^n P(X_{t-1} = s | X_0 = i) p_{sj} = \sum_{s=1}^n p_{is}^{(t-1)} p_{sj},$$

czyli rekurencyjną zależność $p_{ij}^{(t)}$ od prawdopodobieństw niższego rzędu $p_{is}^{(t-1)}$. Stąd jeśli $\mathbf{P}^{(t)} = (p_{ij}^{(t)})_{i,j=1}^n$ jest macierzą przejścia wyższego rzędu, to dostajemy zależność

$$\mathbf{P}^{(t)} = \mathbf{P}^{(t-1)}\mathbf{P} \implies \mathbf{P}^{(t)} = \mathbf{P}^t.$$

Wnioskujemy zatem, że rozkład łańcucha w dowolnym kroku jest jednoznacznie wyznaczony przez macierz przejścia i początkowy rozkład prawdopodobieństwa $p_i^0 = P(X_0 = i)$, $i = 1, \dots, n$.

W kontekście opisywanych powyżej doświadczeń genetycznych chcemy poznać asymptotykę dwóch konkretnych typów łańcuchów Markowa. Zanim jednak przejdziemy do badania asymptotyki, podamy potrzebne definicje, w tym klasyfikację stanów łańcucha.

9.2. Klasyfikacja stanów i łańcuchów

Zacniemy od podziału stanów łańcucha na pewne specjalne grupy.

Definicja 9.3. Stan \mathcal{W}_i nazwiemy **nieistotnym**, jeśli istnieje taki stan \mathcal{W}_j i liczba $t_0 \in \mathbb{N} \setminus \{0\}$, że $p_{ij}^{(t_0)} > 0$ oraz $p_{ji}^{(t)} = 0$ dla dowolnego $t \in \mathbb{N}$. W przypadku przeciwnym stan \mathcal{W}_i nazywamy **istotnym**.

Definicja 9.4. Stany istotne $\mathcal{W}_i, \mathcal{W}_j$ nazywamy **komunikującymi się**, jeśli istnieją takie $t, s \in \mathbb{N} \setminus \{0\}$, dla których $p_{ij}^{(t)} > 0$ i $p_{ji}^{(s)} > 0$.

Podzielmy teraz wszystkie stany danego łańcucha na podklasy: niech \mathcal{S}^0 oznacza klasę wszystkich stanów nieistotnych, następnie wybierzmy jakikolwiek stan istotny \mathcal{W}_i i oznaczmy przez \mathcal{S}^i zbiór wszystkich stanów komunikujących się z \mathcal{W}_i . Ten zbiór stanów nazywa się także czasem zbiorem **stochastycznie zamkniętym**. W taki sposób podzielimy łańcuch na rozłączne klasy stanów komunikujących się $\mathcal{S}^1, \dots, \mathcal{S}^N$, $N \leq n$. Jeśli łańcuch znalazł się w pewnym momencie w stanie istotnym \mathcal{W}_i , to już nigdy nie wyjdzie z odpowiadającej mu klasy.

Definicja 9.5. Jeśli klasa \mathcal{S}^i składa się z jednego stanu \mathcal{W}_i , to taki stan nazywamy **pochłaniającym** (absorbującym).

Definicja 9.6. Łańcuch Markowa składający się z jednej klasy stanów istotnych komunikujących się nazywamy łańcuchem **nieprzywiedlnym**. Jeśli łańcuch zawiera więcej niż jedną klasę, to jest **przywiedlny**.

Badając łańcuch Markowa sprowadzamy najpierw jego macierz do postaci kanonicznej numerując stany w taki sposób, aby na początku znalazły się stany nieistotne, a następnie poszczególne stany z kolejnych klas. Macierz ma wtedy postać

$$\mathbf{P} = \begin{array}{c} \begin{array}{c} \mathcal{S}^0 \\ \mathcal{S}^1 \\ \mathcal{S}^2 \\ \vdots \end{array} \begin{array}{cccc} & \mathcal{S}^0 & \mathcal{S}^1 & \mathcal{S}^2 \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \begin{array}{c} \vdots \\ \vdots \end{array} & 0 & 0 \\ 0 & 0 & \begin{array}{c} \vdots \\ \vdots \end{array} & 0 \\ \vdots & & & \end{array} \end{array},$$

a powstałe macierze $\begin{pmatrix} \vdots & \vdots \\ \vdots & \vdots \end{pmatrix}$ są macierzami stochastycznymi i każda z nich odpowiada łańcuchowi nieprzywiedlnemu.

Oznaczmy

$$f_j^{(t)} = P(X_t = j, X_{t-1} \neq j, \dots, X_1 \neq j | X_0 = j), \quad F_j = \sum_{t=1}^{\infty} f_j^{(t)},$$

gdzie $f_j^{(t)}$ oznacza prawdopodobieństwo, że układ po wyjściu ze stanu \mathcal{W}_j wróci do tego stanu po raz pierwszy po t krokach, a F_j to prawdopodobieństwo, że po wyjściu z tego stanu wróci do niego kiedykolwiek.

- Stan \mathcal{W}_j nazywamy **powracającym**, jeśli $F_j = 1$, natomiast **niepowracającym**, gdy $F_j < 1$.
- Stan \mathcal{W}_j nazywamy **zerowym**, jeśli $p_{jj}^{(t)} \rightarrow 0$ przy $t \rightarrow \infty$.
- Stan \mathcal{W}_j nazywamy **okresowym** o okresie d_j , jeśli powrót z dodatnim prawdopodobieństwem możliwy jest w liczbie kroków równej wielokrotności d_j i d_j jest najmniejszą liczbą o tej własności (czyli d_j jest NWD zbioru $\{t : p_{jj}^{(t)} > 0\}$).
- Stan powracający, niezerowy i nieokresowy nazywamy **ergodycznym**.

Wymienimy teraz dwa typy łańcuchów, które są istotne z punktu widzenia doświadczeń genetycznych, jakie za ich pomocą chcemy opisywać.

Definicja 9.7. Łańcuchem **pochłaniającym** (absorbującym) nazwiemy taki łańcuch przywiedlny, którego klasy stanów istotnych składają się z pojedynczych stanów pochłaniających.

Drugi typ stanowią oczywiście łańcuchy nieprzywiedlne. Każdy typ łańcucha przywiedlnego możemy analizować korzystając z własności łańcuchów pochłaniających i łańcuchów nieprzywiedlnych. Dopóki łańcuch nie trafi do któregoś ze zbiorów \mathcal{S}^i , $i > 0$, możemy traktować łańcuch jak łańcuch pochłaniający, natomiast jak już trafi do zbioru \mathcal{S}^i , $i > 0$, to ten zbiór tworzy łańcuch nieprzywiedlny.

Wróćmy teraz do naszych prostych przykładów i spróbujmy w ich kontekście zastosować poznane definicje.

- Rzut monetą: zauważmy, że prawdopodobieństwa przejścia z każdego z dwóch stanów łańcucha \mathcal{O} , \mathcal{R} do każdego z nich jest zawsze równe $1/2$, bez względu na stan w kroku poprzednim,

zatem $\mathbf{P}^t = \mathbf{P}$, czyli oba stany są istotne i komunikujące się, tworzą więc łańcuch nieprzywiedlny.

- Rosyjska ruletka II: stan \mathcal{M} jest oczywiście pochłaniający, natomiast stan \mathcal{Z} tworzy zbiór \mathcal{S}^0 i jest stanem nieistotnym, ponieważ $p_{21} = 1/6$, natomiast $p_{12}^{(t)} = 0$ dla dowolnego $t \in \mathbb{N}$, gdyż $\mathcal{M} = \mathcal{W}_1$ jest pochłaniający, czyli łańcuch ma własności wymienione w def. 9.7, zatem stanowi łańcuch pochłaniający.

Zauważmy, że asymptotycznie te dwa typy łańcuchów (czyli nieprzywiedlny i pochłaniający) mają diametralnie różne własności. W łańcuchu nieprzywiedlnym każdy stan jest osiągalny z dodatnim prawdopodobieństwem w dowolnej próbie, natomiast w łańcuchu pochłaniającym w pewnej próbie łańcuch wynosi trafia do jakiegoś stanu pochłaniającego, z którego już nie wychodzi. W przypadku rzutu monetą, ponieważ dla dowolnego $t \in \mathbb{N} \setminus \{0\}$ zachodzi $\mathbf{P}_{\text{rm}}^t = \mathbf{P}_{\text{rm}}$, więc bez względu na liczbę prób zawsze średnio dostajemy połowę orłów i połowę reszek. Z kolei w przypadku rosyjskiej ruletki prawdopodobieństwo pozostania żywym maleje do 0 wraz z rosnącą liczbą prób, co łatwo możemy bezpośrednio obliczyć iterując macierz przejścia

$$\mathbf{P}_{\text{rr}}^t = \begin{pmatrix} 1 & 0 \\ 1 - \left(\frac{5}{6}\right)^t & \left(\frac{5}{6}\right)^t \end{pmatrix}.$$

Co ciekawe, oczekiwane liczba kroków przed absorpcją wynosi 6, czyli tyle, ile maksymalna liczba kroków możliwa w wersji I. Wykażemy to poniżej.

10. Doświadczenia Mendla: łańcuchy Markowa w klasycznej genetyce II

10.1. Łańcuchy pochłaniające i ciągłe krzyżowanie z dominantą

Dla łańcuchów pochłaniających postać kanoniczna macierzy przejścia przedstawia się następująco

$$\mathbf{P} = \begin{pmatrix} \mathbb{I} & 0 \\ R & Q \end{pmatrix},$$

gdzie \mathbb{I} oznacza macierz jednostkową wymiaru $m \times m$ reprezentującą m stanów pochłaniających, R jest macierzą wymiaru $(n - m) \times m$ przejścia ze stanów nieistotnych do stanów pochłaniających, a Q — macierzą $(n - m) \times (n - m)$ przejścia ze stanów nieistotnych do stanów nieistotnych. Zauważmy, że dla takich łańcuchów wygodniej jest zdefiniować postać kanoniczną inaczej niż poprzednio — zaczynamy numerowanie od stanów pochłaniających, dopiero później bierzemy pod uwagę stany nieistotne.

Z własności stanów nieistotnych i stanów pochłaniających wynika, że

1. $Q^t \rightarrow 0$ dla $t \rightarrow \infty$ (zbieżność po wyrazach);
2. macierz $\mathbb{I} - Q$ jest odwracalna;
3. $(\mathbb{I} - Q)^{-1} = \sum_{s=0}^{\infty} Q^s$.

Własność 1. wynika z ogólnego twierdzenia dotyczącego łańcuchów Markowa, które orzeka, że niezależnie od stanu początkowego prawdopodobieństwo trafienia do stanu komunikującego się po t krokach dąży do 1 przy $t \rightarrow \infty$, a ponieważ Q odpowiada stanom nieistotnym, zatem każde z pozostałych prawdopodobieństw dąży do 0. Dalej zauważmy, że

$$\mathbb{I} - Q^t = (\mathbb{I} - Q)(\mathbb{I} + Q + Q^2 + \dots + Q^{t-1}).$$

Skoro $Q^t \rightarrow 0$, to $\mathbb{I} - Q^t \rightarrow \mathbb{I}$, a z ciągłości wyznacznika wynika, że dla dostatecznie dużych t zachodzi $\det(\mathbb{I} - Q^t) \neq 0$, czyli także $\det(\mathbb{I} - Q) \neq 0$, więc macierz $\mathbb{I} - Q$ jest odwracalna. Stąd

$$(\mathbb{I} - Q)^{-1}(\mathbb{I} - Q^t) = \mathbb{I} + Q + Q^2 + \dots + Q^{t-1}$$

i przechodząc do granicy $t \rightarrow \infty$ dostajemy wzór 3.

W przypadku łańcuchów pochłaniających interesują nas głównie następujące zagadnienia dotyczące łańcucha, dla którego stanem początkowym jest pewien nieistotny stan \mathcal{W}_i .

1. Jaka jest oczekiwana liczba przejść przez stan \mathcal{W}_j , zakładając że stanem początkowym jest \mathcal{W}_i ?
2. Jaka jest oczekiwana liczba kroków przed absorpcją, jeśli stanem początkowym jest \mathcal{W}_i ?
3. Jakie jest prawdopodobieństwo absorpcji przez dany stan pochłaniający \mathcal{W}_j , jeśli stanem początkowym jest \mathcal{W}_i ?

Dla macierzy przejścia w postaci kanonicznej definiujemy następującą macierz fundamentalną

$$\mathbf{N} = (\mathbb{I} - Q)^{-1} = (\eta_{ij})_{i,j=1}^{n-m}$$

łańcucha pochłaniającego. Macierz ta zawiera wszystkie istotne informacje dotyczące zachowań asymptotycznych. W szczególności

Twierdzenie 10.1. *Niech \mathbf{N} będzie macierzą fundamentalną łańcucha pochłaniającego.*

1. *Oczekiwana liczba przejść przez stan \mathcal{W}_j przy stanie początkowym \mathcal{W}_i jest równa η_{ij} .*
2. *Oczekiwana liczba kroków przed absorpcją dla łańcucha o stanie początkowym \mathcal{W}_i zadana jest jako suma wyrazów w i . wierszu macierzy \mathbf{N} .*
3. *Niech $\mathbf{B} = (b_{ij})$ będzie $(n - m) \times m$ macierzą prawdopodobieństw absorpcji przez stan \mathcal{W}_j przy stanie początkowym \mathcal{W}_i . Wtedy $\mathbf{B} = \mathbf{N}R$.*

Dowód. Niech e_{ij} oznacza oczekiwaną liczbę przejść przez stan \mathcal{W}_j przy stanie początkowym \mathcal{W}_i . Wprowadźmy zmienną losową

$$\zeta_j^{(t)} = \begin{cases} 1 & \text{gdy } X_t = j, \\ 0 & \text{wpp.} \end{cases}$$

Niech $E_i(x)$ oznacza wartość średnią x przy warunku, że proces zaczął się w stanie \mathcal{W}_i . Wtedy

$$e_{ij} = E_i \left(\sum_{s=0}^{\infty} \zeta_j^{(s)} \right) = \sum_{s=0}^{\infty} E_i \left(\zeta_j^{(s)} \right) = \sum_{s=0}^{\infty} \left((1 - p_{ij}^{(s)}) \cdot 0 + p_{ij}^{(s)} \cdot 1 \right),$$

czyli ostatecznie $e_{ij} = \sum_{s=0}^{\infty} p_{ij}^{(s)}$. Skoro $p_{ij}^{(s)}$ jest (i, j) . wyrazem macierzy Q^s , to e_{ij} jest (i, j) . wyrazem macierzy \mathbf{N} , czyli η_{ij} .

Bezpośrednio z tego wzoru otrzymujemy także, że średni czas do absorpcji jest sumą wyrazów w wierszu i .

Wyprowadzimy teraz wzór rekurencyjny na prawdopodobieństwo b_{ij} . Niech \mathcal{W}_i będzie stanem początkowym nieistotnym. W pierwszym kroku łańcuch może przejść do interesującego nas stanu pochłaniającego \mathcal{W}_j , do innego stanu pochłaniającego \mathcal{W}_k , $k \neq j$, albo do któregoś ze stanów nieistotnych \mathcal{W}_l z prawdopodobieństwami zadanymi przez macierz \mathbf{P} . Prawdopodobieństwa absorpcji przez stan \mathcal{W}_j ze stanu \mathcal{W}_j , \mathcal{W}_k oraz \mathcal{W}_l są odpowiednio równe 1, 0 oraz b_{lj} . Stąd

$$b_{ij} = p_{ij} + \sum_{l=m+1}^n p_{il} b_{lj}.$$

Skoro \mathcal{W}_i oraz \mathcal{W}_j są odpowiednio stanem nieistotnym i pochłaniającym, więc p_{ij} jest (i, j) . wyrazem macierzy R i analogicznie p_{il} jest (i, l) . wyrazem macierzy Q . Zatem

$$\mathbf{B} = R + Q\mathbf{B} \implies \mathbf{B} = (\mathbb{I} - Q)^{-1}R = \mathbf{N}R.$$

□

Zastosujemy teraz powyższe twierdzenie do opisu asymptotyki rosyjskiej ruletki oraz do opisu doświadczenia polegającego na ciągłym krzyżowaniu z dominantą.

W przypadku rosyjskiej ruletki mamy jeden stan pochłaniający \mathcal{M} i jeden stan nieistotny \mathcal{Z} . Odpowiednie podmacierze macierzy \mathbf{P} są jednoelementowe

$$Q = (5/6), \quad R = (1/6), \quad \mathbf{N} = (1 - 5/6)^{-1} = (6), \quad \mathbf{B} = \mathbf{NR} = 1.$$

Ponieważ jest tylko jeden stan pochłaniający, więc oczywiście prawdopodobieństwo znalezienia się w tym stanie po dostatecznie długim czasie wynosi 1, natomiast oczekiwana liczba kroków do absorpcji równa się 6.

Ciągle krzyżowanie z dominantą Rozpatrzmy teraz następujący ciąg doświadczeń tworzący łańcuch Markowa. Bierzemy ustalonego osobnika o genotypie dominującym DD i krzyżujemy go z nieznanym osobnikiem. W wyniku eksperymentu dostajemy potomka o genotypie zależnym od genotypu drugiego rodzica z prawdopodobieństwami wynikającymi z prawa Mendla. Mamy zatem 3 możliwe wyniki eksperymentu, ponieważ są 3 genotypy. Ponumerujmy je w następujący sposób:

$$DD = \mathcal{W}_1, \quad DR = \mathcal{W}_2, \quad RR = \mathcal{W}_3.$$

Odpowiednie prawdopodobieństwa wynoszą

- $p_{11} = 1$, gdyż zawsze ze skrzyżowania dominandy z dominantą otrzymujemy ten sam genotyp. Stąd $p_{1j} = 0$ dla $j = 1, 2$ i stan \mathcal{W}_1 jest pochłaniający.
- $p_{21} = 1/2 = p_{22}$ oraz $p_{23} = 0$, gdyż krzyżując dominantę z hybrydą otrzymujemy z prawdopodobieństwem $1/2$ albo dominantę, albo hybrydę, nie możemy natomiast dostać osobnika recesywnego;
- $p_{31} = 0 = p_{33}$, $p_{32} = 1$, ponieważ krzyżując osobnika recesywnego z dominantą zawsze dostajemy hybrydę.

Dla tego łańcucha Markowa macierz przejścia ma postać

$$\mathbf{P}_D = \begin{pmatrix} 1 & 0 & 0 \\ 0,5 & 0,5 & 0 \\ 0 & 1 & 0 \end{pmatrix},$$

dla której

$$R = \begin{pmatrix} 0,5 \\ 0 \end{pmatrix} \quad \text{oraz} \quad Q = \begin{pmatrix} 0,5 & 0 \\ 1 & 0 \end{pmatrix}.$$

Policzmy macierz fundamentalną

$$\mathbf{N} = (\mathbb{I} - Q)^{-1} = \begin{pmatrix} 0,5 & 0 \\ -1 & 1 \end{pmatrix}^{-1} = 2 \begin{pmatrix} 1 & 0 \\ 1 & 0,5 \end{pmatrix}.$$

Na jej podstawie wnioskujemy, że jeśli wybranym rodzicem była hybryda, to średni czas do absorpcji wynosi 2, natomiast jeśli osobnik recesywny, to $2 + 1 = 3$. Możemy jeszcze sprawdzić punkt 3. twierdzenia wykonując mnożenie

$$\mathbf{B} = \mathbf{NR} = 2 \begin{pmatrix} 1 & 0 \\ 1 & 0,5 \end{pmatrix} \begin{pmatrix} 0,5 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

co oczywiście potwierdza, że z prawdopodobieństwem 1 nasz eksperyment kończy się osobnikami o genotypie DD .

Zauważmy jeszcze, że dokładnie w taki sam sposób możemy opisać ciągłe krzyżowanie z osobnikiem recesywnym. Wtedy macierz przejścia ma postać

$$\mathbf{P}_R = \begin{pmatrix} 1 & 0 & 0 \\ 0,5 & 0,5 & 0 \\ 0 & 1 & 0 \end{pmatrix},$$

gdzie tym razem $\mathcal{W}_1 = RR$, $\mathcal{W}_2 = DR$ oraz $\mathcal{W}_3 = DD$, zatem odpowiednio numerując stany dostajemy dokładnie taką samą dynamikę łańcucha pochłaniającego jak w przypadku krzyżowania z dominantą.

10.2. Łańcuchy regularne i ciągłe krzyżowanie z hybrydą

Zajmiemy się teraz pewnym szczególnym przypadkiem łańcuchów nieprzywiedlnych.

Definicja 10.1. Łańcuch nieprzywiedlny nazwiemy **regularnym**, jeśli istnieje taka liczba $k \in \mathbb{N} \setminus \{0\}$, że z każdego stanu \mathcal{W}_i , $i \in \{1, \dots, n\}$ można dojść do dowolnego stanu \mathcal{W}_j , $j \in \{1, \dots, n\}$ w dokładnie k krokach.

Zauważmy, że rzut monetą jest łańcuchem regularnym, dla którego $k = 1$. Natomiast łatwo podać przykład łańcucha nieprzywiedlnego, który nie jest regularny. Weźmy dla przykładu łańcuch o dwóch stanach $\mathcal{W}_1, \mathcal{W}_2$, dla którego macierz przejścia ma postać

$$\mathbf{P} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \implies \mathbf{P}^2 = \mathbb{I},$$

zatem np. ze stanu \mathcal{W}_1 do \mathcal{W}_2 przechodzimy zawsze w nieparzystej liczbie kroków, podczas gdy ze stanu \mathcal{W}_1 do niego samego — w parzystej. Tak zdefiniowany łańcuch jest oczywiście łańcuchem okresowym. Co więcej, można udowodnić następujące twierdzenie

Twierdzenie 10.2. *W nieprzywiedlnym łańcuchu Markowa wszystkie stany są tego samego typu, tzn. jeśli choć jeden stan jest powracający, to wszystkie są powracające, jeśli choć jeden jest zerowy, to wszystkie są zerowe, a jeśli choć jeden jest okresowy o okresie d , to wszystkie są okresowe o okresie d .*

Dowód. Niech $\mathcal{W}_i, \mathcal{W}_j$ będą dwoma różnymi stanami. Ponieważ stany te komunikują się, więc istnieją takie $M, N \in \mathbb{N} \setminus \{0\}$, że

$$p_{ij}^{(M)} > 0 \quad \text{oraz} \quad p_{ji}^{(N)} > 0.$$

Stosując wzór na prawdopodobieństwo całkowite

$$p_{ii}^{(N+M+t)} = \sum_{l,s=1}^n p_{il}^{(M)} p_{ls}^{(t)} p_{si}^{(N)}$$

dostajemy nierówność

$$p_{ii}^{(N+M+t)} \geq p_{ij}^{(M)} p_{jj}^{(t)} p_{ji}^{(N)} = \alpha \beta p_{jj}^{(t)},$$

gdzie $\alpha = p_{ij}^{(M)} > 0$, $\beta = p_{ji}^{(N)} > 0$ oraz $t \in \mathbb{N} \setminus \{0\}$. Postępując analogicznie otrzymujemy nierówność

$$p_{jj}^{(N+M+t)} \geq \alpha \beta p_{ii}^{(t)}.$$

Stąd

$$\frac{1}{\alpha \beta} p_{ii}^{(N+M+t)} \geq p_{jj}^{(t)} \geq \alpha \beta p_{ii}^{(t-M-N)}.$$

Powyższe nierówności wykazują, że własności asymptotyczne ciągów $(p_{ii}^{(t)})_{t \in \mathbb{N}}$ oraz $(p_{jj}^{(t)})_{t \in \mathbb{N}}$ są takie same. W szczególności, jeśli stan \mathcal{W}_i jest zerowy, $p_{ii}^{(t)} \rightarrow 0$ przy $t \rightarrow \infty$, to także \mathcal{W}_j jest zerowy. Z kolei jeśli \mathcal{W}_i jest powracający, czyli $P_i = \sum_{t=1}^{\infty} p_{ii}^{(t)} = \infty$, to \mathcal{W}_j ma tę samą własność.

Załóżmy teraz, że \mathcal{W}_i jest okresowy o okresie d_i . Ponieważ $p_{ii}^{(M+N)} \geq \alpha \beta > 0$, więc $d_i/(M+N)$ (d_i dzieli $M+N$). Pokażemy, że \mathcal{W}_j jest okresowy i jego okres $d_j = d_i$. Zauważmy, że jeśli dla pewnego l zachodzi $p_{jj}^{(l)} > 0$, to także $p_{ii}^{(l+M+N)} > 0$, zatem $d_i/(l+M+N)$, czyli d_i/l , więc $d_i \leq d_j$. Analogicznie $d_j \leq d_i$, skąd $d_i = d_j$. \square

W ogólnym przypadku dla łańcuchów nieprzywiedlnych asymptotykę opisuje następujące twierdzenie ergodyczne, którego dowód pomijamy.

Twierdzenie 10.3. *Łańcuch Markowa jest nieprzywiedlny wtw jeśli dla każdego $j \in \{1, \dots, n\}$ istnieje niezależna od i granica*

$$\lim_{t \rightarrow \infty} p_{ij}^{(t)} = p_j, \quad i, j \in \{1, \dots, n\},$$

przy czym p_j są jednoznacznym rozwiązaniem układu

$$p_j = \sum_{i=1}^n p_i p_{ij}, \quad j = 1, \dots, n, \quad \sum_{j=1}^n p_j = 1.$$

Taki asymptotyczny rozkład prawdopodobieństwa nazywamy rozkładem stacjonarnym. Wynika stąd, że zachowanie asymptotyczne badanego łańcucha nie zależy od stanu początkowego — łańcuch ma własność „zapominania” rozkładu początkowego.

Dla łańcuchów regularnych można udowodnić twierdzenie silniejsze.

Twierdzenie 10.4. *Macierz \mathbf{P} łańcucha regularnego ma pojedynczą wartość własną równą 1. Macierz \mathbf{P}^t zbiega do macierzy stochastycznej \mathbf{W} o wyrazach dodatnich, której wszystkie wiersze w są jednakowe. Zachodzi $\mathbf{wP} = \mathbf{w}$.*

Dowód. Ponieważ dla łańcucha regularnego wszystkie wyrazy macierzy \mathbf{P}^k są dodatnie, więc twierdzenie to jest bezpośrednim wnioskiem z [twierdzenia Frobeniusa – Perrona](#) dla macierzy o wszystkich wyrazach dodatnich. \square

Niech \mathbf{P} będzie macierzą przejścia łańcucha regularnego, natomiast W jej macierzą graniczną. Można pokazać, że macierz

$$\mathbb{I} - (\mathbf{P} - W)$$

jest odwracalna, co więcej

$$\mathbf{Z} = (\mathbb{I} - (\mathbf{P} - W))^{-1} = \mathbb{I} + \sum_{s=1}^{\infty} \mathbf{P}^s - W.$$

Macierz \mathbf{Z} nazywamy macierzą fundamentalną łańcucha regularnego i dzięki niej możemy wyznaczyć czasy przejścia i powrotów opisane macierzą $\mathbf{E} = (e_{ij})_{i,j=1}^n$. Mamy

$$\mathbf{E} = (\mathbb{I} - \mathbf{Z} + J\mathbf{Z}_p)D,$$

gdzie J jest macierzą o wszystkich wyrazach równych 1, \mathbf{Z}_p to macierz, która ma na przekątnej wyrazy takie same jak \mathbf{Z} i zera poza przekątną, natomiast D jest macierzą diagonalną o wyrazach na przekątnej równych $1/w_i$, gdzie $w = (w_1, \dots, w_n)$ oznacza rozkład stacjonarny. W szczególności wyrazy na przekątnej e_{jj} zadają średni pierwszy czas powrotu do stanu \mathcal{W}_j .

Zastosujmy najpierw powyższe twierdzenia do rzutu monetą. W tym przypadku rozkład stacjonarny znajdujemy bez trudu, gdyż $\mathbf{P}_{rm}^t = \mathbf{P}_{rm}$ dla dowolnego t , zatem $w_1 = w_2 = 0,5$. Policzmy jeszcze macierz \mathbf{E} , choć możemy się spodziewać rezultatu — średnio co drugi rzut powinien dawać \mathcal{O} i co drugi \mathcal{R} , zatem średni czas powrotu szacujemy jako 2.

$$\mathbf{Z} = (\mathbb{I} - (\mathbf{P}_{rm} - W))^{-1} = \mathbb{I} \implies \mathbf{E} = J\mathbf{Z}_p D = D,$$

co oczywiście implikuje, że średni czas powrotu wynosi 2 rzuty, natomiast średni czas przejścia od \mathcal{O} do \mathcal{R} i odwrotnie wynosi 1.

Ciągłe krzyżowanie z hybrydą W tym paragrafie omówimy doświadczenie genetyczne polegające na ciągłym krzyżowaniu z hybrydą. Przebieg doświadczenia jest analogiczny jak w omówionym już przypadku krzyżowania z dominantą, ale matematyczny opis doświadczenia prowadzi do innego typu łańcucha Markowa. W tym łańcuchu mamy także 3 możliwe stany $\mathcal{W}_1 = DD$, $\mathcal{W}_2 = DR$ i $\mathcal{W}_3 = RR$. Macierz przejścia tego łańcucha ma postać

$$\mathbf{P} = \begin{pmatrix} 0,5 & 0,5 & 0 \\ 0,25 & 0,5 & 0,25 \\ 0 & 0,5 & 0,5 \end{pmatrix}$$

i łatwo sprawdzimy, że \mathbf{P}^2 jest macierzą o wszystkich wyrazach dodatnich. Zatem mamy do czynienia z łańcuchem regularnym. Policzmy jego rozkład stacjonarny

$$(w_1, w_2, w_3)\mathbf{P} = (w_1, w_2, w_3), \quad w_1 + w_2 + w_3 = 1,$$

skąd dostajemy

$$w = (0,25, 0,5, 0,25).$$

Wobec tego przy ciągłym krzyżowaniu z hybrydą dostaniemy następujący wynik eksperymentu: bez względu na genotyp drugiego praprzdka po dostatecznie długim czasie średnio 3/4 potomków będzie wykazywało fenotyp dominany, a 1/4 fenotyp recesywny. Otrzymaliśmy zatem potwierdzenie pierwotnych wyników uzyskanych przez Mendla przy krzyżowaniu groszku, gdzie średnio było 3/4 strąków zielonych i 1/4 strąków żółtych.

11. Modele z zależnością przestrzenną: dyfuzja w procesach biologicznych

Do tej pory rozważaliśmy modele populacyjne, w których nie uwzględnialiśmy zależności od położenia osobników w przestrzeni, a tylko średnie zagęszczenia i ich zmiany w czasie. Takie podejście do modelowania dynamiki populacji kryje za sobą dodatkowe założenia dotyczące równomiernego rozłożenia przestrzennego osobników danej populacji oraz braku istotnego wpływu zmian położenia na opisywany proces. Oczywiście wydaje się, że w wielu przypadkach położenie i ruch w przestrzeni odgrywają dużą, jeśli nie zasadniczą rolę w przebiegu danego procesu. W związku z tym kolejną część wykładu poświęcimy omówieniu jednego z możliwych sposobów modelowania ruchu, tzw. procesu dyfuzji. W przypadku procesów chemicznych dyfuzja opisuje ruch substancji w kierunku od większych do mniejszych stężeń. Możliwe są oczywiście także przepływy innego typu — np. chemotaksja działa w pewnym sensie przeciwnie do dyfuzji.

Wyprowadzimy najpierw równanie dyfuzji substancji chemicznej w przestrzeni jednowymiarowej (dla przestrzeni dwu i trójwymiarowej rozumowanie przebiega analogicznie). Załóżmy, że mamy bardzo wąską rurkę o długości L i przekroju S , wzdłuż której przemieszcza się pewna substancja A . Rurka jest na tyle wąska, że ruch cząsteczek odbywa się tylko wzdłuż niej. Niech stężenie cząsteczek tej substancji w chwili t w miejscu x wynosi $A(t, x)$ i załóżmy, że A jest dostatecznie gładką funkcją obu zmiennych.

Z punktem x wiążemy wycinek rurki $(x, x + \Delta x)$ o objętości $\Delta V = S\Delta x$. Zmiana stężenia substancji A w tym wycinku następuje wyniku przepływu — mamy z jednej strony napływ do wycinka, z drugiej odpływ z niego. Zgodnie z prawem Ficka masa substancji przepływająca przez wąski przekrój w ciągu krótkiego przedziału czasu $(t, t + \Delta t)$ jest proporcjonalna do gradientu stężenia (w naszym jednowymiarowym przypadku do $\frac{\partial A}{\partial x}$), długości przedziału czasu i pola przekroju rurki, zatem

$$Q_x = DS \frac{\partial A(t, x)}{\partial x} \Delta t$$

w punkcie x oraz

$$Q_{x+\Delta x} = DS \frac{\partial A(t, x + \Delta x)}{\partial x} \Delta t$$

w punkcie $x + \Delta x$, gdzie $D > 0$ jest pewnym współczynnikiem proporcjonalności. Zatem zmiana masy w objętości ΔV wynosi $\Delta Q = Q_{x+\Delta x} - Q_x$, czyli zmiana stężenia substancji w wycinku wyraża się wzorem

$$\Delta A = \frac{\Delta Q}{\Delta V} = \frac{\left(\frac{\partial A(t, x + \Delta x)}{\partial x} - \frac{\partial A(t, x)}{\partial x} \right) DS \Delta t}{S \Delta x} = D \frac{\frac{\partial A(t, x + \Delta x)}{\partial x} - \frac{\partial A(t, x)}{\partial x}}{\Delta x} \Delta t.$$

Stąd, dzieląc stronami powyższą równość przez Δt i przechodząc do granicy $\Delta x \rightarrow 0$ oraz $\Delta t \rightarrow 0$ dostajemy

$$\frac{\partial A}{\partial t} = D \frac{\partial^2 A}{\partial x^2}, \quad (11.1)$$

czyli jednowymiarowe równanie dyfuzji. W ogólnym przypadku równanie dyfuzji ma postać

$$\frac{\partial A}{\partial t} = D \Delta A,$$

gdzie

$$\Delta A = \sum_{i=1}^n \frac{\partial^2 A}{\partial x_i^2},$$

n oznacza wymiar przestrzeni (w przypadku procesów obserwowanych w naturze mamy $n = 1, 2$ lub 3) i Δ nazywamy operatorem Laplace'a lub krócej laplasjanem, a $D > 0$ — współczynnikiem dyfuzji.

Z naszego populacyjnego punktu widzenia będziemy rozpatrywać dyfuzję jako proces losowego przemieszczania się osobników w przestrzeni. Najprostsze podejście do tego typu zagadnień reprezentowane jest przez tzw. ruchy Browna. Ponownie dla uproszczenia omówimy sytuację jednowymiarową. Rozważamy prostą podzieloną na odcinki o długości Δx . Niech osobnik znajduje się w chwili t w odcinku $[x, x + \Delta x)$ i w przedziale czasu o długości Δt może się przemieścić do odcinka na lewo, na prawo, albo pozostać w tym samym odcinku. Załóżmy, że prawdopodobieństwo przemieszczenia się do odcinka $[x - \Delta x, x)$ wynosi p_l , a do odcinka $[x + \Delta x, x + 2\Delta x)$ — p_p . Niech $a(t, x)$ będzie liczbą osobników w chwili t w odcinku $[x, x + \Delta x)$. Po upływie czasu Δt w tym samym odcinku mamy

$$a(t + \Delta t, x) = a(t, x) + p_p a(t, x - \Delta x) - p_p a(t, x) + p_l a(t, x + \Delta x) - p_l a(t, x), \quad (11.2)$$

gdzie $p_p a(t, x - \Delta x)$ oznacza średnią liczbę osobników, które w czasie t były na lewo od rozpatrywanego odcinka i przemieściły do niego, $p_l a(t, x + \Delta x)$ to osobniki, które były na prawo i przeszły na lewo, natomiast $p_p a(t, x)$, $p_l a(t, x)$ odpowiadają tym osobnikom, które były w omawianym odcinku w czasie t i w przedziale czasu $(t, t + \Delta t)$ przemieściły się odpowiednio na prawo lub lewo. Załóżmy teraz, że funkcję a możemy rozwinąć w szereg Taylora w punkcie (t, x) zarówno jako funkcję zmiennej t z parametrem x , jak i zmiennej x z parametrem t . Otrzymujemy odpowiednio

$$\begin{aligned} a(t + \Delta t, x) &= a(t, x) + \frac{\partial a(t, x)}{\partial t} \Delta t + \frac{1}{2} \frac{\partial^2 a(t, x)}{\partial t^2} (\Delta t)^2 + \dots \\ a(t, x + \Delta x) &= a(t, x) + \frac{\partial a(t, x)}{\partial x} \Delta x + \frac{1}{2} \frac{\partial^2 a(t, x)}{\partial x^2} (\Delta x)^2 + \dots \end{aligned}$$

Wstawiając powyższe rozwinięcia do równania (11.2) dostajemy

$$\begin{aligned} \frac{\partial a(t, x)}{\partial t} \Delta t + \frac{1}{2} \frac{\partial^2 a(t, x)}{\partial t^2} (\Delta t)^2 &= p_p \left(-\frac{\partial a(t, x)}{\partial x} \Delta x + \frac{1}{2} \frac{\partial^2 a(t, x)}{\partial x^2} (\Delta x)^2 \right) + \\ &+ p_l \left(\frac{\partial a(t, x)}{\partial x} \Delta x + \frac{1}{2} \frac{\partial^2 a(t, x)}{\partial x^2} (\Delta x)^2 \right) + \dots \end{aligned}$$

Założmy teraz, że $p_l = p_p = \frac{1}{2}$, czyli osobnik z jednakowym prawdopodobieństwem przemieszcza się zarówno w lewo, jak i w prawo. Mamy wtedy

$$\frac{\partial a(t, x)}{\partial t} \Delta t + \frac{1}{2} \frac{\partial^2 a(t, x)}{\partial t^2} (\Delta t)^2 = \frac{1}{2} \frac{\partial^2 a(t, x)}{\partial x^2} (\Delta x)^2 + \dots$$

Dzieląc to równanie przez Δt , przechodząc do granicy $\Delta t \rightarrow 0$ i $\Delta x \rightarrow 0$ w taki sposób, że $\frac{(\Delta x)^2}{2\Delta t} = D$ dostajemy równanie (11.1).

Otrzymaliśmy więc związek losowego przemieszczenia się osobników z równaniem dyfuzji z jednej strony i pewnym łańcuchem Markowa z drugiej.

11.1. Równania ewolucyjne: równanie Fishera – Kołmogorowa

W przypadku rzeczywistych populacji oprócz przemieszczania się rozpatrujemy zwykle także inne procesy, np. rozrodczości/śmiertelności. W ogólnym przypadku otrzymujemy równanie reakcji-dyfuzji

$$\frac{\partial N(t, x)}{\partial t} = f(t, N(t, x)) + \Delta N(t, x), \quad (11.3)$$

zwane także równaniem ewolucyjnym, które opisuje zmiany zagęszczenia populacji w chwili t w miejscu o położeniu x w przestrzeni. Czasami także w odniesieniu do funkcji f używa się nazwy kinetyka oddziaływań (albo reakcji w kontekście biochemicznym), a odpowiadające równaniu (11.3) równanie zwyczajne $\dot{N} = f(t, N)$ zwane jest wtedy równaniem kinetycznym.

Równanie Skellama Najprostsze tego typu równanie z liniową kinetyką zostało użyte przez Skellama w 1951 roku do opisu rozprzestrzeniania się populacji. Dokładniej, za pomocą równania (11.3) z funkcją $f(t, N) = \alpha N$ opisał on rozprzestrzenianie się populacji piżmaków w Europie. Gatunek ten został sprowadzony do Europy w 1905 roku przez węgierskiego posiadacza ziemskiego i szybko rozprzestrzenił się wokół. Mimo że obszar Europy jest ograniczony, Skellam rozpatrywał to równanie w całym \mathbb{R}^2 z warunkiem początkowym, który odpowiada introdukcji kilku osobników w punkcie $(0, 0)$, czyli $N(0, x, y) = M\delta(x, y)$, gdzie M to początkowa liczba osobników, a $\delta(x, y) = \delta(x)\delta(y)$, gdzie $\delta(x)$ oznacza deltę-Diraca, czyli $\delta(x) = 0$ dla $x \neq 0$ oraz $\int_{-\infty}^{\infty} \delta(x)dx = 1$. Aby prześledzić rozprzestrzenianie się osobników zgodnie z modelem Skellama zbadajmy najpierw zagadnienie prostsze.

Rozpatrzmy równanie dyfuzji (11.1) w całym \mathbb{R} z warunkiem początkowym $A(0, x) = \delta(x)$, gdzie δ oznacza jednowymiarową deltę-Diraca. Można pokazać, że rozwiązaniem tego zagadnienia początkowego jest funkcja

$$A(t, x) = \frac{1}{\sqrt{4\pi Dt}} \exp\left(-\frac{x^2}{4Dt}\right),$$

którą nazywamy rozwiązaniem fundamentalnym. Oczywiście spełnienie warunku początkowego należy rozumieć w sensie granicznym. Ogólniej, jeśli rozważamy dyfuzję w przestrzeni n -wymiarowej z warunkiem początkowym $\delta(\mathbf{x}) = \Pi\delta(x_i)$, $\mathbf{x} = (x_1, x_2, \dots, x_n)$, to rozwiązanie fundamentalne ma postać

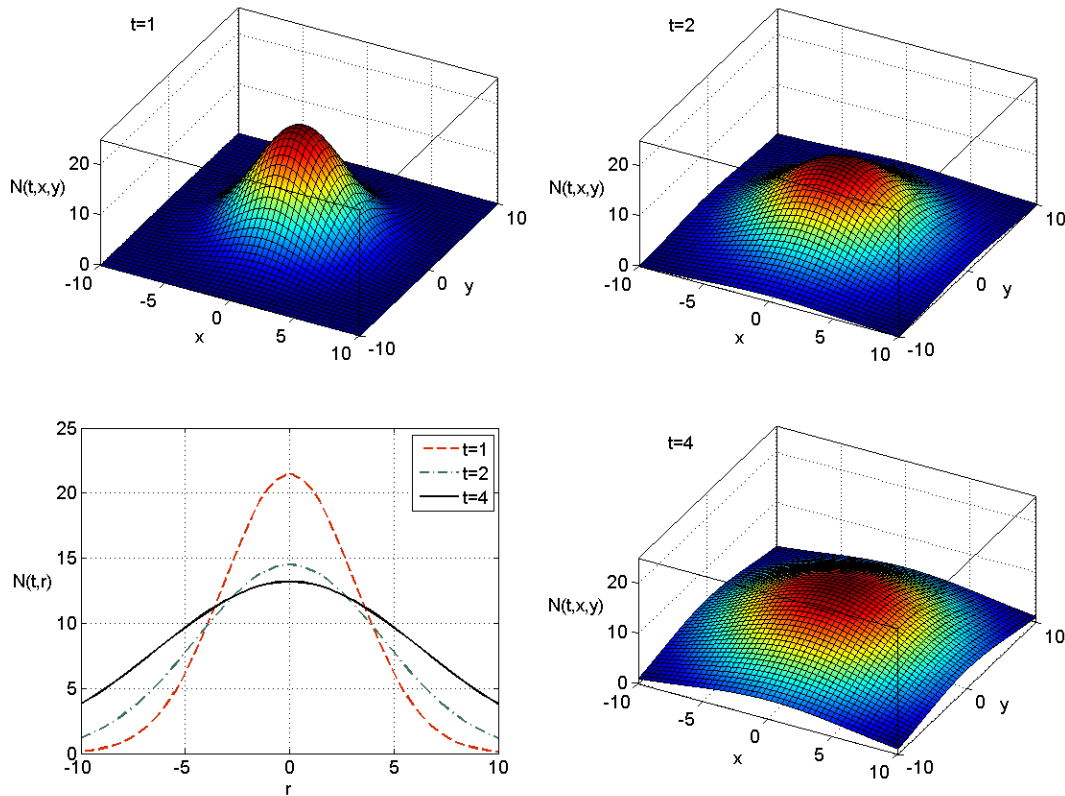
$$A(t, \mathbf{x}) = \frac{1}{(4\pi Dt)^{n/2}} \exp\left(-\frac{|\mathbf{x}|^2}{4Dt}\right).$$

Teraz połączmy znajomość rozwiązania fundamentalnego równania dyfuzji dla $n = 2$ i rozwiązania równania kinetycznego $\dot{N} = \alpha N$ w celu otrzymania rozwiązania równania Skellama

$$N(t, r) = \frac{M}{4\pi Dt} \exp\left(\alpha t - \frac{r^2}{4Dt}\right), \quad r = \sqrt{x^2 + y^2},$$

które zadaje falę inwazyjną. W punktach o ustalonym promieniu r w pewnej chwili t pojawiają się osobniki, których tam wcześniej w zasadzie nie było (należy to interpretować w taki sposób, że fizycznie tych osobników nie było, mamy więc do czynienia z przełączeniem, przy czym funkcja A w sposób ciągły przybliża taką dynamikę), a choć ich liczebność wraz z upływem czasu maleje, ponieważ stale rozprzestrzeniają się dalej, to populacja jako taka w tym miejscu egzystuje, por rys. 11.1.

Równanie Fishera – Kołmogorowa Równanie Fishera [3] lub inaczej Fishera – Kołmogorowa jest z kolei najprostszym nieliniowym równaniem typu ewolucyjnego i stanowi naturalne



Rysunek 11.1. Przykładowe rozwiązanie równania Skellama.

uogólnienie równania logistycznego na przypadek populacji złożonej z osobników, które mogą się przemieszczać i ruch ten ma znaczenie dla opisywanego procesu. Jest to także prototypowe równanie, w którym występują rozwiązania w postaci fal biegnących. Typowo jest to równanie rozpatrywane w całej przestrzeni \mathbb{R}^n , gdzie n zwykle odpowiada naturalnym wymiarom 1, 2 lub 3. Zostało zaproponowane przez Fishera w 1937 roku do opisu rozprzestrzeniania się genu dominującego w populacji.

Równanie to ma postać

$$\frac{\partial N(t, x)}{\partial t} = D\Delta N(t, x) + rN(t, x)(1 - N(t, x)), \quad (11.4)$$

przy czym w najprostszym przypadku, który omówimy, $x \in \mathbb{R}$, $\Delta N = \frac{\partial^2 N}{\partial x^2}$, a $D > 0$ jest współczynnikiem dyfuzji.

Pokażemy, że równanie (11.4) ma rozwiązania w postaci fal biegnących. Zaczniemy od przeprowadzenia ubezwymiarowania w celu zredukowania liczby parametrów. Przeprowadzając zamianę zmiennych

$$\hat{t} = rt, \quad \hat{x} = x \left(\frac{r}{D} \right)^{1/2}$$

i wracając do standardowych oznaczeń zmiennych dostajemy

$$\frac{\partial N(t, x)}{\partial t} = \Delta N(t, x) + N(t, x)(1 - N(t, x)). \quad (11.5)$$

Odpowiadające równaniu (11.5) równanie kinetyczne, czyli równanie logistyczne, ma dwa stany stacjonarne, niestabilny $N \equiv 0$ i stabilny $N \equiv 1$. Będziemy więc poszukiwać fali biegnącej od $N = 0$ do $N = 1$.

Definicja 11.1. Falą biegnącą nazywamy nietrywialne, ograniczone rozwiązanie równania (11.3) postaci

$$N(t, x) = U(x - ct) = U(z), \quad z = x - ct,$$

dla pewnego $c > 0$, które nazywamy prędkością fali.

Zauważmy, że

Uwaga 11.1. Równanie (11.1) nie ma rozwiązań w postaci fali biegnącej.

Dowód. Załóżmy, że takie rozwiązanie $U(z)$ istnieje. Skoro tak, to spełnia równanie różniczkowe

$$D \frac{d^2 U}{dz^2} + c \frac{dU}{dz} = 0,$$

zatem wartości własne tego równania spełniają równanie kwadratowe $D\lambda^2 + c\lambda = 0$, czyli rozwiązania mają postać $U(z) = A + B \exp(-\frac{c}{D}z)$, gdzie A, B są stałymi. Ponieważ fala biegnąca jest ograniczona dla wszystkich $z \in \mathbb{R}$, więc $B = 0$, czyli $U(z) = A$, zatem nie jest to rozwiązanie w postaci fali. \square

Dla równania Fishera – Kołmogorowa (11.5) rozwiązanie w postaci fali biegnącej musi spełniać równanie

$$\frac{d^2 U}{dz^2} + c \frac{dU}{dz} + U(1 - U) = 0 \quad (11.6)$$

i postulujemy, że $\lim_{z \rightarrow -\infty} U(z) = 0$, $\lim_{z \rightarrow +\infty} U(z) = 1$, zatem fala łączy dwa stany stacjonarne — jest odpychana od niestabilnego i przyciągana przez stabilny.

Równanie (11.6) jest równoważne układowi równań

$$\begin{aligned} U' &= V, \\ V' &= -cV - U(1 - U), \end{aligned} \quad (11.7)$$

gdzie prim ($'$) oznacza różniczkowanie po z . W przestrzeni fazowej (U, V) układ (11.7) ma oczywiście dwa punkty krytyczne $(0, 0)$ oraz $(1, 0)$. Macierz Jacobiego tego układu ma postać

$$\text{MJ}(U, V) = \begin{pmatrix} 0 & 1 \\ -1 + 2U & -c \end{pmatrix}.$$

Badając lokalną stabilność wyznaczamy wielomian charakterystyczny

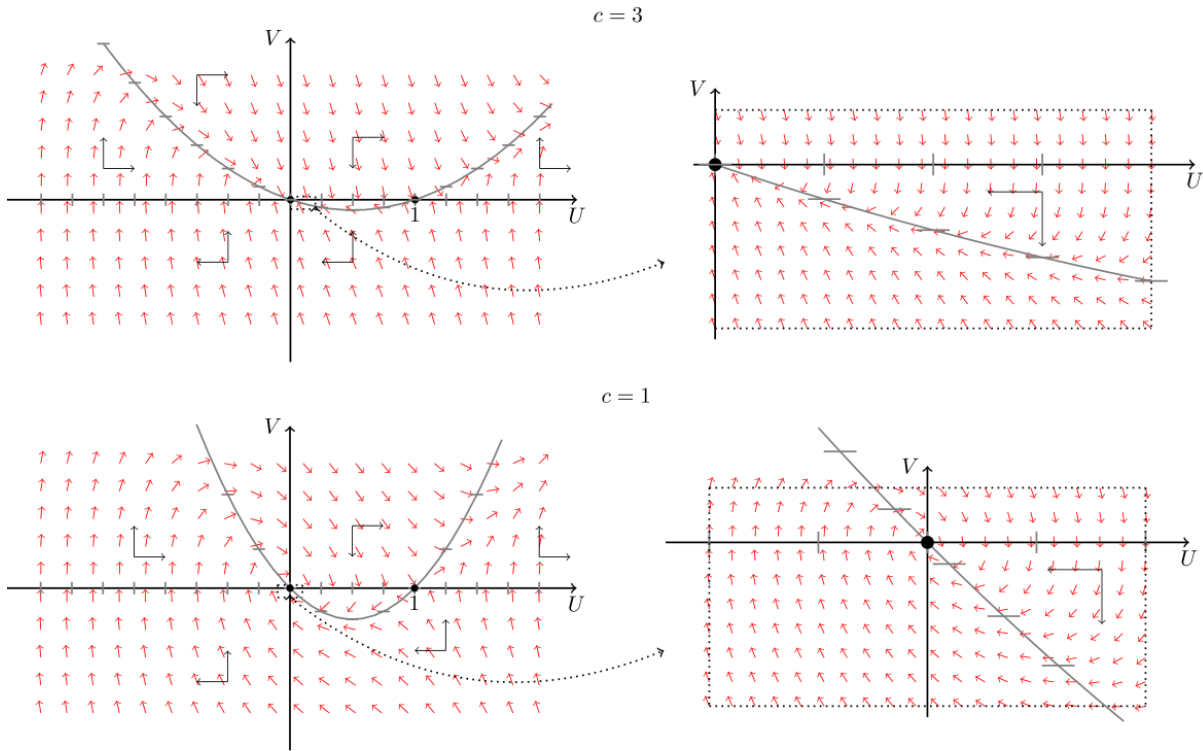
$$\lambda^2 + c\lambda + 1 - 2U = 0$$

i wartości własne

$$\begin{aligned} \lambda_{1,2} &= \frac{1}{2} \left(-c \pm \sqrt{c^2 - 4} \right) \quad \text{dla } (0, 0), \\ \lambda_{1,2} &= \frac{1}{2} \left(-c \pm \sqrt{c^2 + 4} \right) \quad \text{dla } (1, 0), \end{aligned}$$

skąd wynika, że $(1, 0)$ jest zawsze punktem siodłowym, natomiast $(0, 0)$ jest stabilnym ogniskiem dla $c^2 < 4$ i węzłem stabilnym dla $c^2 > 4$.

Stwierdzenie 11.1. Dla $c \geq 2$ równanie (11.5) ma rozwiązanie w postaci fali biegnącej o prędkości c .



Rysunek 11.2. Portret fazowy układu (11.7) dla $c = 3$ (na górze) oraz dla $c = 1$ (na dole).

Dowód. Dla układu (11.7) przy $c \geq 2$ z kierunku pola wektorowego łatwo wnioskujemy, że istnieje orbita łącząca punkt krytyczny $(0,0)$ z punktem krytycznym $(1,0)$, przy czym można pokazać, że pozostaje ona w dodatniej półpłaszczyźnie, por. rys. 11.2. Orbita ta tworzy rozwiązanie w postaci fali biegnącej łączącej punkty $(0,0)$ oraz $(1,0)$. \square

Uwaga 11.2. Dla $c < 2$ także istnieje orbita łącząca dwa punkty krytyczne $(0,0)$ i $(1,0)$, ale ze względu na typ punktu $(0,0)$ nie jest zachowana nieujemność rozwiązań, zatem takie rozwiązania są nieistotne z biologicznego punktu widzenia, por. rys. 11.2.

Zagadnienie występowania fal biegnących (bądź też fal innego typu) jest często badane w kontekście tego typu równań. Należy jednak zdawać sobie sprawę z tego, że nie zawsze rozważana dziedzina jest na tyle duża, że można założyć iż jest to cała przestrzeń. Wtedy, przy założeniu ograniczonej dziedziny, rozważamy zagadnienia początkowo-brzegowe. Równanie (11.4) było także rozpatrywane w kontekście rozwoju nowotworu przy założeniu warunku brzegowego. Załóżmy dla uproszczenia jednowymiarową propagację komórek na odcinku $[0,1]$ z warunkiem początkowym

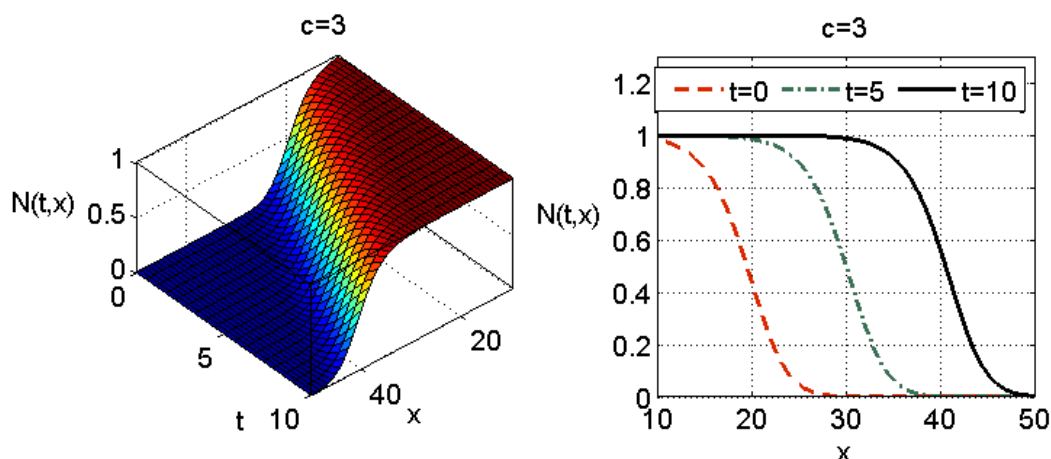
$$N(0, x) = N_0(x) \quad (11.8)$$

i warunkiem brzegowym Neumanna, czyli

$$\frac{\partial N}{\partial x} \Big|_{x=0,1} = 0 \quad (11.9)$$

Dla zagadnienia początkowo-brzegowego (11.4) (11.8) (11.9) lokalne istnienie i jednoznaczność wynika z własności kinetyki.

Korzystając z twierdzenia porównawczego 11.1 i własności rozwiązań równania logistycznego możemy wykazać, że dla warunku początkowego $N_0(x) \in (0,1)$, $x \in [0,1]$ rozwiązanie zagadnienia początkowo-brzegowego (11.4) (11.8) (11.9) pozostaje w przedziale $[0,1]$.



Rysunek 11.3. Przykładowe rozwiązanie równania (11.5) w postaci fali biegnącej o prędkości $c = 3$.

Twierdzenie 11.1. Niech $u(t, x)$, $v(t, x)$ będą ograniczonymi funkcjami ciągłymi spełniającymi nierówności

$$\begin{aligned} A &\leq u(t, x), \quad v(t, x) \leq B, \\ \frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} - f(u) &\geq \frac{\partial v}{\partial t} - \frac{\partial^2 v}{\partial x^2} - f(v) \quad w \quad [0, T] \times (a, b), \\ A &\leq v(0, x) \leq u(0, x) \leq B \quad w \quad (a, b), \end{aligned}$$

gdzie $-\infty \leq a < b \leq \infty$ i $0 < T \leq \infty$ oraz f jest funkcją klasy C^1 . Załóżmy także, że

— jeśli $a > -\infty$, to

$$A \leq v(t, a) \leq u(t, a) \leq B \quad w \quad [0, T];$$

— jeśli $b < \infty$, to

$$A \leq v(t, b) \leq u(t, b) \leq B \quad w \quad [0, T].$$

Wtedy albo $u \equiv v$ w $[0, T] \times [a, b]$, albo $u > v$ w $[0, T] \times (a, b)$.

Badając dynamikę modelu zwykle zaczynamy od zbadania stabilności rozwiązań stacjonarnych. Dla modeli niejednorodnych przestrzennie, oprócz rozważanych do tej pory rozwiązań stacjonarnych niezależnych od przestrzeni możemy mieć także rozwiązania stacjonarne, które zależą od położenia w dziedzinie równania. Jednak typowo analizę takiego zagadnienia zaczynamy od zbadania istnienia i stabilności rozwiązań stacjonarnych jednorodnych przestrzennie, a potem przechodzimy do znacznie trudniejszych zagadnień związanych z niejednorodnością przestrzenną. Jednym z takich zagadnień jest powstawanie tzw. wzorów Turinga — od strony analitycznej wiąże się to z bifurkacją stabilnego przestrzennie stanu stacjonarnego do niestabilnego, kiedy w otoczeniu stanu stacjonarnego powstają skupiska osobników — grupy o różnym zagęszczeniu — tworzące wzory przestrzenne.

Niech $\eta = N - \bar{N}$ będzie odchyleniem rozwiązania równania (11.4) od jednorodnego przestrzennie stanu stacjonarnego \bar{N} (gdzie $\bar{N} = 0$ lub $\bar{N} = 1$). Równanie zlinearyzowane wokół \bar{N}

ma postać

$$\eta_t = a\eta + D\eta_{xx}, \quad (11.10)$$

gdzie $a = r$ dla stanu $\bar{N} = 0$ i $a = -r$ dla stanu $\bar{N} = 1$, przy czym równanie (11.10) badamy z warunkiem brzegowym $\frac{\partial \eta}{\partial x}|_{x=0,1} = 0$ i początkowym $\eta(0, x) = \eta^0$. Z twierdzenia o linearyzacji wynika, że jeśli kinetyka jest funkcją różniczkowalną, to ze stabilności rozwiązania równania zlinearyzowanego wynika stabilność rozwiązania równania wyjściowego. Badając stabilność równania liniowego szukamy rozwiązań postaci $\eta(t, x) = \eta_0 \exp(\lambda t)W(x)$, gdzie $W''(x) + k^2W(x) = 0$, $k \in \mathbb{Z}$, jest rozwiązaniem fundamentalnym równania dyfuzji. Istnienie takiego rozwiązania wymaga spełnienia zależności $\lambda = a - Dk^2$. Zatem dla $a = -r$ mamy $\lambda = -r - Dk^2 < 0$, czyli $\bar{N} = 1$ jest rozwiązaniem stabilnym, gdyż wszystkie wartości własne są ujemne, zaś dla $\bar{N} = 0$ mamy $\lambda = r - Dk^2$ i przy $k = 0$ wartość własna $\lambda_0 = r > 0$, zatem to rozwiązanie jest niestabilne. Dokładnie tak samo kształtuje się stabilność w przypadku bez dyfuzji, zatem dla równania Fishera – Kolmogorowa formowanie się wzorów Turinga nie jest możliwe. Co więcej widać, że aby takie wzory powstały, model musi być opisany za pomocą układu co najmniej dwóch równań.

W przypadku równania (11.4) możemy udowodnić znacznie więcej, czyli globalną stabilność $\bar{N} = 1$ przy założeniu, że N jest oddzielone od zera.

Twierdzenie 11.2. *Jeśli $N_0(0) \geq d > 0$ dla pewnego $d \in (0, 1)$ oraz $t \geq 0$, to $N(t, x) \rightarrow 1$ przy $t \rightarrow \infty$.*

Dowód. Twierdzenie porównawcze pokazuje, że $N(t, x) \geq d$ dla $t \geq 0$ i $x \in [0, 1]$. Jeśli $\eta = N - 1$, to

$$\dot{\eta} = -r\eta(\eta - 1) + D\Delta\eta, \quad \frac{\partial \eta}{\partial x}|_{x=0,1} = 0 \quad (11.11)$$

oraz $\eta(t, x) \geq -(1 - d)$ dla $t \geq 0$ i $x \in [0, 1]$. Zdefiniujmy

$$L(t) = \frac{1}{2} \int_{[0,1]} \eta^2(t, x) dx.$$

Wtedy

$$\dot{L}(t) = \int_{[0,1]} \dot{\eta}(t, x)\eta(t, x) dx.$$

Mnożąc równanie (11.11) przez η i całkując przez części dostajemy

$$\dot{L}(t) = -r \int_{[0,1]} \eta^2(t, x) dx - r \int_{[0,1]} \eta^3(t, x) dx - D \int_{[0,1]} \left(\frac{\partial \eta(t, x)}{\partial x} \right)^2 dx.$$

Widzimy, że jeśli $N > 1$, to $\eta > 0$ oraz $\dot{L}(t) < -2rL$. Jeśli zaś $N < 1$, to $0 > \eta > -(1 - d)$ i $\eta^3 > -(1 - d)\eta^2$, skąd wynika $\dot{L}(t) < -2rdL$. W obu przypadkach $L(t) \rightarrow 0$ przy $t \rightarrow \infty$, skąd wynika $N(t, x) \rightarrow 1$. \square

11.2. Wzory Turinga

Jak już wspomnieliśmy, w przypadku jednego równania kinetycznego formowanie się wzorów Turinga nie jest możliwe, gdyż stabilność nie zmienia się pod wpływem dyfuzji. Natomiast

dla $n = 2$ istnieje taka możliwość. W szczególności można w taki sposób opisać tworzenie się wzorów umaszczenia u ssaków i próbować na tym gruncie wyjaśnić, dlaczego zebry czy żyrafy mają ustalony wzór umaszczenia, a u kotów domowych ten wzór jest zmienny. Podobnie — dlaczego wzór na tułowi różni się od wzoru na kończynach lub ogonie. Taki model nazywamy modelem melanogenezy, ponieważ zakłada się, że za umaszczenie odpowiada stężenie barwnika zwanego melaniną. Melanina powstaje w reakcji substratu (tyrozyny) z enzymem (tyrozinazą). Model może być zatem opisany za pomocą układu dwóch równań reakcji-dyfuzji. Ponieważ w przypadku tułowia w uproszczeniu możemy powierzchnię skóry rozpatrywać jako pewną figurę płaską, to taki układ rozważa się najczęściej w ograniczonym obszarze w \mathbb{R}^2 utożsamianym zwykle z prostokątem. Zadaje się zatem warunek początkowy oraz warunek brzegowy, który tu jest warunkiem Neumanna. Jeśli natomiast rozpatrujemy kończyny lub ogon, to rozpatrujemy taki układ np. na walcu lub stożku.

Badając możliwość występowania wzorów Turinga przeanalizujemy lokalną stabilność jednorodnych rozwiązań stacjonarnych w układzie dwóch równań reakcji-dyfuzji

$$\begin{aligned} s_t^1 &= f_1(s^1, s^2) + D_1 \Delta s^1, \\ s_t^2 &= f_2(s^1, s^2) + D_2 \Delta s^2, \end{aligned} \quad (11.12)$$

gdzie s^i , $i = 1, 2$, są stężeniami substancji chemicznych, f_i oznaczają templa produkcji tych reakcji, a D_i są ich współczynnikami dyfuzji. W przypadku melanogenezy funkcje produkcji mają postać

$$f_1(s^1, s^2) = s_0^1 - s^1 - \frac{\rho s^1 s^2}{1 + s^1 + k(s^1)^2}, \quad f_2(s^1, s^2) = \alpha(s_0^2 - s^2) - \frac{\rho s^1 s^2}{1 + s^1 + k(s^1)^2},$$

gdzie s^1 to stężenie tyrozyny, a s^2 — tyrozinazy. Zauważmy, że dla $k = 0$ nieliniowy składnik f_i ma postać funkcji Michaelisa – Menten, natomiast dla $k > 0$ wraz z rosnącą ilością substratu reakcja ulega zahamowaniu.

Dla uproszczenia układ (11.12) będziemy rozpatrywać z jednorodnym warunkiem brzegowym Neumanna na $[0, \pi]$ — analiza w przypadku kwadratu $[0, \pi] \times [0, \pi]$ przebiega podobnie. Także dla uproszczenia założymy, że $(0, 0)$ jest jednorodnym stanem stacjonarnym układu (11.12) (jeśli stanem stacjonarnym jest (\bar{s}^1, \bar{s}^2) , to dokonujemy zamiany zmiennych $s^i \rightarrow s^i - \bar{s}^i$).

Badając stabilność stanu stacjonarnego zlinearyzujemy układ (11.12)

$$\begin{aligned} s_t^1 &= \alpha_{11}s^1 + \alpha_{12}s^2 + D_1 \Delta s^1, \\ s_t^2 &= \alpha_{21}s^1 + \alpha_{22}s^2 + D_2 \Delta s^2, \end{aligned} \quad (11.13)$$

przy czym $\alpha_{ij} = \frac{\partial f_i}{\partial s^j}$, $i, j \in \{1, 2\}$. Wzory Turinga mogą pojawić się w sytuacji, gdy w modelu bez dyfuzji, czyli dla $D_1 = D_2 = 0$ stan stacjonarny jest stabilny, a dla pewnych $D_i > 0$ destabilizuje się.

Okazuje się, że liniowe równanie (11.13) można rozwiązać metodą rozdzielenia zmiennych, czyli $s^i(t, x) = F(t)G(x)$, gdzie $F(t)$ to rozwiązanie układu kinetycznego (dla $D_i = 0$), a $G(x)$ jest rozwiązaniem zagadnienia własnego operatora Laplace'a. W ogólnym przypadku na odcinku $[0, \pi]$ takie rozwiązanie jest kombinacją sinusów i kosinusów, jednak ponieważ badamy układ (11.13) z warunkiem brzegowym Neumanna, to spełnia go tylko kosinus. Stąd

$$s^i(t, x) = A_i e^{Bt} \cos(Cx), \quad i = 1, 2. \quad (11.14)$$

Podstawiając postać rozwiązania (11.14) do układu (11.13) otrzymujemy

$$\begin{aligned} A_1 B &= \alpha_{11} A_1 + \alpha_{12} A_2 - D_1 C^2 A_1, \\ A_2 B &= \alpha_{21} A_1 + \alpha_{22} A_2 - D_2 C^2 A_2, \end{aligned} \quad (11.15)$$

skąd dostajemy układ równań na współczynniki A_1, A_2

$$\begin{aligned} (B - \alpha_{11} + D_1 C^2) A_1 - \alpha_{12} A_2 &= 0, \\ -\alpha_{21} A_1 + (B - \alpha_{22} + D_2 C^2) A_2 &= 0. \end{aligned} \quad (11.16)$$

Rozwiązanie dostajemy przyrównując wyznacznik układu (11.16) do 0

$$B^2 + B(-\alpha_{11} - \alpha_{22} + C^2(D_1 + D_2)) + (\alpha_{11} - D_1 C^2)(\alpha_{22} - D_2 C^2) - \alpha_{12}\alpha_{21} = 0. \quad (11.17)$$

W przypadku bez dyfuzji równanie (11.17) ma postać

$$B^2 - B(\alpha_{11} + \alpha_{22}) + \alpha_{11}\alpha_{22} - \alpha_{12}\alpha_{21} = 0$$

i stabilność stanu stacjonarnego zależy od śladu i wyznacznika macierzy Jacobiego układu kinetycznego. Precyzyjniej, jeśli

1. $\alpha_{11} + \alpha_{22} < 0$,
2. $\alpha_{11}\alpha_{22} - \alpha_{12}\alpha_{21} > 0$,

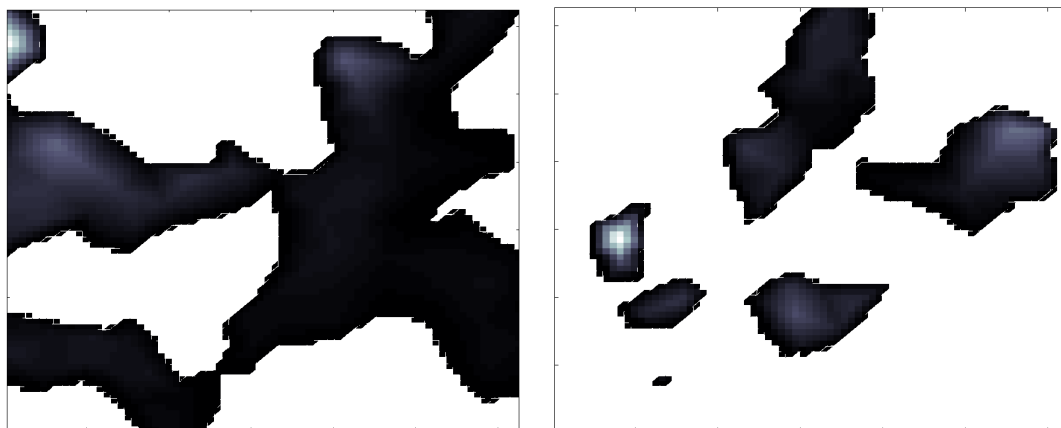
to stan stacjonarny dla $D_1 = D_2 = 0$ jest asymptotycznie stabilny. Jeśli teraz chcemy, by był on niestabilny dla $D_i > 0$, to któraś z powyższych nierówności powinna mieć przeciwny znak. Nierówność I w przypadku z dyfuzją przyjmuje postać $\alpha_{11} + \alpha_{22} - C^2(D_1 + D_2) < 0$, więc dla dowolnych $D_i > 0$ jest spełniona przy założeniu $\alpha_{11} + \alpha_{22} < 0$. Zatem tylko nierówność II może zmienić znak dla $D_i > 0$. Postulując

$$(\alpha_{11} - D_1 C^2)(\alpha_{22} - D_2 C^2) - \alpha_{12}\alpha_{21} < 0$$

dostajemy nierówność na C^2 , skąd otrzymujemy warunki destabilizacji stanu stacjonarnego pod wpływem dyfuzji

1. $\alpha_{11} + \alpha_{22} < 0$,
2. $\alpha_{11}\alpha_{22} - \alpha_{12}\alpha_{21} > 0$,
3. $D_2\alpha_{11} + D_1\alpha_{22} > 2\sqrt{D_1 D_2 (\alpha_{11}\alpha_{22} - \alpha_{12}\alpha_{21})}$.

W analogiczny sposób możemy badać tworzenie się wzorów przestrzennych w modelu melanoogenezy, przy czym wielkość i różnorodność formujących się wzorów zależy oczywiście od parametrów modelu. Jeśli dla danego gatunku dominują składniki kinetyczne, to powstające wzory są powtarzalne (jak np. u żyraf czy zebr), a jeśli składniki losowe, to wzory mogą się bardzo różnić w zależności od osobników (jak w przypadku kotów domowych).



Rysunek 11.4. Graficzne przedstawienie rozwiązań układu (11.12) dla różnych układów parametrów. Ciemne plamy odzwierciedlają stężenie substratu s^1 przewyższające wartość rozwiązania stacjonarnego, natomiast obszary z jasnymi plamami - stężenie poniżej tej wartości.

12. Modelowanie odpowiedzi odpornościowej

Dwa kolejne wykłady zostaną poświęcone modelowaniu [reakcji odpornościowych](#). Na początku przedstawimy podstawowe informacje dotyczące działania układu odpornościowego człowieka.

12.1. Układ odpornościowy człowieka

[Układ odpornościowy](#) człowieka (i nie tylko człowieka) jest bardzo skomplikowany. Odpowiedź odpornościową wywołuje [antygen](#). Antygenem nazywamy substancję, która wykazuje się:

1. [immunogennością](#), czyli zdolnością indukowania [swoistej](#) (specyficznej) odpowiedzi odpornościowej przeciw sobie;
2. [antygenowością](#), czyli zdolnością do swoistego łączenia się z immunoglobulinami (zarówno wolnymi jak i stanowiącymi receptory [limfocytów B](#)) oraz receptorami [limfocytów T](#).

Okazuje się, że w obrębie jednego antygeny może znajdować się wiele miejsc wiązanych przez przeciwciała. Są to tzw. [epitopy](#) lub determinanty antygenowe. Epitopy obecne w jednej części antygeny mogą być identyczne lub różne i mogą być wiązane przez [przeciwciała](#) o tej samej lub różnej swoistości. Antygen zawierający wiele epitopów nazywamy wielowartościowym lub poliwalentnym.

Reakcja odpornościowa składa się z wielu różnorodnych procesów, w których bierze udział wiele składników układu odpornościowego. Najważniejsze z tych składników to:

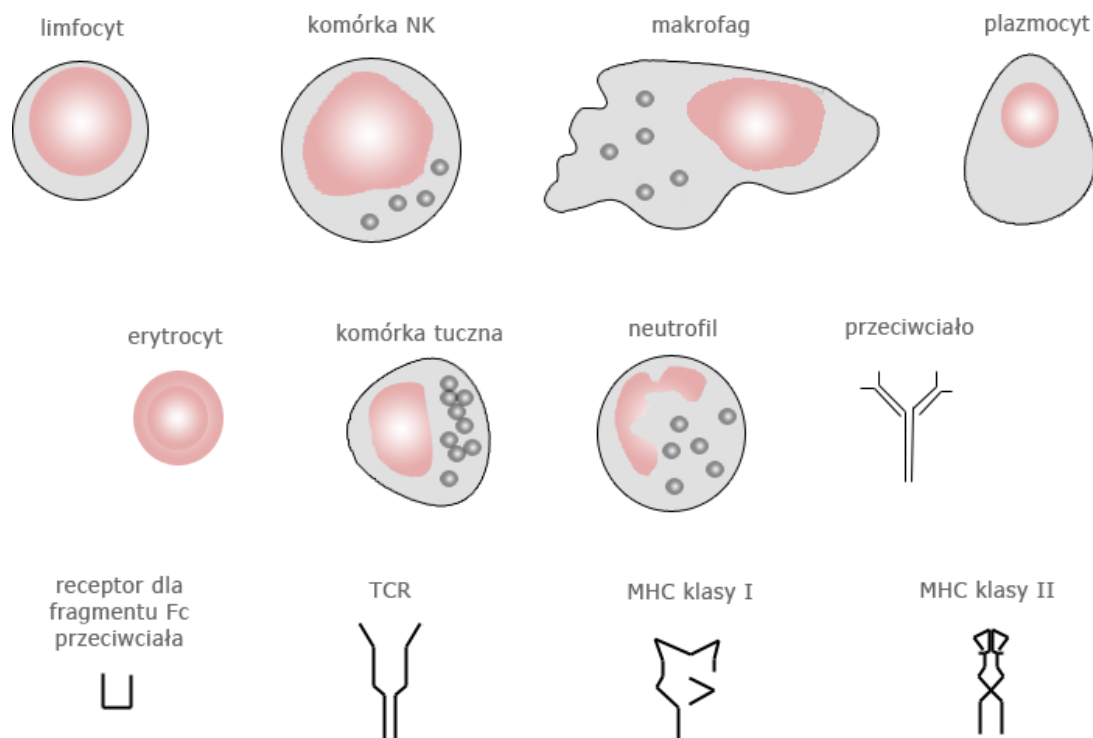
- wśród komórek odpornościowych:
 - limfocyty *T*;
 - limfocyty *B*;
 - komórki prezentujące antygen, ang. *antigen presenting cells* ([APC](#));
- wśród cząstek:
 - immunoglobuliny (przeciwciała);
 - receptory limfocytów wiążące się z antygenem.

Zestawienie najważniejszych składników układu odpornościowego przedstawiamy na rys. [12.1](#) i w tabeli [12.1](#).

W każdej odpowiedzi odpornościowej oprócz komórek efektorowych (np. limfocytów *B* wytwarzających przeciwciała lub cytotoksycznych limfocytów *T*) uczestniczą także komórki regulujące tę odpowiedź.

Rodzinę komórek zwanych limfocytami *T* dzielimy na kilka grup, w zależności od spełnianych przez nie funkcji:

- [limfocyty T cytotoksyczne](#) (*Tc*), które są zdolne do zabicia innych komórek, takich jak komórki obce, komórki nowotworowe, komórki zainfekowane przez wirusy;
- [limfocyty T pomocnicze](#), ang. *T helper cells* (*Th*), które wspomagają odpowiedź odpornościową;
- [limfocyty T regulatorowe](#), ang. *regulatory T cells* (*Treg*), które hamują odpowiedź odpornościową.



Rysunek 12.1. Najważniejsze składniki układu odpornościowego.

Z kolei limfocyty *B* należą do klasy komórek prezentujących antygen APC i uczestniczą w procesie rozpoznawania antygeny i przygotowywania swoistej odpowiedzi odpornościowej przeciw temu antygenowi. Następnie w procesie różnicowania stają się **komórkami plazmatycznymi**, które produkują przeciwciała zwalczające dany antygen.

Typy odpowiedzi odpornościowej Odpowiedź odpornościową możemy podzielić na:

- swoistą (specyficzną) i **nieswoistą** (niespecyficzną);
- typu **humoralnego** i **komórkowego**.

Mechanizmy nieswoiste nazywane są też odpornością natywną lub wrodzoną. Jednak żadna z tych nazw nie oddaje istoty tego typu reakcji. Podobnie zresztą dzieje się w przypadku mechanizmów swoistych nazywanych też odpornością adoptywną.

Mechanizmy nieswoiste rozwinęły się wcześniej w **filogenezie** (tzn. są starsze ewolucyjnie), są mniej precyzyjne, ale reakcje zachodzą szybko. Mówi się, że stanowią pierwszą linię obrony organizmu. Z kolei mechanizmy swoiste są filogenetycznie młodsze. W trakcie reakcji odpornościowej pojawienie się tych mechanizmów wymaga czasu, ale nakierowane są dokładnie przeciw konkretnemu rodzajowi antygeny, więc działają precyzyjniej i skuteczniej.

W odpowiedzi swoistej biorą udział głównie antygeny swoiste wyprodukowane przez komórki plazmatyczne powstałe w procesie rozpoznawania antygeny przez limfocyty *B* oraz limfocyty *T* ze swoistymi receptorami wiążącymi konkretny antygen.

Z kolei podział na odpowiedź humoralną i komórkową jest cokolwiek sztuczny. W większości przypadków obserwuje się oba typy reakcji, chociaż w konkretnych przypadkach dla danego antygeny jedna z nich może dominować nad drugą. W odpowiedzi humoralnej decydującą rolę w zwalczaniu antygeny odgrywają przeciwciała obecne głównie we krwi odwodowej i limfie, ale także w innych płynach organicznych, natomiast za odpowiedź komórkową odpowiadają limfocyty *T*, por. rys. 12.2. Limfocyty cytotoksyczne *T_c* bezpośrednio zwalczają antygen za pomocą

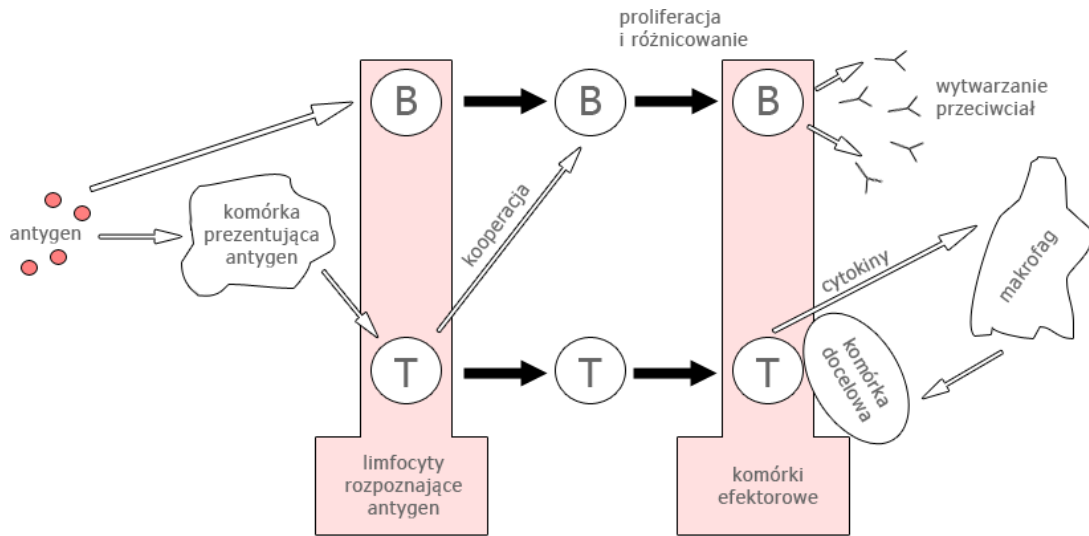
Limfocyty <i>B</i>	rozpoznają antygeny i wytwarzają przeciwciała
Komórki plazmatyczne (plazmocyty)	wytwarzają przeciwciała
Limfocyty <i>Th</i>	wspomagają odpowiedź odpornościową; wydzielają cytokiny
Limfocyty <i>Treg</i>	hamują odpowiedź odpornościową; wydzielają cytokiny
Limfocyty <i>Tc</i>	są zdolne do zabijania komórek docelowych
Komórki <i>K</i>	są zdolne do cytotoksyczności komórkowej zależnej od przeciwciał
Komórki <i>NK</i>	są zdolne do spontanicznego zabijania komórek nowotworowych i zakażonych przez wirusy
Komórki dendrytyczne	prezentują antygeny limfocytom <i>Th</i>
Monocyty, makrofagi	są zdolne do fagocytozy i zabicia mikroorganizmów oraz innych obcych komórek; wydzielają cytokiny
Neutrofile	są zdolne do fagocytozy i zabijania mikroorganizmów
Komórki tuczne, bazofile	biorą udział w nadwrażliwości I (anafilaksji)
Eozynofile	są zdolne do fagocytozy i zbijania niektórych pasożytów z użyciem przeciwciał
Megakariocyty	wytwarzają trombocyty , które uczestniczą w krzepnięciu krwi i reakcjach zapalnych
Erytrocyty	poprzez receptor dla dopełniacza wiążą kompleksy immunologiczne i uczestniczą w ich usuwaniu

Tabela 12.1. Zasadnicze czynności komórek uczestniczących w odpowiedzi odpornościowej.

swoistych receptorów, natomiast limfocyty pomocnicze *Th* w reakcji ze swoistym antygenem wydzielają cytokiny, które aktywują makrofagi i granulocyty niszczące antygen.

12.2. Proste modele odpowiedzi odpornościowej

Zajmiemy się teraz zaproponowaniem możliwie najprostszych modeli odpowiedzi odpornościowej, gdzie ponownie wykorzystamy znane już sposoby modelowania oddziaływań między



Rysunek 12.2. Rozwój odpowiedzi odpornościowej typu komórkowego (limfocyt T) i humoralnego (limfocyt B).

różnymi populacjami. Zaczniemy od opisu wewnętrznej dynamiki antygeny. Pod względem tejże dynamiki możemy antygeny podzielić na:

1. aktywne, jak bakterie, grzyby i wirusy, mające zdolność aktywnego namnażania się w organizmie;
2. pasywne, jak trucizny i jady, które się nie namnażają;
3. własne, czyli produkowane przez organizm, jak autoantygeny czy antygeny transplantacyjne.

Niech $V(t)$ oznacza zagęszczenie antygeny w organizmie w chwili t . W najprostszym przypadku wewnętrzną dynamikę pierwszego typu antygeny możemy opisać równaniem

$$\dot{V} = \beta V, \quad (12.1)$$

gdzie β oznacza współczynnik namnażania netto (czyli różnicę między współczynnikiem rozrodczości i śmiertelności). Oczywiście tylko gdy $\beta > 0$ antygen istotnie jest aktywny. Z kolei dla antygeny pasywnego możemy rozważyć takie samo równanie wewnętrznej dynamiki przy założeniu $\beta < 0$. Natomiast trzeciemu typowi antygeny odpowiada w najprostszym przypadku równanie

$$\dot{V} = \alpha - \mu V, \quad (12.2)$$

gdzie $\alpha > 0$ oznacza stałe źródło, a $\mu > 0$ — współczynnik śmiertelności, przy czym $1/\mu$ to średnia długość życia takiego antygeny.

Zakładamy teraz, że dany antygen wywołuje odpowiedź odpornościową, którą w modelu reprezentują efektor (limfocyty lub przeciwciała) o zagęszczeniu $E(t)$. Wobec tego w równaniu (12.1) (odpowiednio (12.2)) musimy uwzględnić kolejny składnik opisujący oddziaływanie antygen – efektor. Bazując na modelu drapieżnik – ofiara otrzymamy zamiast (12.1) równanie

$$\dot{V} = \beta V - \gamma V E,$$

gdzie γ odzwierciedla skuteczność efektorów w zwalczaniu antygenów.

Układ odpornościowy aktywuje się po wykryciu antygeny przez komórki odpornościowe. Wiemy też, że efektor zużywają się w trakcie odpowiedzi odpornościowej, zatem równanie ich dynamiki możemy zapisać w następującej postaci

$$\dot{E} = \eta V - \nu V E,$$

gdzie η oznacza współczynnik aktywacji, a ν — śmiertelność efektorów w wyniku reakcji odpornościowej.

Ostatecznie model odpowiedzi odpornościowej przeciw antygenowi aktywnemu opiszemy następującym układem dwóch równań różniczkowych zwyczajnych

$$\begin{aligned}\dot{V} &= \beta V - \gamma V E, \\ \dot{E} &= \eta V - \nu V E,\end{aligned}\tag{12.3}$$

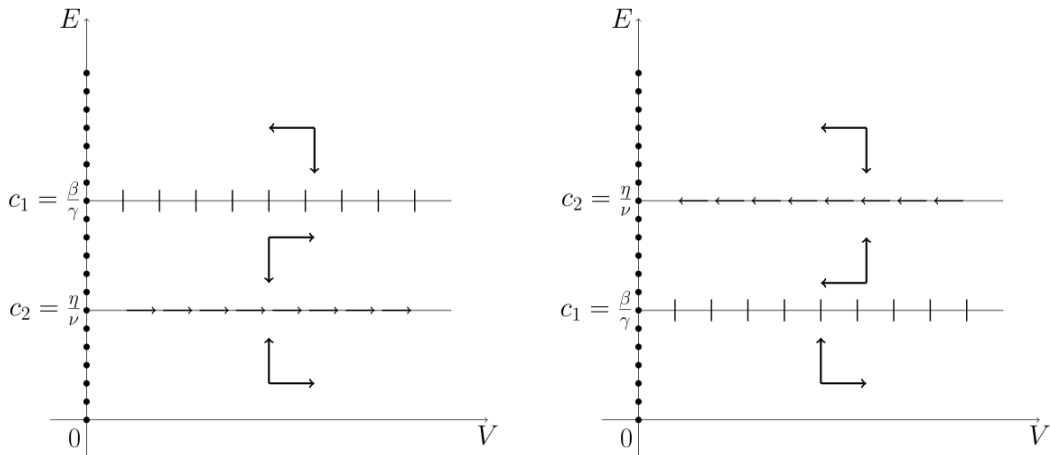
z dodatnimi współczynnikami. Analogicznie możemy modelować odpowiedź przeciwko antygenowi pasywnemu oraz autoantygenowi, co pozostawiamy jako ćwiczenie.

Przejdziemy teraz do analizy układu (12.3). Podstawowe własności jak istnienie, jednoznaczność, nieujemność i przedłużalność rozwiązań badamy w sposób standardowy. Istnienie i jednoznaczność rozwiązań dla dowolnego warunku początkowego (V_0, E_0) wynika z gładkości prawej strony układu. Nieujemność pierwszej zmiennej dla nieujemnego warunku początkowego wynika z tego, że oś $V \equiv 0$ jest zbiorem niezmienniczym. W przypadku drugiej zmiennej E , jeśli założymy, że istnieje $\bar{t} > 0$, dla którego $E(\bar{t}) = 0$, to otrzymamy $\dot{E}(\bar{t}) > 0$, zatem rozwiązanie nie może opuścić pierwszej ćwiartki układu współrzędnych. Nieujemność rozwiązań implikuje też ich przedłużalność dla wszystkich $t > 0$, ponieważ prawa strona układu ma w tym przypadku oszacowanie liniowe — obie pochodne szacują się przez cV , gdzie odpowiednio $c = \beta$ dla zmiennej V i $c = \eta$ dla E .

W kolejnym kroku analizy znajdziemy izokliny zerowe i naszkicujemy portrety fazowe w zależności od parametrów układu (12.3). Widzimy, że

$$\begin{aligned}\dot{V} = 0 &\iff V = 0 \text{ lub } E = \frac{\beta}{\gamma}, \\ \dot{E} = 0 &\iff V = 0 \text{ lub } E = \frac{\eta}{\nu}.\end{aligned}$$

Zauważmy, że oś $V = 0$ jest izokliną zarówno dla V jak i dla E . Mamy więc do czynienia z przypadkiem niehiperbolicznym. Cała oś $V = 0$ składa się z punktów krytycznych układu (12.3), por. rys. 12.3.



Rysunek 12.3. Szkice portretu fazowego układu (12.3) z zaznaczonymi izoklinami zerowymi i kierunkami przebiegu poszczególnych zmiennych w przypadku gdy $c_1 > c_2$ (po lewej) oraz w przypadku gdy $c_1 < c_2$ (po prawej).

Oznaczmy $c_1 = \frac{\beta}{\gamma}$, $c_2 = \frac{\eta}{\nu}$. Widzimy, że możemy wyznaczyć równanie orbit układu (12.3). Mamy

$$\frac{dV}{dE} = \frac{\beta - \gamma E}{\eta - \nu E} \quad \text{dla } E \neq c_2. \quad (12.4)$$

Zauważmy, że $E \equiv c_2$ stanowi rozwiązanie drugiego równania układu, gdyż wtedy $\dot{E} \equiv 0$. Stąd $\dot{V} = (\beta - \gamma c_2)V$ i funkcja V zmienia się w sposób wykładniczy. Dla tego rozwiązania w generycznych przypadkach mamy, jeśli $c_1 < c_2$, to $V \rightarrow 0$ przy $t \rightarrow +\infty$ oraz jeśli $c_1 > c_2$, to $V \rightarrow +\infty$. Niegeneryczny przypadek $c_1 = c_2$ pozostawiamy jako ćwiczenie.

Jeśli $E \neq c_2$, to możemy rozwiązać równanie (12.4). Dostajemy

$$\int_{V_0}^{V(t)} dV = \int_{E_0}^{E(t)} \frac{\beta - \gamma E}{\eta - \nu E} dE,$$

skąd

$$V(t) - V_0 = \frac{\gamma}{\nu} \int_{E_0}^{E(t)} \frac{E - c_1}{E - c_2} dE \implies V(t) = V_0 + \frac{\gamma}{\nu} \left(E(t) - E_0 + (c_2 - c_1) \ln \frac{E(t) - c_2}{E_0 - c_2} \right). \quad (12.5)$$

Rozpatrzmy przypadek $c_1 > c_2$, który opisuje silny antygen lub słaby układ odpornościowy, gdyż wtedy $\beta > \gamma\eta/\nu$, co oznacza duży współczynnik namnażania antygeny β w stosunku do jego śmiertelności w wyniku odpowiedzi odpornościowej (współczynnik γ) i aktywacji układu odpornościowego (η), przy dużej śmiertelności efektorów (ν). Izoklina $E = c_2$ stanowi separatrycę. Poniżej tej izokliny obie zmienne są rosnące i E jest ograniczona. Wynika stąd, że E ma granicę, natomiast V nie może mieć granicy, gdyż w obszarze $E < c_2$ nie ma punktów krytycznych o dodatniej współrzędnej V . Zatem $V \rightarrow +\infty$, co implikuje, że $E \rightarrow c_2$. W obszarze $c_2 < E < c_1$ obie zmienne są również monotoniczne, przy czym V rośnie, a E maleje. Skoro obie zmienne pozostają w tym obszarze, to analogicznie jak w pierwszym przypadku otrzymujemy $(V, E) \rightarrow (+\infty, c_2)$. Jeśli orbita zaczyna się w obszarze $E > c_1$, to obie zmienne maleją. Wobec tego orbita albo przechodzi do poprzedniego obszaru, albo ma granicę, zatem dąży do któregoś z punktów krytycznych na osi $V = 0$. Wtedy na mocy zależności (12.5) granica E_g spełnia równanie

$$0 = V_0 + \frac{\gamma}{\nu} \left(E_g - E_0 + (c_2 - c_1) \ln \frac{E_g - c_2}{E_0 - c_2} \right). \quad (12.6)$$

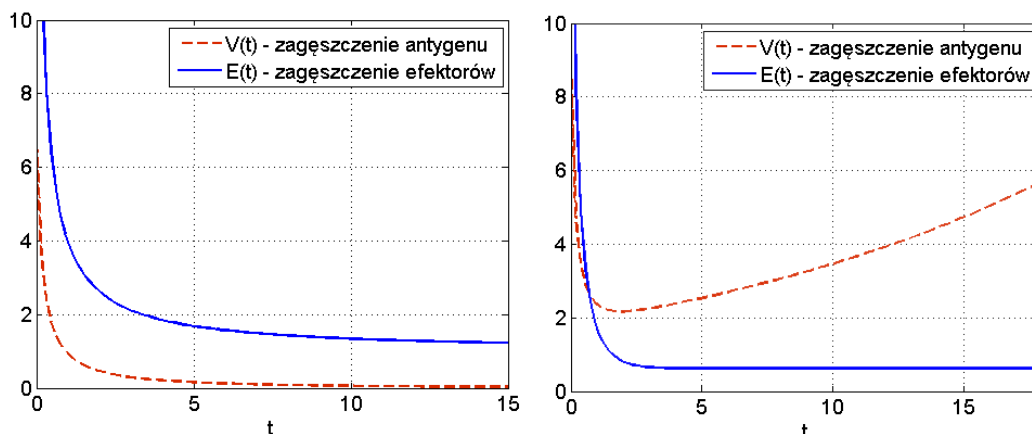
Taka dynamika jest w tym przypadku możliwa pod warunkiem, że $c_1 < E_g < E_0$, czyli

$$V_0 = \frac{\gamma}{\nu} \left(E_0 - E_g + (c_1 - c_2) \ln \frac{E_g - c_2}{E_0 - c_2} \right) \leq \frac{\gamma}{\nu} \left(E_0 - c_1 + (c_1 - c_2) \ln \frac{E_0 - c_2}{c_1 - c_2} \right). \quad (12.7)$$

Wniosek 12.1. *Jeśli $E_0 > c_1 > c_2$ i V_0 spełnia nierówność (12.7), to rozwiązania układu (12.3) dążą do punktu krytycznego $(0, E_g)$ opisanego zależnością (12.6). W przeciwnym przypadku $V \rightarrow +\infty$ i $E \rightarrow c_2$ przy $t \rightarrow +\infty$.*

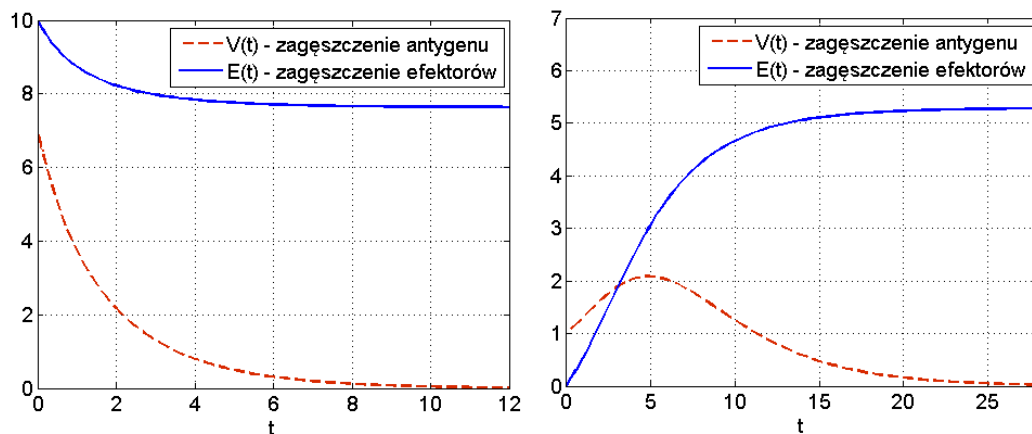
Wniosek 12.1 pokazuje, że jeśli antygen jest silny (spełnia nierówność $c_1 > c_2$), to wyzdrowienie jest możliwe tylko w przypadku dostatecznie dużego początkowego poziomu efektorów $E_0 > c_1$ oraz stosunkowo niewielkiej początkowej dawki antygeny V_0 spełniającej nierówność (12.7), por. rys. 12.4.

W przypadku słabego antygeny lub dostatecznie silnego układu odpornościowego, czyli dla $c_1 < c_2$, analogiczna analiza pokazuje, że wszystkie rozwiązania zbiegają do któregoś z punktów krytycznych $(0, E_g)$ i oczywiście granica E_g spełnia zależność (12.6). W tej sytuacji mamy



Rysunek 12.4. Rozwiązania układu (12.3) w przypadku silnego antygeny (spełniona jest nierówność $c_1 > c_2$) i wysokiego początkowego poziomu efektorów (w szczególności $E_0 > c_1$) oraz niewielkiej początkowej dawki antygeny V_0 spełniającej nierówność (12.7) (po lewej) lub nie spełniającej nierówności (12.7) (po prawej).

zawsze do czynienia z wyzdrowieniem, bez względu na początkową dawkę antygeny. Jeśli poziom efektorów jest dostatecznie duży, to zagęszczenie antygeny stale maleje w czasie, a jeśli jest niezbyt duży w stosunku do początkowej dawki antygeny, to w pierwszej fazie odpowiedzi odpornościowej układ nie może sobie poradzić i zagęszczenie antygeny wzrasta, ale po pewnym czasie wyprodukowana zostanie dostateczna liczba efektorów i poziom antygeny zaczyna maleć, w efekcie maleje do 0, por. rys. 12.5.



Rysunek 12.5. Rozwiązania układu (12.3) w przypadku słabego antygeny lub dostatecznie silnego układu odpornościowego, czyli dla $c_1 < c_2$.

Widzimy więc, że układ (12.3) jest zbyt prosty, aby mógł w odpowiedni sposób odzwierciedlić rzeczywiste procesy. W szczególności w rzeczywistości nie jest możliwe, by nawet bardzo silny układ odpornościowy zawsze poradził sobie z antygenem, bez względu na wielkość dawki początkowej.

13. Model Marczuka humoralnej odpowiedzi odpornościowej

Kolejny wykład poświęcimy na zapoznanie się z nieco bardziej skomplikowanym modelem odpowiedzi odpornościowej. W 1975 r. [Guri I. Marczuk](#) (właściwie G. I. Marchuk według obowiązującej transkrypcji rosyjskiego nazwiska, Marczuk to spolszczona wersja tego nazwiska) zaproponował dość prosty model, który jest w stanie odzwierciedlić najważniejsze przebiegi odpowiedzi odpornościowej, w tym również chorobę przewlekłą (chroniczną). Model opisany jest za pomocą czterech równań z opóźnionym argumentem i z założenia odzwierciedla humoralną odpowiedź odpornościową, choć podobną strukturę modelu można także zaproponować w przypadku odpowiedzi komórkowej. Model ten będziemy w dalszym ciągu nazywać modelem Marczuka. W 1980 r. ukazała się książka Marczuka [11], w której oprócz tego najprostszego modelu przedstawił także kilka wersji bardziej rozbudowanych, w tym model opisujący równoległy przebieg odpowiedzi humoralnej i komórkowej. W późniejszych latach ukazały się także tłumaczenia tej książki na język angielski [12] i polski [13]. Wyniki doświadczalne dla pewnych chorób okazały się bardzo podobne do teoretycznych wyników uzyskanych na podstawie modelu Marczuka.

Z powstaniem modelu wiąże się ciekawa historia. Marczuk przez wiele lat był przewodniczącym [Akademii Nauk ZSSR](#). Niestety zachorował na [żółtaczkę zakaźną](#), a pobyt w szpitalu urozmaicał sobie poznawaniem mechanizmów odpornościowych. W rezultacie doprowadziło to do powstania modelu reakcji odpornościowej i zaowocowało wieloma pracami naukowymi w tej dziedzinie. Po powrocie do zdrowia Marczuk w dalszym ciągu prowadził działalność naukową, w ramach której odwiedził Wydział Matematyki, Informatyki i Mechaniki UW, gdzie zaprezentował model i wyniki symulacji wskazujące możliwości leczenia choroby przewlekłej za pomocą zaostżenia infekcji. Zarówno sam model, jak i możliwości jego zastosowania, zafascynowały jednego z pionierów polskiej biomatematyki, [prof. Wiesława Szlenka](#). Tematyka ta znalazła się później wśród zagadnień badawczych w powstałym po pewnym czasie Instytucie Matematyki Stosowanej i Mechaniki.

13.1. Prezentacja modelu Marczuka

Zajmiemy się teraz omówieniem konstrukcji prostego modelu odpowiedzi odpornościowej zaproponowanego przez G. I. Marczuka, a następnie przedstawimy podstawowe własności układu.

Przypomnijmy, że reakcja odpornościowa następuje w wyniku wniknięcia do organizmu obcego czynnika, który nazywamy antygenem. W przypadku omawianego modelu antygenem są głównie bakterie, ponieważ np. wirusy wywołują znacznie bardziej skomplikowane reakcje, które wymagają bardziej skomplikowanego opisu matematycznego, w szczególności uwzględnienia także komórkowej odpowiedzi odpornościowej. Pojawienie się antygeny wywołuje kaskadę reakcji, które w efekcie doprowadzają do rozpoznania antygeny i uruchomienia produkcji przeciwciał swoistych przez komórki plazmatyczne. Mechanizmem uruchamiającym tę kaskadę jest powstanie tzw. kompleksu *VT*, czyli przyłączenie do antygeny przeciwciała prezentującego. Wobec tego zakładamy, że produkcja komórek plazmatycznych zależy od liczby wiązań antygen – prze-

ciwciało, przy czym obserwujemy opóźnienie między powstaniem takiego wiązania a wyprodukowaniem przed układ odpornościowy komórek plazmatycznych produkujących przeciwciała swoiste.

W modelu występują następujące podstawowe zmienne

- $V(t)$ — zagęszczenie populacji bakterii (antygeny) w organizmie w chwili t ,
- $C(t)$ — zagęszczenie komórek plazmatycznych w chwili t ,
- $F(t)$ — stężenie przeciwciał w chwili t ,

przy czym (jak dotychczas) używamy określenia zagęszczenie w stosunku do osobników pewnej populacji (tu populacji komórkowych), natomiast stężenie dotyczy głównie substancji, a wiemy, że przeciwciała są białkami.

Dynamikę antygeny opisujemy w modelu Marczuka w taki sam sposób jak w najprostszym modelu odpowiedzi odpornościowej prezentowanym poprzednio. Zakładamy, że bakterie rozmnażają się za stałym współczynnikiem reprodukcji $\beta > 0$ i giną w wyniku reakcji odpornościowych. Zatem śmiertelność spowodowana jest wiązaniem antygenów z przeciwciałami, proporcjonalnie do liczby tych wiązań ze współczynnikiem $\gamma > 0$, który odzwierciedla prawdopodobieństwo związania antygeny z przeciwciałem i siłę ich reakcji. Stąd

$$\dot{V}(t) = \beta V(t) - \gamma V(t)F(t).$$

Przeciwciała giną nie tylko w wyniku wiązania z antygenem, ale też w procesie naturalnej degradacji. Są one produkowane przez komórki plazmatyczne w stałym tempie $\rho > 0$, czyli

$$\dot{F}(t) = \rho C(t) - \eta \gamma V(t)F(t) - \mu_F F(t),$$

gdzie $\eta > 0$ opisuje frakcję związanych przeciwciał ginących w reakcji odpornościowej, a $1/\mu_F$ odzwierciedla średni czas przetrwania przeciwciała.

Zauważyliśmy wcześniej, że produkcja komórek plazmatycznych uruchamiana jest przez powstanie kompleksów VT i jest opóźniona w stosunku do mechanizmu uruchamiającego. Dla uproszczenia zakładamy, że opóźnienie to jest stałe, równe τ . W organizmie utrzymuje się też stały fizjologiczny poziom przeciwciał C^* . Skoro tak, to jeśli nie ma antygeny, wewnętrzną dynamikę komórek plazmatycznych możemy opisać równaniem

$$\dot{C}(t) = \mu_C(C^* - C(t)),$$

gdzie $1/\mu_C$ oznacza średnią długość życia komórki plazmatycznej. Wtedy $C(t) = C^* + (C_0 - C^*)e^{-\mu_C t} \rightarrow C^*$ przy $t \rightarrow \infty$, co oznacza, że w zdrowym organizmie po odchyleniu od poziomu fizjologicznego zawsze następuje samoczynny powrót do poziomu.

Jeśli antygen występuje w organizmie, to w powyższym równaniu uwzględniamy jeszcze składnik związany ze stymulacją odpowiedzi odpornościowej

$$\dot{C}(t) = \alpha V(t - \tau)F(t - \tau) - \mu_C(C(t) - C^*),$$

przy czym $\alpha > 0$ jest współczynnikiem immunogenności antygeny odzwierciedlającym siłę odpowiedzi odpornościowej na dany antygen.

Prócz tych trzech podstawowych zmiennych w modelu występuje zmienna dodatkowa, która opisuje stopień uszkodzenia organu zaatakowanego przez antygen, ponieważ Marczuk budował swój model z myślą o żółtacze i stopniu uszkodzenia wątroby w wyniku tej choroby. Niech M oznacza charakterystykę (masę lub powierzchnię) zdrowego organu, a $M(t)$ — zdrową część organu uszkodzonego przez antygen. Definiujemy

$$m(t) = \frac{M - M(t)}{M}.$$

Mamy zatem $m \in [0, 1]$, co opisuje stosunek części uszkodzonej do całości. Oczywiście $m = 0$ dla organu zdrowego oraz $m = 1$ dla organu całkowicie zniszczonego. Dynamikę $m(t)$ opisujemy za pomocą prostego równania liniowego

$$\dot{m}(t) = \sigma V(t) - \mu_m m(t) \quad \text{dla } m(t) \leq 1,$$

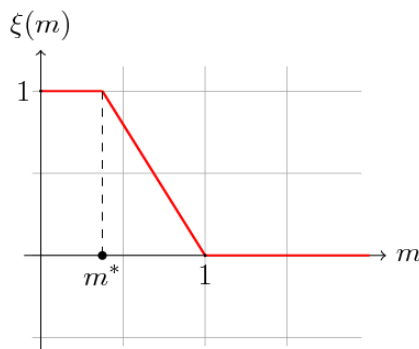
gdzie $\sigma > 0$ odzwierciedla zdolność antygeny do niszczenia organu, a $1/\mu_m$ oznacza średni czas regeneracji organu.

Wielkość $m(t)$ osłabia odpowiedź odpornościową poprzez zmniejszenie stymulacji produkcji komórek plazmatycznych, zatem zamiast współczynnika immunogenności α rozpatrujemy w modelu współczynnik $\alpha\xi(m)$, gdzie $\xi(m)$ jest nierosnącą funkcją zmiennej m , $\xi(0) = 1$ i $\xi(1) = 0$.

Ostatecznie model Marczuka humoralnej odpowiedzi odpornościowej opisujemy układem 4. równań różniczkowych z opóźnionym argumentem (rrzoa)

$$\begin{aligned} \dot{V}(t) &= (\beta - \gamma F(t))V(t), \\ \dot{C}(t) &= \alpha\xi(m(t))V(t - \tau)F(t - \tau) - \mu_C(C(t) - C^*), \\ \dot{F}(t) &= \varrho C(t) - (\mu_F + \eta\gamma V(t))F(t), \\ \dot{m}(t) &= \sigma V(t) - \mu_m m(t) \quad \text{dla } m(t) \leq 1. \end{aligned} \tag{13.1}$$

W większości przypadków istnieje pewna progowa wielkość m^* , do której m nie wpływa znacząco na reakcję odpornościową i można założyć, że $\xi(m) = 1$ dla $m < m^*$.



Rysunek 13.1. Przebieg typowej funkcji $\xi(m)$, czyli funkcja tożsamościowo równa 1 na odcinku $[0, m^*]$ oraz malejąca liniowo do 0 na $[m^*, 1]$.

Typowo w analizie matematycznej układu (13.1) zakładamy, że $m < m^*$ i wtedy możemy zredukować układ do 3 rrzoa, choć dla konkretnej postaci funkcji ξ można też badać np. stabilność całego układu (13.1).

13.2. Podstawowe własności układu (13.1)

Przejdziemy teraz do zbadania podstawowych własności układu (13.1). Jest on układem równań z opóźnionym argumentem, zatem jako warunek początkowy należy zadać funkcję określoną na odcinku $[-\tau, 0]$ o wartościach w $(\mathbb{R}^+)^3 \times [0, 1]$. Ponieważ jednak tylko współrzędne $V(t)$ i $F(t)$ występują z opóźnionym argumentem, więc możemy zadać tylko ich wartości na całym przedziale $[-\tau, 0]$, ograniczając się do wartości C_0 i m_0 dla zmiennych C i m . Należy przy tym pamiętać, że z formalnego punktu widzenia potrzebne są wszystkie cztery współrzędne funkcji początkowej.

Istnienie, jednoznaczność i nieujemność rozwiązań dla nieujemnego warunku początkowego możemy wykazać stosując metodę kroków. Przypomnijmy, że opiera się ona na zasadzie indukcji matematycznej i pozwala zastosować znane techniki dotyczące równań różniczkowych zwyczajnych do analizy rzroa.

Istnienie lokalnego rozwiązania układu (13.1). Niech $V_0(s), F_0(s) : [-\tau, 0] \rightarrow \mathbb{R}^+$. Wtedy dla $t \in [0, \tau]$ układ (13.1) sprowadza się do niejednorodnego układu rzr postaci

$$\begin{aligned}\dot{V}(t) &= (\beta - \gamma F(t))V(t), \\ \dot{C}(t) &= \alpha \xi(m(t))G_0(t) - \mu_C(C(t) - C^*), \\ \dot{F}(t) &= \varrho C(t) - (\mu_F + \eta\gamma V(t))F(t), \\ \dot{m}(t) &= \sigma V(t) - \mu_m m(t) \quad \text{dla } m(t) \leq 1,\end{aligned}\tag{13.2}$$

gdzie $G_0(t) = V_0(t - \tau)F_0(t - \tau)$ jest znaną funkcją, gdyż dla $t \in [0, \tau]$ argument $t - \tau \in [-\tau, 0]$. Funkcje opisujące współrzędne prawej strony układu (13.2) są wielomianami kwadratowym względem współrzędnych, czyli są klasy \mathbf{C}^1 (faktycznie \mathbf{C}^∞), więc spełniają założenia twierdzeń o lokalnym istnieniu i jednoznaczności rozwiązań równań różniczkowych zwyczajnych. Istnieje zatem odcinek $[0, \bar{t}) \subset [0, \tau]$, na którym mamy jednoznaczne rozwiązanie układu.

Nieujemność lokalnego rozwiązania układu (13.1). Pokażemy najpierw, że rozwiązanie to jest nieujemne. Pierwsze równanie układu (13.2) zapisujemy w równoważnej postaci całkowej

$$\frac{\dot{V}}{V} = \beta - \gamma F \implies \ln \left| \frac{V}{V_0} \right| = \int_0^t (\beta - \gamma F(s)) ds \implies V(t) = V_0 e^{\beta t - \int_0^t \gamma F(s) ds},$$

gdzie w powyższym wzorze $V_0 = V_0(0)$. Zatem $V(t) \geq 0$ dla $V_0 \geq 0$, co więcej, $V(t) > 0$ dla $V_0 > 0$, czyli z matematycznego punktu widzenia wyzdrowienie może wystąpić tylko asymptotycznie. W praktyce zagęszczenie antygeny niższe od czułości urządzeń pomiarowych jest niewykrywalne i takie zagęszczenia uznajemy za równoważne wyzdrowieniu.

Dla $\xi(m) = 1$ drugie równanie układu (13.2) jest równaniem liniowym niejednorodnym. Rozwiązujemy je stosując metodę uzmienniania stałej i otrzymujemy

$$C(t) = C^* + (C_0 - C^*)e^{-\mu_C t} + \int_0^t G_0(s)e^{-\mu_C(t-s)} ds.\tag{13.3}$$

Ponieważ $G(s) \geq 0$ oraz $C^* + (C_0 - C^*)e^{-\mu_C t} = C^*(1 - e^{-\mu_C t}) + C_0 e^{-\mu_C t} > 0$ dla $t > 0$, to $C(t) > 0$ dla $t \in [0, \tau]$. Zauważmy też, że w tym przypadku współrzędna C jest zdefiniowana za pomocą wzoru (13.3) na całym przedziale $[0, \tau]$.

W ogólnym przypadku, gdy $\xi(m) < 1$, oszacujemy wartość pochodnej \dot{C} z dołu przez $-\mu_C C(t)$, więc

$$\dot{C} \geq -\mu_C C \implies C(t) \geq C_0 e^{-\mu_C t}.$$

Zatem $C(t) \geq 0$ dla $C_0 \geq 0$. Zauważmy także, że jeśli $C_0 = C^*$, to z powyższych nierówności wynika, że $C(t) > C^*$ dla dowolnego $t > 0$.

W analogiczny sposób dowodzimy nieujemności zmiennej $m(t)$.

Skoro $C(t) \geq 0$ dla $t \in [0, \bar{t})$, to podobnie oszacujemy \dot{F} , czyli

$$F(t) \geq F_0 e^{-\int_0^t (\eta\gamma V(s) + \mu_F) ds}.$$

Zatem $F(t) \geq 0$ dla $F_0 \geq 0$.

Istnienie rozwiązania układu (13.1) na przedziale $[0, \tau]$. Pokażemy teraz, że z nieujemności rozwiązań wynika ich istnienie na całym przedziale $[0, \tau]$. W tym celu oszacujemy nasze rozwiązanie z góry. Dla pierwszej zmiennej mamy

$$\dot{V} \leq \beta V \implies V(t) \leq V_0 e^{\beta t},$$

zatem wzrost $V(t)$ jest co najwyżej wykładniczy.

Dla zmiennej $C(t)$ mamy

$$\dot{C} \leq \alpha G_0 \implies C(t) \leq C_0 + \alpha \int_0^t G_0(s) ds,$$

czyli $C(t)$ jest również ograniczona na każdym podprzedziale odcinka $[0, \tau]$.

Z ograniczoności $C(t)$ i $V(t)$ wynika ograniczoność $F(t)$ i $m(t)$. Skoro funkcje te są ograniczone, więc ograniczone są także ich pochodne, czyli rozwiązanie można przedłużyć biorąc jako nowy warunek początkowy granicę $\lim_{t \rightarrow \bar{t}^-} (V(t), C(t), F(t), m(t))$ (istnienie granicy wynika bezpośrednio z ograniczoności pochodnej) oraz punkt \bar{t} i stosując twierdzenie o istnieniu rozwiązań. Wobec dowolności \bar{t} wnioskujemy, że rozwiązanie istnieje na całym $[0, \tau]$.

Globalne istnienie rozwiązań układu (13.1). Stosując metodę kroków zakładamy, że istnieje nieujemne rozwiązanie układu (13.1) na odcinku $[(n-1)\tau, n\tau]$ i powtarzamy powyższe rozumowanie w celu wykazania istnienia nieujemnego rozwiązania na kolejnym odcinku $[n\tau, (n+1)\tau]$. Na mocy zasady indukcji matematycznej jednoznacznie, nieujemne rozwiązanie układu (13.1) istnieje dla dowolnego $t \geq 0$.

Punkty krytyczne układu (13.1). Znajdziemy teraz punkty krytyczne badanego układu. Rozwiązujemy następujący układ równań algebraicznych

$$\begin{aligned} 0 &= (\beta - \gamma \bar{F}) \bar{V}, \\ 0 &= \alpha \xi(\bar{m}) \bar{V} \bar{F} - \mu_C (\bar{C} - C^*), \\ 0 &= \varrho \bar{C} - (\mu_F + \eta \gamma \bar{V}) \bar{F}, \\ 0 &= \sigma \bar{V} - \mu_m \bar{m}. \end{aligned} \tag{13.4}$$

W ogólnym przypadku, aby rozwiązać układ (13.4), musimy zadać konkretną postać funkcji ξ . Dla uproszczenia założmy więc, że $m < m^*$, czyli $\xi(m) = 1$. W tym przypadku badanie możemy ograniczyć do układu trzech równań, a następnie sprawdzić zachowanie ostatniej zmiennej.

Łatwo widać, że układ (13.4) ma rozwiązanie opisujące stan zdrowia, czyli $V = 0$ i $C = C^*$. Wtedy $F = \frac{\varrho}{\mu_F} C^* = F^*$ i oczywiście $m = 0$. F^* nazwiemy poziomem fizjologicznym przeciwciał.

Niech $X_1 = (0, C^*, F^*)$ odzwierciedla stan zdrowia organizmu.

Jeśli $V \neq 0$, to z pierwszego równania układu (13.4) mamy $\bar{F} = \frac{\beta}{\gamma}$. Wstawiając tę zależność do drugiego i trzeciego równania obliczmy

$$\bar{V} = \frac{\mu_C (\mu_F \beta - \gamma \varrho C^*)}{\beta (\alpha \varrho - \eta \gamma \mu_C)}, \quad \bar{C} = \frac{\alpha \mu_F \beta - \eta \gamma^2 \mu_C C^*}{\gamma (\alpha \varrho - \eta \gamma \mu_C)}.$$

Punkt krytyczny $X_2 = (\bar{V}, \bar{F}, \bar{C})$ odzwierciedla chorobę przewlekłą, oczywiście pod warunkiem, że wszystkie jego współrzędne są dodatnie. Współrzędne \bar{V} i \bar{C} możemy przepisać w nieco innej postaci

$$\bar{V} = \frac{\mu_C \mu_F (\beta - \gamma F^*)}{\beta (\alpha \varrho - \eta \gamma \mu_C)}, \quad \bar{C} = \frac{\mu_F (\alpha \beta \varrho - \eta \gamma^2 \mu_C F^*)}{\gamma \varrho (\alpha \varrho - \eta \gamma \mu_C)},$$

z czego wynika, że jeśli

1. $\beta > \gamma F^*$ i $\alpha \varrho > \eta \gamma \mu_C$ albo
2. $\beta < \gamma F^*$ i $\alpha \varrho < \eta \gamma \mu_C$,

to $\bar{V} > 0$. Zauważmy dalej, że przy warunku A mamy:

$$\alpha \beta \varrho - \eta \gamma^2 \mu_C F^* > \gamma F^* (\alpha \varrho - \eta \gamma \mu_C) > 0, \quad (13.5)$$

zatem $\bar{C} > 0$. Przy warunku B nierówność (13.5) zmienia znak, ale znak zmienia też mianownik \bar{C} , czyli $\bar{C} > 0$. Widzimy więc, że warunki A, B są konieczne i dostateczne dla istnienia dodatniego punktu krytycznego. Zauważmy także, że w tym przypadku $\bar{m} = \frac{\sigma}{\mu_m} \bar{V} > 0$ oraz $\bar{V} < \frac{\mu_m}{\sigma} m^*$. Ostatnia nierówność oznacza, że w omawianym przypadku $m < m^*$ poziom stacjonarny antygeny \bar{V} nie może być zbyt duży. W literaturze mówi się o lekkiej chorobie przewlekłej dla $m < m^*$, natomiast gdy $m > m^*$ — o ciężkiej.

Lokalna stabilność punktów krytycznych układu (13.1). Badanie lokalnej stabilności punktów krytycznych rzosa nie różni się w początkowej fazie od analogicznej analizy dla rzr. Niech $(\tilde{V}, \tilde{C}, \tilde{F})$ oznacza punkt krytyczny. Zdefiniujmy nowe zmienne: $v = V - \tilde{V}$, $c = C - \tilde{C}$, $f = F - \tilde{F}$. Układ w nowych zmiennych ma postać

$$\begin{aligned} \dot{v}(t) &= (\beta - \gamma \tilde{F} - \gamma f(t))v(t) - \gamma \tilde{V} f(t), \\ \dot{c}(t) &= \alpha(\tilde{V} f(t - \tau) + \tilde{F} v(t - \tau) + v(t - \tau)f(t - \tau)) - \mu_C c(t), \\ \dot{f}(t) &= \varrho c(t) - \mu_F f(t) - \eta \gamma (\tilde{F} v(t) + \tilde{V} f(t) + v(t)f(t)). \end{aligned} \quad (13.6)$$

Linearyzujemy układ (13.6) wokół punktu $(0, 0, 0)$ zakładając, że $|v(t)| < \varepsilon$, $|c(t)| < \varepsilon$, $|f(t)| < \varepsilon$ dla $t > 0$. Składniki wyższego rzędu, czyli $v(t)f(t)$ i $v(t - \tau)f(t - \tau)$ szacują się przez ε^2 , co możemy pominać. Po zlinearyzowaniu układ (13.6) przyjmuje postać

$$\begin{aligned} \dot{v}(t) &= (\beta - \gamma \tilde{F})v(t) - \gamma \tilde{V} f(t), \\ \dot{c}(t) &= \alpha(\tilde{V} f(t - \tau) + \tilde{F} v(t - \tau)) - \mu_C c(t), \\ \dot{f}(t) &= \varrho c(t) - \mu_F f(t) - \eta \gamma (\tilde{F} v(t) + \tilde{V} f(t)). \end{aligned} \quad (13.7)$$

W kolejnym kroku znajdujemy rozwiązania układu (13.7) postaci $(v(t), c(t), f(t)) = (v_0, c_0, f_0)e^{\lambda t}$, $\lambda \in \mathbb{C}$, $\lambda \neq 0$. Otrzymujemy więc następujące zależności

$$\begin{aligned} 0 &= (-\lambda + \beta - \gamma \tilde{F})v_0 - \gamma \tilde{V} f_0, \\ 0 &= \alpha \tilde{F} v_0 e^{-\lambda \tau} + \alpha \tilde{V} f_0 e^{-\lambda \tau} - (\mu_C + \lambda)c_0, \\ 0 &= -\eta \gamma \tilde{F} v_0 + \varrho c_0 - (\eta \gamma \tilde{V} + \mu_F + \lambda)f_0. \end{aligned} \quad (13.8)$$

Układ (13.8) ma nietrywialne rozwiązania, jeśli

$$\det \begin{pmatrix} \beta - \gamma \tilde{F} - \lambda & 0 & -\gamma \tilde{V} \\ \alpha \tilde{F} e^{-\lambda \tau} & -\mu_C - \lambda & \alpha \tilde{V} e^{-\lambda \tau} \\ -\eta \gamma \tilde{F} & \varrho & -\eta \gamma \tilde{V} - \mu_F - \lambda \end{pmatrix} = 0. \quad (13.9)$$

Twierdzenie o linearyzacji gwarantuje, że jeśli wartości własne, czyli rozwiązania równania (13.8) mają części rzeczywiste ujemne, to rozwiązanie $(0, 0, 0)$ układu (13.6) jest lokalnie asymptotycznie stabilne, natomiast jeśli istnieje wartość własna o części rzeczywistej dodatniej, to rozwiązanie jest niestabilne.

Zbadamy wartości własne dla poszczególnych punktów krytycznych. Zaczniemy od X_1 . Wtedy macierz układu ma postać dolnietrójkątną, zatem łatwo wyznaczamy wartości własne, gdyż spełniają równanie

$$(\beta - \gamma\tilde{F} - \lambda)(\mu_C + \lambda)(\eta\gamma\bar{V} + \mu_F + \lambda) = 0,$$

czyli wartości własne są równe: $\lambda_1 = \beta - \gamma F^*$, $\lambda_2 = -\mu_C$, $\lambda_3 = -(\eta\gamma\bar{V} + \mu_F)$, zatem są rzeczywiste i ujemne dla $\beta < \gamma F^*$.

Wniosek 13.1. *Jeśli $\beta < \gamma F^*$, to rozwiązanie X_1 jest lokalnie asymptotycznie stabilne. Natomiast jeśli $\beta > \gamma F^*$, to X_1 jest niestabilne.*

Widzimy, że stabilność rozwiązania X_1 , czyli wyzdrowienie jest możliwe tylko przy niewielkiej reproduktywności antygeny β w stosunku do poziomu fizjologicznego przeciwciał F^* .

Badając stabilność rozwiązania X_2 otrzymujemy następujący pseudowielomian charakterystyczny

$$W(\lambda) = \lambda^3 + a\lambda^2 + b\lambda - d + (f - g\lambda)e^{-\lambda\tau}, \quad (13.10)$$

gdzie

$$\begin{aligned} a &= \mu_C + \mu_F + \eta\gamma\bar{V}, \quad b = \mu_C(\mu_F + \eta\gamma\bar{V}) - \eta\beta\gamma\bar{V} \\ d &= \eta\gamma\mu_C\beta\bar{V}, \quad g = \alpha\varrho\bar{V}, \quad f = \beta g, \end{aligned} \quad (13.11)$$

a wyznaczenie współczynników pozostawiamy jako ćwiczenie. Przypomnijmy, że dla $\tau > 0$ pseudowielomian ma nieskończenie wiele pierwiastków, co utrudnia jego analizę (podczas gdy dla wielomianów możemy skorzystać np. z [kryt. Routha – Hurwitza](#)).

Wiemy, że jeśli rozwiązanie jest stabilne dla $\tau = 0$, to z ciągłej zależności od parametrów wynika, że albo rozwiązanie jest stabilne dla wszystkich $\tau > 0$, albo istnieje krytyczna wartość parametru τ_c , dla której może nastąpić destabilizacja. Dla $\tau = 0$ możemy zbadać stabilność punktu X_2 za pomocą kryterium Routha – Hurwitza, skąd dostajemy warunek stabilności

$$0 < f - d < a(b - g). \quad (13.12)$$

Warunek $f - d > 0$ jest równoważny $\alpha\varrho > \eta\gamma\mu_C$, zatem stabilność X_2 dla $\tau = 0$ możliwa jest tylko w przypadku A. Natomiast w przypadku B mamy stabilność punktu X_1 .

Zauważmy, że zmiana stabilności może nastąpić tylko gdy pseudowielomian charakterystyczny (13.10) ma pierwiastki czysto urojone $\lambda = i\omega$. Wtedy

$$-i\omega^3 - a\omega^2 + i b\omega - d = -(f - i g\omega) \exp(-i\omega\tau_c) \implies \|-i\omega^3 - a\omega^2 + i b\omega - d\| = \|f - i g\omega\|,$$

czyli

$$\omega^6 + (a^2 - 2b)\omega^4 + (b^2 + 2ad - g^2)\omega^2 + d^2 - f^2 = 0.$$

Zdefiniujmy funkcję pomocniczą

$$\Phi(z) = z^3 + (a^2 - 2b)z^2 + (b^2 + 2ad - g^2)z + d^2 - f^2, \quad z = \omega^2.$$

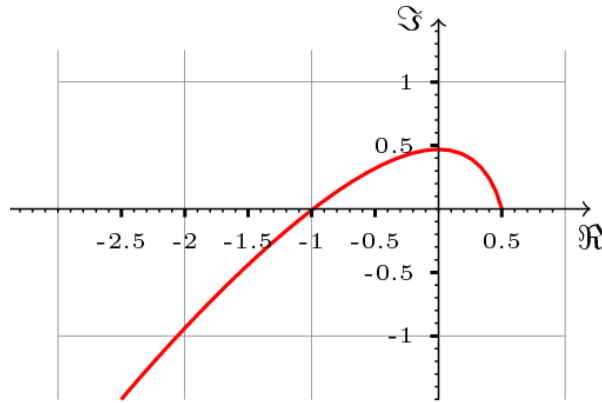
Jej dodatnie miejsca zerowe determinują potencjalne zmiany stabilności punktu krytycznego X_2 . Z warunku (13.12) wynika, że wyraz wolny wielomianu Φ jest ujemny, zatem istnieje co najmniej jedno dodatnie miejsce zerowe. Warunek A lub B implikuje też $a^2 - 2b > 0$, czyli dodatnie miejsce zerowe Φ jest wyznaczone jednoznacznie. Niech $\omega_0 = \sqrt{z_0}$, gdzie $\Phi(z_0) = 0$. Podstawiając

$\lambda = i\omega_0$ do pseudowielomianu (13.10) dostajemy układ równań, z którego wyznaczamy sinus i kosinus wartości krytycznej $\omega_0\tau_c$. Mamy zatem krytyczną wartość opóźnienia zdefiniowaną w sposób uwikłany za pomocą tych funkcji trygonometrycznych. Skoro jednak znamy tylko wartości sinusa i kosinusa, to oczywiście mamy nie jedną wartość krytyczną, a cały ciąg takich wartości $(\tau_{c_n})_{n=1}^{\infty}$ z pierwszą wartością τ_{c_1} , taką że $\omega_0\tau_{c_1} \in (0, 2\pi)$ i $\tau_{c_n} = \tau_{c_1} + 2(n-1)\pi/\omega_0$.

Zmiana stabilności następuje przy takiej krytycznej wartości τ_c , dla której istnieje para czysto urojonych wartości własnych $\pm i\omega$ i przy przekraczaniu osi urojonej wartości własne przechodzą z lewej półpłaszczyzny zespolonej na prawą, czyli $\frac{d}{d\tau}\Re W(\lambda)\Big|_{\tau=\tau_c} > 0$. Przypomnijmy, że znak tej pochodnej jest taki sam jak znak pochodnej funkcji pomocniczej Φ w odpowiednim miejscu zerowym — tutaj jedynym miejscem zerowym z_0 . W tej sytuacji mamy oczywiście dodatnią pochodną $\Phi'(z_0)$ i wartości własne przechodzą z lewej półpłaszczyzny zespolonej na prawą.

Wniosek 13.2. *Jeśli spełnione są nierówności (13.12) i istnieje dodatni punkt krytyczny X_2 , to punkt ten jest stabilny dla $\tau = 0$ i istnieje krytyczna wartość opóźnienia τ_c , przy której X_2 traci stabilność i pozostaje niestabilny dla $\tau > \tau_c$. W $\tau = \tau_c$ zachodzi bifurkacja Hopfa, czyli pojawiają się nietrywialne rozwiązania okresowe o okresie $\frac{2\pi}{\omega_0}$.*

Przypomnijmy, że z nierówności $f - d > 0$ wynika, że stabilność punktu X_2 przy $\tau = 0$ jest możliwa tylko w przypadku spełnienia warunku A.



Rysunek 13.2. Hodograf Michajłowa dla parametrów: $a = f = 1$, $d = g = \tau = 0.5$, $b = 2$.

Korzystając z kryterium Michajłowa, można udowodnione następujące twierdzenie dotyczące stabilności punktu X_2 , które daje nam oszacowanie wielkości opóźnienia krytycznego τ_c

Twierdzenie 13.1. *Niech $\mu_C\tau \leq 1$ oraz*

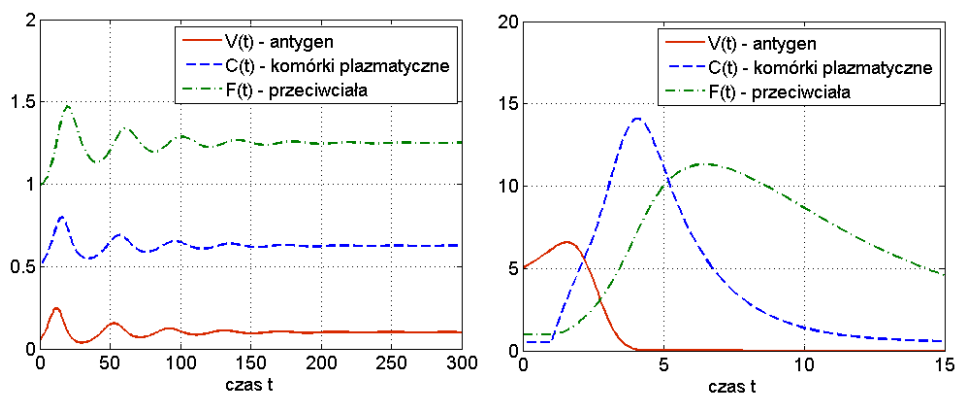
$$0 < \frac{f-d}{a-g\tau} < b-g-f\tau. \quad (13.13)$$

Jeśli dodatkowo spełnione są nierówności (13.12) i istnieje dodatni punkt krytyczny X_2 , to punkt ten jest lokalnie asymptotycznie stabilny.

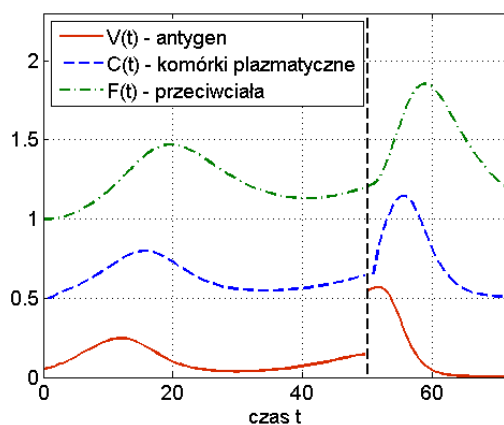
Twierdzenie 13.1 zadaje warunki wystarczające stabilności punktu X_2 przy ustalonych parametrach. Warunki te implikują, że hodograf Michajłowa ma kształt przedstawiony na rys. 13.2.

Widzimy więc, że układ (13.1) jest w stanie odzwierciedlić dwa podstawowe przebiegi choroby: zakażenie z wyzdrowieniem oraz chorobę przewlekłą (chroniczną), por. rys. 13.3. Ciekawą

własnością symulacji tego modelu jest to, że zwiększając dawkę antygeny możemy wyprowadzić organizm z choroby chronicznej — mówimy o wyzdrowieniu poprzez zaostrenie choroby, por. rys. 13.4. Trzeba jednak podkreślić, że taki efekt obserwuje się tylko w symulacjach, natomiast wszystkie wyniki analityczne wskazują na globalną stabilność punktu X_2 przy założeniu $m < m^*$.



Rysunek 13.3. Rozwiązania układu (13.1) ilustrujące dwa podstawowe przebiegi choroby: zakażenie z wyzdrowieniem (po prawej) oraz chorobę przewlekłą (po lewej).



Rysunek 13.4. Symulacja ilustrująca sytuację, w której poprzez zwiększenie dawki antygeny możemy wyprowadzić organizm z choroby przewlekłej.

14. Łańcuchy pokarmowe

Spróbujemy teraz zmierzyć się z problemem złożoności i stabilności ekosystemów w kontekście łańcuchów pokarmowych. Modelując łańcuchy pokarmowe skorzystamy z teorii grafów [6, 4]. Powszechnie uważa się, że bardziej stabilne są ekosystemy złożone, które na każdym poziomie troficznym mają wiele gatunków. Będziemy chcieli wyjaśnić te pojęcia ekologiczne w oparciu o podstawowe pojęcia teorii grafów. Wobec tego zaczniemy od wprowadzenia tych pojęć.

14.1. Podstawy teorii grafów

W teorii grafów wyróżniamy dwa podstawowe typy grafów — skierowane i nieskierowane. Ponieważ będziemy wykorzystywać oba te typy, więc zdefiniujemy je poniżej.

Definicja 14.1. Grafem nieskierowanym (krócej — grafem) G nazywamy parę (V, E) , gdzie zbiór $V = V(G)$ nazywamy zbiorem wierzchołków, natomiast zbiór $E = E(G)$ będący zbiorem nieuporządkowanych par elementów zbioru V nazywamy zbiorem krawędzi.

Definicja 14.2. Grafem skierowanym D nazywamy parę (V, A) , gdzie V jest zbiorem wierzchołków, natomiast A będący zbiorem par uporządkowanych elementów zbioru V nazywamy zbiorem krawędzi uporządkowanych.

Podkreślmy, że w teorii grafów nieskierowanych, którą będziemy się posługiwać, nie można połączyć dwóch wierzchołków więcej niż jedną krawędzią. Grafy nieskierowane z wieloma krawędziami opisuje teoria multigrafów. Natomiast w przypadku grafów skierowanych możemy mieć krawędź od stanu a do stanu b oraz od stanu b do stanu a . Z takimi grafami mamy do czynienia np. przy grafach przejścia łańcuchów Markowa, gdzie możemy z pewnym prawdopodobieństwem przejść ze stanu a do stanu b jak i ze stanu b do stanu a .

Grafy są często wykorzystywane w teorii transportu, więc wiele pojęć w tej teorii nosi nazwy związane z przemieszczeniem się.

Definicja 14.3. Ścieżką w D (odpowiednio w G) nazywamy ciąg wierzchołków wraz z łączącymi je krawędziami: $u_1, a_1, u_2, a_2, \dots, u_{t+1}, t \geq 0$, gdzie $a_i = (u_i, u_{i+1})$ (odpowiednio $a_i = \{u_i, u_{i+1}\}$) są krawędziami grafu D (odpowiednio G).

Definicja 14.4. Długością ścieżki nazywamy liczbę krawędzi t w tej ścieżce.

Definicja 14.5. Ścieżką prostą nazywamy ścieżkę, w której każda krawędź występuje tylko raz: $a_i \neq a_j$ dla $i \neq j$.

Definicja 14.6. Ścieżkę nazywamy zamkniętą, jeśli zaczyna się i kończy w tym samym wierzchołku: $u_1 = u_{t+1}$.

Definicja 14.7. Ścieżkę prostą zamkniętą nazywamy cyklem.

Zauważmy, że ponieważ długość ścieżki $t \geq 0$, to w szczególności przy $t = 0$ mamy ścieżkę u_1 , która ma długość 0. Natomiast nie ma innych ścieżek łączących wierzchołek z samym sobą.

Ponieważ ścieżka jest w jednoznaczny sposób wyznaczona przez występujące w niej krawędzie, więc będziemy w skrócie pisać u_1, u_2, \dots, u_{t+1} zamiast $u_1, a_1, u_2, \dots, u_{t+1}$.

Z istnieniem ścieżek łączy się też pojęcie osiągalności.

Definicja 14.8. Wierzchołek v jest osiągalny z u w grafie D (odpowiednio G), jeśli istnieje ścieżka o początku w u i końcu w v , czyli ścieżka $u = u_1, u_2, \dots, u_{t+1} = v$.

Oczywiście w grafie nieskierowanym, jeśli v jest osiągalne z u , to u jest osiągalne z v , natomiast w grafie skierowanym taka implikacja w ogólnym przypadku nie musi być prawdziwa. W związku z tym w przypadku grafu nieskierowanego możemy mówić o tym, że dwa wierzchołki komunikują się (lub są ze sobą połączone). Zauważmy też, że

Twierdzenie 14.1. *Jeśli w grafie D wierzchołek v jest osiągalny z u , to istnieje ścieżka prosta z u do v .*

Dowód. Jeśli $u = v$, to ścieżką prostą jest u . Jeśli $u \neq v$, to skoro v jest osiągalny z u , to istnieje co najmniej jedna ścieżka $u = u_1, u_2, \dots, u_{t+1} = v$, $t \geq 1$. Załóżmy, że nie jest to ścieżka prosta. Wtedy istnieją takie indeksy i, j , $i \neq j$, że $u_i = u_j$. Zbudujmy nową ścieżkę $u = u_1, \dots, u_i, u_j, \dots, u_{t+1} = v$, biegnącą z u do v , usuwając krawędzie od indeksu $i + 1$ do $j - 1$. Jeśli jest to ścieżka prosta, to kończymy dowód, jeśli nie, to powtarzamy poprzedni krok. Ponieważ liczba możliwych kroków jest skończona, to w którymś z kroków dostajemy ścieżkę prostą. \square

Korzystając z pojęcia długości ścieżki możemy teraz zdefiniować odległość dwóch wierzchołków.

Definicja 14.9. Jeśli wierzchołek v jest osiągalny z u , to odległością z u do v nazywamy długość najkrótszej ścieżki (czyli ścieżki o najmniejszej długości) z u do v .

Zauważmy, że tak zdefiniowana odległość w grafie skierowanym nie spełnia wszystkich aksjomatów metryki, ponieważ nie musi być symetryczna. Co więcej, jeśli wierzchołek v jest osiągalny z u , to nie oznacza, że u jest osiągalne z v , zatem odległość $d(u, v)$ może być zdefiniowana, natomiast odległość $d(v, u)$ — nie musi. Dla przykładu — w systemie komunikacyjnym miasta, jeśli wprowadzamy ulice jednokierunkowe, to musimy zadbać, żeby wszystkie istotne miejsca miały zdefiniowaną odległość.

Zauważmy dalej, że $d(u, u) = 0$, ponieważ założyliśmy, że u jest ścieżką. Odległość w grafie nieskierowanym spełnia też nierówność trójkąta.

Twierdzenie 14.2. *Jeśli w grafie D wierzchołek v jest osiągalny z u i w jest osiągalny z v , to zachodzi*

$$d(u, w) \leq d(u, v) + d(v, w).$$

Dowód. Z osiągalności wynika istnienie odpowiednich najkrótszych ścieżek $u = u_1, \dots, u_{m+1} = v$ długości m oraz $v = v_1, \dots, v_{n+1} = w$ długości n .

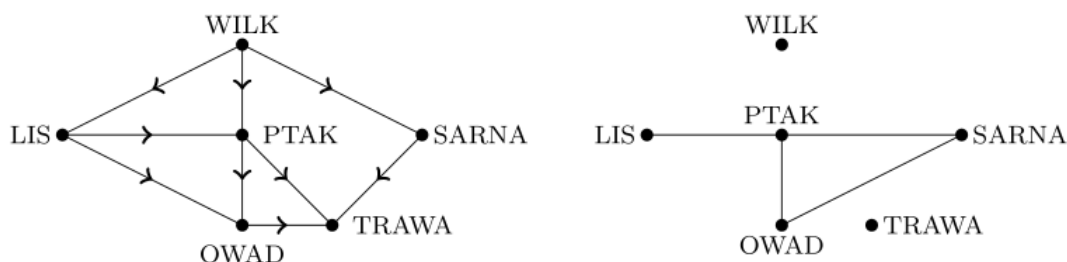
Wtedy ścieżka $u = u_1, \dots, u_{m+1}, v_1, \dots, v_{n+1}$ długości $m + n$ zaczyna się w u i kończy w w , ale nie musi to być ścieżka najkrótsza. Zatem

$$d(u, v) + d(v, w) = m + n \geq d(u, w).$$

\square

14.2. Łańcuchy pokarmowe

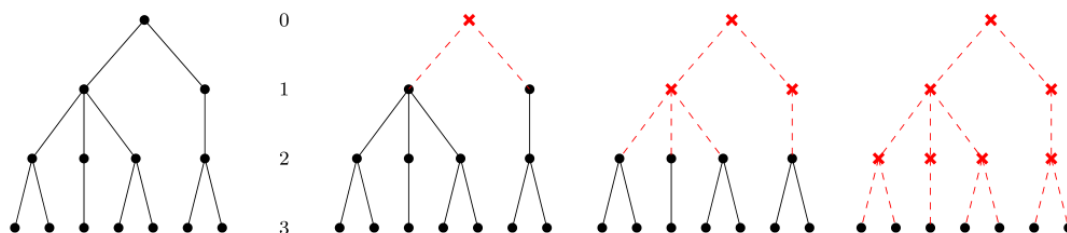
Teraz spróbujemy opisać ekosystem siedliska za pomocą łańcuchów pokarmowych. Siedliskiem nazwiemy obszar ograniczony i w pewien sposób wydzielony, w którym żyją różne gatunki roślin i zwierząt. Tworzą one *łańcuch pokarmowy*, który odzwierciedlamy za pomocą grafu skierowanego, w którym wierzchołkami są gatunki, natomiast krawędź skierowana prowadzi od gatunku u do v , jeśli u jest drapieżnikiem dla v . Z danym łańcuchem pokarmowym wiążemy też graf konkurencji, który z kolei stanowi graf nieskierowany z tymi samymi wierzchołkami, a krawędzie łączą dwa wierzchołki, jeśli odpowiadające im gatunki konkurują o to samo pożywienie.



Rysunek 14.1. Przykładowy łańcuch pokarmowy (po lewej) wraz z powiązaniem z nim grafem konkurencji (po prawej).

Łańcuchy pokarmowe możemy badać na różne sposoby. Przedstawimy tutaj koncepcję statusu troficznego (pokarmowego) zaadaptowaną z socjologicznej [teorii organizacji](#). W ekologii socjologiczne pojęcie statusu osobnika w schemacie organizacji odpowiada poziomowi (stopniowi) troficznemu gatunku. Chcemy w taki sposób odzwierciedlić złożoność i różnorodność łańcuchów pokarmowych. Jeśli na każdym poziomie troficznym w danym łańcuchu pokarmowym jest wiele gatunków, to powiemy, że taki łańcuch jest różnorodny. Wspominaliśmy już, że w ekologii uważa się, iż łańcuchy bardziej złożone są też bardziej stabilne. Chcemy więc móc określić co to znaczy, że dany łańcuch jest bardziej złożony.

Pojęcie poziomu troficznego dość łatwo zdefiniować w przypadku prostych łańcuchów pokarmowych opisywanych za pomocą specyficznych grafów zwanych drzewami. W strukturze drzewa występuje jeden wyróżniony wierzchołek zwany korzeniem, do którego nie prowadzą żadne krawędzie wchodzące, zatem nie ma też żadnych ścieżek prowadzących od jakiegokolwiek innego wierzchołka do korzenia. Struktura drzewa jest rekurencyjna — jeśli usuniemy korzeń wraz z prowadzącymi zeń krawędziami, to powstaną oddzielne grafy o analogicznej strukturze, których korzeniami są wierzchołki bezpośrednio osiągalne z pierwotnego korzenia, por. rys. 14.2.



Rysunek 14.2. Ilustracja wielokrotnego przeprowadzenia procedury usuwania korzenia grafu wraz z prowadzącymi z niego krawędziami.

W takiej rekurencyjnej strukturze, jeśli procedurę usuwania korzeni przeprowadzimy dostatecznie wiele razy, to na końcu dostaniemy pojedyncze wierzchołki (czasem mówi się o liściach drzewa). Liczba usunięć przeprowadzająca dany wierzchołek na korzeń nowego grafu jest ściśle związana z poziomem troficznym gatunku reprezentowanego przez ten wierzchołek w drzewie. Jeśli maksymalna liczba możliwych usunięć w danym drzewie wynosi m , a liczba usunięć, która tworzy z danego wierzchołka korzeń — k , to poziom troficzny gatunku odpowiadającego temu wierzchołkowi równa się $m - k$.

W ogólnym przypadku dla złożonego łańcucha pokarmowego znacznie trudniej zdefiniować analog tego pojęcia. Zdefiniujemy miarę statusu troficznego w celu odróżnienia od „zwykłego” poziomu troficznego, łatwego do wyznaczenia w prostych strukturach. Zanim przejdziemy do zaproponowania takiej definicji wprowadzimy odpowiednie założenia. Zakładamy, że graf opisujący łańcuch pokarmowy jest acykliczny, czyli wśród ścieżek tego grafu nie da się wyróżnić cyklu. W szczególności nie ma dwóch gatunków u i v , takich że jednocześnie u jest drapieżnikiem dla v i v jest drapieżnikiem dla u . Dla takich grafów sformułujemy system aksjomatów, którym powinna podlegać miara statusu troficznego. Zauważmy, że w zależności od liczby przyjętych aksjomatów odpowiadająca im miara może być jednoznaczna, niejednoznaczna (przy mniejszej liczbie aksjomatów), a nawet może nie istnieć, jeśli aksjomatów jest zbyt dużo. Odpowiednia liczba aksjomatów z jednej strony prowadzi do barku możliwości wyboru i implikuje jednoznaczność miary, z drugiej — gwarantuje jej istnienie. Jeśli dany układ aksjomatów jest zbyt obszerny i odpowiadająca mu miara nie istnieje, to należy się zastanowić, które z aksjomatów należy usunąć, aby zagwarantować istnienie i jednocześnie nie usunąć zbyt wielu aksjomatów.

Wprowadzając system aksjomatów analogiczny do systemu aksjomatów dla struktury organizacji posłużymy się pojęciami bezpośredniego i pośredniego drapieżnika (analogicznie bezpośredniej i pośredniej ofiary). Powiemy, że gatunek u jest bezpośrednim drapieżnikiem dla v jeśli istnieje krawędź z u do v , natomiast jest pośrednim drapieżnikiem, jeśli v jest osiągalne z u . Będziemy mówić, że u jest drapieżnikiem dla v , jeśli u albo bezpośrednim, albo pośrednim drapieżnikiem dla v . Analogicznie v jest wtedy ofiarą u (bezpośrednią/pośrednią). Zauważmy, że w dowolnym łańcuchu pokarmowym możliwa jest sytuacja, gdy ani u nie jest drapieżnikiem dla v , ani v nie jest drapieżnikiem dla u .

Jeśli u jest drapieżnikiem dla v , to zdefiniujemy

Definicja 14.10. Stopień troficzny u w stosunku do v równy odległości z u do v .

Zauważmy, że powyższa definicja ma sens, gdyż v jest w tym przypadku osiągalne z u , co znaczy, że istnieje najkrótsza ścieżka z u do v , długość której wyznacza odległość z u do v .

Niech $t_D(u)$ oznacza miarę statusu troficznego gatunku u w łańcuchu pokarmowym D . Spełnia ona następujące aksjomaty

1. Jeśli gatunek u w łańcuchu D nie jest drapieżnikiem dla żadnego innego gatunku, to $t_D(u) = 0$.
2. Na bazie łańcucha D zbudujmy łańcuch \tilde{D} dodając nowy wierzchołek — gatunek, dla którego u jest bezpośrednim drapieżnikiem. Wtedy $t_{\tilde{D}}(u) > t_D(u)$.
3. Na bazie łańcucha D zbudujmy łańcuch \tilde{D} dodając wierzchołki lub krawędzie w taki sposób, że dla pewnych gatunków będących drapieżnikami wyróżnionego gatunku u ich status troficzny względem u rośnie, nie ma natomiast takiego gatunku v będącego ofiarą u , dla którego by zmalał. Wtedy $t_{\tilde{D}}(u) > t_D(u)$.

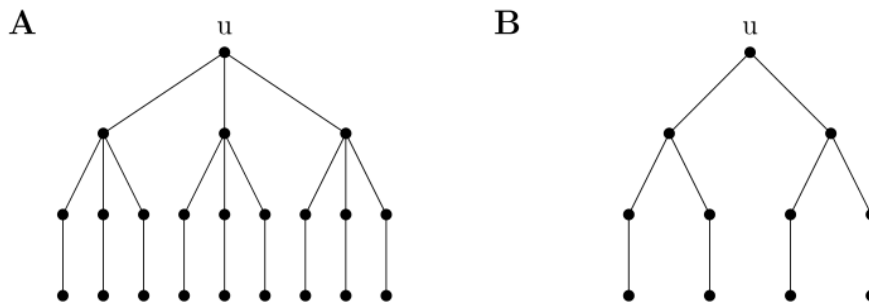
W strukturze organizacji aksjomat I oznacza, że jeśli dany pracownik nie ma żadnych podwładnych, to jego status jest równy 0. W aksjomacie II pracownikowi dajemy nowego bezpośredniego podwładnego, zatem status tego pracownika w organizacji wzrasta. Aksjomat III odzwierciedla sytuację awansu.

Wykażemy teraz istnienie miary spełniającej aksjomaty I – III. Jest to dowód konstrukcyjny. Zdefiniujemy

$$h_D(u) = \sum_k kn_k, \quad (14.1)$$

gdzie n_k jest liczbą gatunków na poziomie troficznym k niżej względem u .

Zauważmy, że tak zdefiniowana miara różni się istotnie od proponowanej wcześniej miary reprezentowanej przez poziom troficzny dla drzew. Na rys. 14.3 możemy porównać te dwa pojęcia — korzeń u będący na tym samym poziomie troficznym w obu drzewach ma różny status troficzny ze względu na zróżnicowanie w złożoności tych łańcuchów pokarmowych.



Rysunek 14.3. Drzewa ilustrujące fakt, że korzeń u będący na tym samym poziomie troficznym w obu drzewach ma różny status troficzny ze względu na zróżnicowanie złożoności tych łańcuchów pokarmowych.

Zauważmy dalej, że $h_D(u) \geq 0$, co nie zostało założone w aksjomatach I – III. Jeśli do tych aksjomatów dołożymy postulat nieujemności, to miara h_D jest minimalna w następującym sensie

Twierdzenie 14.3. *Miara statusu troficznego h_D określona wzorem (14.1) ma następujące własności*

1. *spełnia aksjomaty I – III;*
2. *jeśli t_D jest inną miarą o wartościach naturalnych spełniającą aksjomaty I – III, to dla dowolnego u zachodzi $t_D(u) \geq h_D(u)$;*

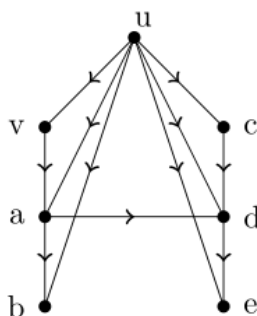
Dowód. Postać wzoru (14.1) implikuje, że aksjomaty I – III są spełnione w sposób oczywisty. W celu udowodnienia minimalności skorzystamy z następującego lematu

Lemat 14.1. *W grafie skierowanym acyklicznym $D = (V, A)$ dla dowolnego wierzchołka $u \in V$ istnieje taki $v \in V$, który jest osiągalny z u i nie ma krawędzi wychodzących.*

Resztę dowodu przeprowadzimy indukcyjnie ze względu na $m = h_D(u)$. Jeśli $m = 0$, to aksjomat I implikuje, że u nie ma żadnej ofiary, więc $t_D(u) = h_D(u) = 0$. Niech dla danego łańcucha pokarmowego $h_D(u) = m$ i jednocześnie $t_D(u) \geq h_D(u)$. Załóżmy dalej, że mamy łańcuch o mierze $h_D(u) = m + 1$. Zatem w D istnieje co najmniej jeden gatunek, dla którego u jest drapieżnikiem. Zgodnie z lematem istnieje w D wierzchołek v , z którego nie ma wychodzących krawędzi, a który jest osiągalny z u . Zbudujemy nowy łańcuch D' na bazie D w następujący sposób. Jeśli v jest bezpośrednią ofiarą u , to usuwamy v , natomiast jeśli v jest ofiarą pośrednią, to istnieje ścieżka $u = u_1, u_2, \dots, u_{k+1} = v$ i usuwamy krawędź (u_k, u_{k+1}) (wraz z wierzchołkiem

v). Wtedy status u względem v zmniejszy się o 1, natomiast względem wszystkich innych się nie zmienia. Stąd $h_{D'}(u) = m$. Z założenia indukcyjnego mamy $t_D(u) \geq m$, ale aksjomaty II i III implikują, że $t_D(u) > t_{D'}(u)$. Zatem zakładając, że miara jest liczbą całkowitą, otrzymujemy $t_D(u) \geq m + 1 = h_D(u)$. \square

Miara h_D ma pewne niepożądane cechy. W szczególności mierząc status troficzny chcielibyśmy mieć następującą własność — jeśli t_D jest miarą statusu troficznego i u jest drapieżnikiem dla v , to chcielibyśmy by $t_D(u) > t_D(v)$. Niestety h_D nie ma tej własności, por. rys. 14.4, gdzie u jest bezpośrednim drapieżnikiem dla v i jednocześnie $h_D(u) = 6 < h_D(v) = 8$.



Rysunek 14.4. Graf ilustrujący niepożądaną cechę miary h_D . Widzimy, że u jest bezpośrednim drapieżnikiem dla v i jednocześnie $h_D(u) = 6 < h_D(v) = 8$.

W celu rozwiązania tego problemu musimy przededefiniować poziom troficzny gatunku u względem gatunku v w łańcuchu pokarmowym. Do tej pory interpretowaliśmy to pojęcie w kontekście odległości od u do v , czyli długości najkrótszej ścieżki, natomiast teraz przeddefiniujemy je jako maksymalną długość ścieżki prostej prowadzącej z u do v . Konstruując miarę statusu troficznego w oparciu o to pojęcie dostaniemy funkcję spełniającą aksjomaty I – III i dodatkowo aksjomat

IV. Jeśli gatunek u jest bezpośrednim/pośrednim drapieżnikiem dla v , to $t_D(u) > t_D(v)$, gdzie t_D jest miarą statusu troficznego gatunku u .

Twierdzenie 14.4. Zdefiniujemy miarę statusu troficznego u względem v jako długość najdłuższej ścieżki prostej z v do u . Wtedy miara h zdefiniowana wzorem (14.1) spełnia aksjomaty I – IV.

Dowód. Dowód tego faktu pozostawiamy jako ćwiczenie. \square

Wracając do łańcucha pokarmowego z ostatniego rysunku obliczamy teraz $h_D(u) = 14 > h_D(v) = 8$.

Miara statusu troficznego pozwala w pewnym sensie zbadać ważność danego gatunku w łańcuchu pokarmowym. Istnieje także inny sposób analizowania roli gatunku w ekosystemie i podatności ekosystemu na zmiany. Opiera się on na pojęciu spójności grafu.

Definicja 14.11. Graf skierowany $D = (V, A)$ nazwiemy

- silnie spójnym, jeśli dla dowolnych dwóch wierzchołków $u, v \in V$ zarówno u jest osiągalny z v jak i v jest osiągalny z u ;
- spójnym, jeśli dla dowolnej pary wierzchołków $u, v \in V$, albo u jest osiągalny z v , albo v jest osiągalny z u ;

- słabo spójnym, jeśli dla dowolnej pary wierzchołków $u, v \in V$ istnieje pseudościeżka między u i v , czyli ciąg wierzchołków i krawędzi $u = u_1, a_1, \dots, a_{n-1}, u_n = v$, takich że albo $a_i = (u_i, u_{i+1}) \in A$, albo $a_i = (u_{i+1}, u_i) \in A$;
- niespójnym, jeśli nie jest słabo spójny.

Oczywiście dla grafu nieskierowanego G wszystkie trzy typy spójności pokrywają się, ponieważ ścieżka jest tożsama z pseudościeżką. Wobec tego graf nieskierowany $G = (V, E)$ jest spójny wtedy i tylko wtedy gdy dla każdej pary wierzchołków $u, v \in V$ istnieje łącząca je ścieżka.

Z podanymi rodzajami spójności wiążemy pojęcie kategorii spójności grafu.

Definicja 14.12. Graf skierowany D ma kategorię spójności

- 0, jeśli jest niespójny;
- 1, jeśli jest słabo spójny i nie jest spójny;
- 2, jeśli jest spójny i nie jest silnie spójny;
- 3, jeśli jest silnie spójny.

Okazuje się, że poszczególne kategorie spójności są równoważne istnieniu tzw. ścieżek kompletnych.

Definicja 14.13. Ścieżką kompletną w grafie D (odpowiednio G) nazwiemy taką ścieżkę, która przechodzi przez wszystkie wierzchołki grafu D (G), czyli u_1, \dots, u_n jest ścieżką kompletną, jeśli $\forall v \in V \exists i \in \{1, \dots, n\}$, takie że $v = u_i$.

Można udowodnić następujące twierdzenie (dowód pomijamy).

Twierdzenie 14.5. Graf D jest

- silnie spójny \iff istnieje kompletna ścieżka zamknięta;
- spójny \iff istnieje ścieżka kompletna;
- słabo spójny \iff istnieje kompletna pseudościeżka.

Na bazie powyższych definicji wprowadzimy kategoryzację wierzchołków.

Definicja 14.14. Wierzchołek $u \in V$ w grafie $D = (V, A)$ nazwiemy wierzchołkiem kategorii (i, j) , $i, j \in \{0, 1, 2, 3\}$, gdy graf D ma kategorię spójności i , a po usunięciu wierzchołka u ma kategorię j .

Dzięki temu pojęciu możemy klasyfikować gatunki w łańcuchu pokarmowym. Za krytyczne w danym ekosystemie uznamy takie gatunki, dla których różnica $i - j$ jest największa, czyli takie, których usunięcie powoduje największą zmianę spójności.

Na koniec omówimy pojęcie wrażliwości na zmiany w danym łańcuchu pokarmowym.

Definicja 14.15. Graf D nazwiemy podatnym krawędziowo (wierzchołkowo), jeśli po usunięciu pewnej liczby krawędzi (wierzchołków) zmniejsza się kategoria spójności grafu. Stopniem podatności krawędziowej (wierzchołkowej) nazwiemy minimalną liczbę krawędzi (wierzchołków), usunięcie których powoduje zmianę kategorii spójności grafu.

Zauważmy, że im większy stopień podatności, tym graf jest mniej wrażliwy na zmiany.

Podsumowując, dla danego łańcucha pokarmowego jego złożoność i stabilność będzie wyrażać się jak największą liczbą gatunków w ekosystemie z jak największą miarą statusu troficznego i jak największym stopniem podatności.

15. Podstawy teorii gier

Na koniec omówimy jeszcze podstawowe pojęcia związane z [teorią gier](#) i w ich języku spróbujemy wyjaśnić pewne zachowania osobników, które na pozór wydają się paradoksalne, ale mają głębokie uzasadnienie z punktu widzenia przetrwania gatunku. Zauważmy, że różne gatunki wykształciły różne typy zachowań. Zdarza się, że osobniki danego gatunku prawie nigdy w trakcie rywalizacji — czy o partnera, czy o pożywienie — nie walczą „na poważnie”, a tylko straszą się wzajemnie w trakcie tzw. walk rytualnych. W przypadku innych — dochodzi do bardzo poważnych potyczek. Analogiczne zachowania mogą być dyskutowane na gruncie socjologicznym w stosunku np. do pracowników danego przedsiębiorstwa czy innych społeczności. Pojawia się zatem pytanie, kiedy osobnikom bardziej opłaca się stosować strategię walki, a kiedy uniku i od czego taki wybór zależy. Odpowiedzi możemy poszukiwać na gruncie teorii gier w oparciu o [strategie ewolucyjnie stabilne](#).

15.1. Podstawy teorii gier

Teoria gier w nierozzerwalny sposób łączy się z dwoma nazwiskami — [Johna Maynarda Smitha](#) [15] i [Johna Nasha](#). Pierwszy z nich wprowadził pojęcie strategii ewolucyjnie stabilnej, natomiast drugi sprecyzował matematyczne podstawy tej teorii i zaproponował pojęcie [równowagi Nasha](#).

W teorii gier zakładamy, że co najmniej dwóch osobników „gra” w pewną grę. Mają do wyboru pewne zachowania, które nazywamy strategiami. W zależności od tego jakie strategie wybiorą poszczególni gracze, dostają odpowiednie wypłaty. Zatem budując model w oparciu o teorię gier musimy zaproponować zbiór graczy, zbiór strategii oraz macierz wypłat związanych z poszczególnymi strategiami.

15.1.1. Gra w postaci normalnej

Załóżmy, że każdy z graczy ma do wyboru jedną z M strategii R_1, \dots, R_M , które nazwiemy strategiami czystymi. Jeśli założymy, że gracz może wybierać te strategie z pewnymi prawdopodobieństwami p_1, \dots, p_M , to taki profil strategii (p_1, \dots, p_M) , $p_i \geq 0$, $\sum_{i=1}^M p_i = 1$, nazywamy strategią mieszaną. Strategie mieszane identyfikujemy z punktami sympleksu

$$S_M = \{p \in \mathbb{R}^M : p_i \in [0, 1], \sum_{i=1}^M p_i = 1\}.$$

Wierzchołki sympleksu S_M są oczywiście strategiami czystymi.

Niech w grze bierze udział dwóch graczy i niech u_{ij} oznacza wypłatę gracza I jeśli gra on strategią R_i przeciwko graczowi II grającemu strategią R_j . Macierz $U = (u_{ij})_{i,j=1}^M$ nazywamy macierzą wypłat.

Definicja 15.1. Gra w postaci normalnej dla dwóch graczy zadana jest za pomocą zbioru strategii czystych R_1, \dots, R_M , sympleksu strategii mieszanych S_M oraz macierzy wypłat U .

Załóżmy teraz, że gracz I gra strategią mieszaną $p = (p_1, \dots, p_M)$ przeciwko graczowi II grającemu strategią mieszaną $q = (q_1, \dots, q_M)$. Ile wyniesie średnia wypłata gracza I? Gdyby gracz I grał czystą strategią R_i przeciwko q , to z prawdopodobieństwami q_j dostawałby wypłatę u_{ij} , $j = 1, \dots, M$, czyli średnio $(Uq^T)_i = \sum_{j=1}^M u_{ij}q_j$. Wobec tego grając p przeciw q dostaje z prawdopodobieństwami p_i wypłatę $(Uq^T)_i$, zatem średnio $pUq^T = \sum_{i,j=1}^M p_i u_{ij} q_j$.

Definicja 15.2. Funkcją odpowiedzi na strategię q nazwiemy funkcję średnich wypłat $F_q(p) = pUq^T$.

Niech $\beta(q)$ oznacza zbiór najlepszych odpowiedzi na strategię q , czyli zbiór tych p , dla których $F_q(p)$ osiąga wartość maksymalną.

Definicja 15.3. Równowagą Nasha nazwiemy taką strategią q , która jest najlepszą odpowiedzią na siebie samą, czyli $q \in \beta(q)$. Ścisła równowaga Nasha jest jedyną taką strategią, czyli $\{q\} = \beta(q)$.

Widzimy więc, że dla równowagi Nasha zachodzi $pUq^T \leq qUq^T$ dla dowolnej strategii p , natomiast dla ścisłej równowagi Nasha mamy $pUq^T < qUq^T$ dla dowolnej strategii $p \neq q$.

Możemy dość łatwo wykazać, że ścisła równowaga Nasha jest strategią czystą.

Twierdzenie 15.1. Niech q będzie równowagą Nasha. Jeśli q jest ścisłą równowagą Nasha, to istnieje i , dla którego $q = R_i$.

Dowód. Załóżmy, że dla wszystkich $j \in \{1, \dots, M\}$ mamy $q \neq R_j$. Oczywiście dla dowolnej strategii p zachodzi nierówność $F_q(p) < F_q(q)$, a w szczególności

$$F_q(R_j) < F_q(q), \quad j = 1, \dots, M. \quad (15.1)$$

Pomnóżmy stronami nierówność (15.1) przez q_j i zsumujmy po wszystkich $j = 1, \dots, M$. Dostaniemy $\sum_{j=1}^M q_j F_q(R_j) < \sum_{j=1}^M q_j F_q(q)$. Zauważmy, że $q = \sum_{j=1}^M q_j R_j$, zatem z liniowości F_q otrzymujemy $F_q(q) < F_q(q)$, czyli q nie może być strategią czystą. \square

Zauważmy, że gdy q jest ścisłą równowagą Nasha, to jeśli w populacji wszyscy stosują tę strategię, to osobnik, który będzie próbował zastosować inną — przegra, gdyż jego wypłata zawsze będzie mniejsza. Taka dynamika gry wiąże się też z koncepcją strategii ewolucyjnie stabilnej, czyli strategii odpornej na niewielkie zaburzenia.

Definicja 15.4. Strategię q nazwiemy ewolucyjnie stabilną (ESS, czyli *evolutionary stable strategie*), jeśli dla wszystkich $p \in S_M$, $p \neq q$ zachodzi nierówność

$$pU(\varepsilon p + (1 - \varepsilon)q)^T < qU(\varepsilon p + (1 - \varepsilon)q)^T$$

dla wszystkich dostatecznie małych $\varepsilon < \tilde{\varepsilon}(p)$.

Stałą $\tilde{\varepsilon}$ nazywamy barierą inwazji. Oczywiście ścisła równowaga Nasha jest strategią ewolucyjnie stabilną. Z drugiej strony

Twierdzenie 15.2. Strategia $q \in S_M$ jest ESS wtv spełnione są następujące warunki

1. równowagi: $pUq^T \leq qUq^T$ dla wszystkich $p \in S_M$;
2. stabilności: jeśli dla pewnego $p \neq q$ zachodzi $pUq^T = qUq^T$, to $pUp^T < qUp^T$.

Dowód tego twierdzenia pozostawiamy jako ćwiczenie.

15.2. Przykłady gier

Omówimy teraz dwa podstawowe przykłady gier i zinterpretujemy ich wyniki.

15.2.1. Gra jastrzęb – gołąb

Gra jastrzęb – gołąb polega na tym, że gracz ma do wyboru strategię agresji, którą stosując zawsze walczy z przeciwnikiem, albo strategię wycofywania się, w której zawsze ustępuje przeciwnikowi. W trakcie walki osobnik może być poważnie zraniony, więc może wiele stracić, ale z drugiej strony może dużo zyskać, jeśli wygra. Z kolei osobnik ustępujący nie traci w trakcie spotkania z przeciwnikiem. Miarą wygranej w tej grze jest dostosowanie (ang. *fitness*), często wyrażane jako liczba potomstwa. Niech osobnik, który wygrywa, dostaje g , natomiast osobnik zraniony traci c . Wtedy jeśli spotyka się jastrzęb z jastrzębiem, to z prawdopodobieństwem $1/2$ może zyskać g i stracić c , więc średnia wygrana to $(g - c)/2$. Wygrana jastrzębia przeciw gołębiowi to oczywiście g . Gołąb w starciu z jastrzębiem wycofuje się, więc jego wygrana to 0 , a w starciu z gołębiem może wygrać z prawdopodobieństwem $1/2$, więc wygrana to $g/2$. Gra jest symetryczna, zatem wystarczy podać jedynie macierz wypłat gracza I przeciwko II. Często mówi się o graczu „kolumnowym” i „wierszowym”. Jeśli przez R_1 oznaczymy strategię jastrzębia, a przez R_2 — gołębia, to macierz wypłat ma postać

$$\begin{pmatrix} \frac{g-c}{2} & g \\ 0 & g/2 \end{pmatrix}.$$

W grze tej dodatkowo zakładamy, że rany otrzymywane w walce są na tyle dotkliwe, że $g - c < 0$.

Sprawdzimy, że gra ta nie ma ścisłej równowagi Nasha. Wiemy, że taka równowaga musiałaby odpowiadać albo czystej strategii jastrzębia, albo czystej strategii gołębia. Policzmy dla obu strategii funkcje najlepszej odpowiedzi.

— Biorąc $R_1 = (1, 0)$ dostajemy

$$F_{R_1}(p) = pUR_1^T = (p_1, 1 - p_1) \begin{pmatrix} \frac{g-c}{2} & g \\ 0 & \frac{g}{2} \end{pmatrix} (1, 0)^T = p_1 \frac{g-c}{2}$$

i maksymalizując to wyrażenie dostajemy $p_1 = 0$ (tu widzimy, do czego potrzebne jest założenie $g - c < 0$), czyli $R_1 \notin \beta(R_1)$.

— Biorąc $R_2 = (0, 1)$ dostajemy

$$F_{R_2}(p) = pUR_2^T = (p_1, 1 - p_1) \begin{pmatrix} \frac{g-c}{2} & g \\ 0 & \frac{g}{2} \end{pmatrix} (0, 1)^T = \frac{(p_1 + 1)g}{2}$$

i maksymalizując to wyrażenie dostajemy $p_1 = 1$, czyli $R_2 \notin \beta(R_2)$.

Widzimy więc, że w grze tej nie ma ścisłej równowagi Nasha. Staje się to oczywiste, jeśli wyobrazimy sobie populację składającą się albo z samych jastrzębi, albo z samych gołębi. Jeśli są same jastrzębie, to pojawiający się gołąb ma lepiej, gdyż w starciu z jastrzębiem nic nie traci, zatem gołębie zaczynają się rozprzestrzeniać. Z kolei jeśli są same gołębie, to pojawiający się jastrząb zawsze wygrywa, więc zaczął rozprzestrzeniać się jastrzębie. Wydaje się więc, że powinna być jakaś odpowiednia proporcja między gołębiami i jastrzębiami.

Rozpatrzmy strategię mieszaną $(x, 1 - x)$, w której z prawdopodobieństwem x osobnik gra strategię jastrzębia i z prawdopodobieństwem $1 - x$ gra gołębia. Wtedy średni wzrost dostosowania jastrzębia wyniesie

$$x \frac{g - c}{2} + (1 - x)g = g - x \frac{g + c}{2},$$

natomiast gołębia

$$(1 - x) \frac{g}{2}.$$

Dopóki któreś z tych dostosowań jest większe, odpowiadająca mu strategia będzie się rozprzestrzeniać. Równowaga ustali się, jeśli oba będą jednakowe, czyli

$$g - x \frac{g + c}{2} = (1 - x) \frac{g}{2} \iff x = \frac{g}{c}.$$

Zatem proporcja jastrzębi będzie tym mniejsza, im więcej można stracić w walce w stosunku do zysku. Ta analiza wyjaśnia też popularność walk rytualnych wśród zwierząt, które mogą się poważnie zranić w trakcie prawdziwej walki — zamiast faktycznie walczyć zwierzęta te wolą w walce rytualnej zaprezentować przed przeciwnikiem swoją siłę licząc na to, że przeciwnik się zniechęci i odejdzie bez prawdziwej walki. W taki sposób postępują np. jelenie na rykowisku.

15.2.2. Dylemat więźnia

Dylemat więźnia jest typową grą socjologiczną. Zakłada się w niej, że dwóch więźniów zostało osadzonych w więzieniu w osobnych celach i każdemu z nich prokurator obiecał zmniejszenie wyroku za wydanie współwinnego. Jeśli jednak żaden z nich się nie przyzna, to niewiele będzie im można udowodnić, więc obaj dostaną najwyżej niewielki wyrok. Z kolei jeśli jeden z nich zdradzi, a drugi nie, to ten, który wydał współwinnego dostanie wyrok w zawieszeniu. Jeśli obaj zdradzą się wzajemnie, obaj dostaną nieco mniejsze wyroki. Niech macierz wypłat w tej grze przy założeniu, że R_1 oznacza zdradę kolegi, a R_2 — odrzucenie propozycji prokuratora, ma postać

$$\begin{pmatrix} a & 0 \\ b & c, \end{pmatrix}$$

gdzie a to wyrok dla więźnia I przy założeniu, że obaj zdradzają się wzajemnie, b — wyrok dla tego więźnia, gdy on odrzuca propozycję prokuratora, a więzień II go zdradza, więc $b > a$, natomiast c to wyrok w sytuacji, gdy obaj odrzucają współpracę z prokuratorem, więc $a > c$.

Zauważmy, że obu więźniom najbardziej opłaca się odrzucić współpracę z prokuratorem, gdyż dzięki temu zapewniają sobie mniejszy wyrok, ale z obawy przed zdradą kolegi wolą sami też zdradzić, bo jeśli kolega zdradzi, a dany więzień nie, to on dostanie cięższy wyrok, niż jeśliby też zdradził. Zatem R_1 jest tu czystą równowagą Nasha. Faktycznie grając R_1 przeciw R_1 mamy wypłatę

$$(1, 0) \begin{pmatrix} a & 0 \\ b & c, \end{pmatrix} (1, 0)^T = a,$$

natomiast grając przeciw dowolnej strategii $(x, 1 - x)$

$$(1, 0) \begin{pmatrix} a & 0 \\ b & c \end{pmatrix} (x, 1 - x)^T = ax$$

i oczywiście $a > ax$ dla $x \neq 1$. Wobec tego R_1 jest ścisłą równowagą Nasha i jednocześnie strategią ewolucyjnie stabilną. Zauważmy jednak, że strategii tej brakuje jednej z podstawowych cech — nie jest optymalna, w tym sensie, że jeśli obaj wybiorą R_2 , dostaną mniejsze wyroki.

Tego typu zagadnienia możemy rozważać badając np. [optymalność w sensie Pareto](#).

Definicja 15.5. Powiemy, że wynik gry jest nieoptymalny w sensie Pareto, jeżeli istnieje inny wynik, dający albo obu graczom wyższe wypłaty, albo jednemu z nich taką samą, a drugiemu wyższą. W przeciwnym przypadku wynik jest optymalny.

Pareto postulował następujące kryterium: tylko wynik optymalny w sensie Pareto może być akceptowany jako rozwiązanie gry. Kryterium Pareto stanowi podstawową zasadą racjonalności zbiorowej. W dylemacie więźnia wchodzi ona w konflikt z zasadą racjonalności indywidualnej reprezentowanej przez równowagę Nasha.

Literatura

- [1] D. Bobrowski. *Ciągi losowe*. Wyd. Naukowe UAM, 2002.
- [2] H. Caswell. *Matrix population models: construction, analysis, and interpretation*. Sunderland, Sinauer Associates, 2001.
- [3] R.A. Fisher. The wave of advance of advantageous genes. *Annals of Eugenics*, 7:355–369, 1937.
- [4] F. Harary. *Graph theory*. Addison-Wesley Publishing Co., 1969.
- [5] C.S. Holling. Some characteristics of simple types of predation and parasitism. *Canadian entomologist*, 91(7):385–398, 1959.
- [6] U. Jupowieckia-Mieszala, E. Krzywiecka. *Wybrane zagadnienia teorii grafów i ich zastosowań*. Katowice, AE im. Karola Adamieckiego, 1997.
- [7] J.G. Kemeny, J.L. Snell. *Finite markov chains*. Springer, 1976.
- [8] P.H. Leslie. On the use of matrices in certain population mathematics. *Biometrika*, 35:183–212, 1945.
- [9] P.H. Leslie. Some further notes on the use of matrices in population mathematics. *Biometrika*, 35(3-4):213–245, 1948.
- [10] A.J. Lotka. *Elements of physical biology*. Williams & Wilkins Co., 1925.
- [11] G.I. Marchuk. *Matematicheskie modeli v immunologii*. Nauka, 1980.
- [12] G.I. Marchuk. *Mathematical Models in Immunology*. Springer-Verlag, New York, 1983.
- [13] G.I. Marczuk. *Modele matematyczne w immunologii*. PWN, 1989.
- [14] A. Palczewski. *Równania Różniczkowe zwyczajne. Teoria i metody numeryczne z wykorzystaniem komputerowego systemu obliczeń symbolicznych*. WNT, 1999.
- [15] J.M. Smith. *Evolution and the Theory of Games*. Cambridge University Press, 1982.
- [16] P.F. Verhulst. Notice sur la loi que la population suit dans son accroissement. *Corr. Math. et Phys*, 10:113–121, 1838.
- [17] V. Volterra. Variazione e fluttuazione del numero d'individui animali conventi. *Memoires de l'Academic Nationale Lincei*, 2:31–113, 1926.