

The background of the cover is a photograph of a mountainous landscape. In the foreground, a steep, green hillside slopes down from the right, dotted with dark, rocky patches. At the bottom right, a small figure of a person is visible on a path. The middle ground shows a deep valley with more green slopes. In the background, several layers of mountain ranges are visible, with the highest peaks covered in snow and partially shrouded in a light mist or haze. The sky is a pale, clear blue.

Iwona Karcz-Dulęba

Dynamika adaptacji ewolucyjnych metod fenotypowych

Iwona KAR CZ-DULĘBA

Dynamics of adaptation
of phenotypic evolution methods

Iwona Karcz-Dulęba

Dynamika adaptacji ewolucyjnych metod fenotypowych



Oficyna Wydawnicza Politechniki Wrocławskiej · Wrocław 2008

Recenzenci

Jarosław Arabas
Andrzej Obuchowicz

Opracowanie redakcyjne

Maria Izbicka

Korekta

Alina Kaczak

Projekt okładki

Zofia i Dariusz Godlewscy

Wszelkie prawa zastrzeżone. Żadna część niniejszej książki, zarówno w całości, jak i we fragmentach, nie może być reprodukowana w sposób elektroniczny, fotograficzny i inny bez pisemnej zgody wydawcy i właściciela praw autorskich.

© Copyright by Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 2008

OFICyna WYDAWNICZA POLITECHNIKI WROCLAWSKIEJ
Wybrzeże Wyspiańskiego 27, 50-370 Wrocław
<http://www.oficyna.pwr.wroc.pl>; e-mail: oficwyd@pwr.wroc.pl

ISSN 0324-9786

Drukarnia Oficyny Wydawniczej Politechniki Wrocławskiej. Zam. nr 603/2008.

*optymalizacja, fenotypowe metody ewolucyjne,
adaptacja lokalna i globalna,
analiza teoretyczna,
nieskończone i bardzo małe populacje*

Iwona KARCZ-DULEBA*

DYNAMIKA ADAPTACJI EWOLUCYJNYCH METOD FENOTYPOWYCH

W dysertacji przedstawiono rezultaty analizy teoretycznej metody poszukiwań ewolucyjnych z miękką selekcją. Metoda bazuje na prostym modelu darwinowskiej ewolucji fenotypowej działającej z selekcją proporcjonalną i mutacją gaussowską (z rozkładem normalnym). Połączenie obu mechanizmów reprodukcji powoduje, że ewoluujące populacje są w stanie eksplorować nieograniczone wielowymiarowe rzeczywiste przestrzenie w poszukiwaniu wyższych przystosowań funkcji celu. Algorytmy optymalizacyjne implementujące metodę mają pożądane w optymalizacji właściwości: zdolność przekraczania siodła pomiędzy optimami funkcji celu, a tym samym znajdowania optimów globalnych, niewielki spadek efektywności wraz ze wzrostem wymiarowości przestrzeni poszukiwań oraz małą wrażliwość na niedokładności i zaszumienia w odczycie wartości funkcji celu. Zaprezentowano analizę teoretyczną dynamiki ewoluujących populacji dla skrajnych liczebności populacji: nieskończonych i dwuelementowych, w jednowymiarowych rzeczywistych przestrzeniach poszukiwań bez ograniczeń. Wyniki teoretyczne uzyskano, jak do tej pory, tylko przy tych upraszczających założeniach, ponieważ metody ewolucyjne są skomplikowanymi nieliniowymi procesami stochastycznymi. Badano zarówno dynamikę lokalną – przy poszukiwaniu jedyne optimum jednomodalnych funkcji celu, jak i globalną – gdy funkcja celu była wielomodalna. Dla jasności wyводу w przypadku dynamiki globalnej ograniczono się głównie do funkcji dwumodalnych i analizy mechanizmu przejścia siodła pomiędzy optimami, aczkolwiek niektóre badania rozszerzono na funkcje o więcej niż dwóch optimach.

Nieskończone populacje scharakteryzowano za pomocą rozkładów cech i opisano przez funkcje gęstości prawdopodobieństwa rozkładu osobników z daną cechą. Ustalono, że dla

* Instytut Informatyki, Automatyki i Robotyki Politechniki Wrocławskiej, ul. Janiszewskiego 11/17, 50-372 Wrocław.

jednomodalnych funkcji celu populacja dąży do stanu równowagi wokół optimum, a jej rozkład jest normalny. Nieskończone populacje są więc zbieżne według średniej i z prawdopodobieństwem do optimum jednomodalnej funkcji celu. Dla funkcji wielomodalnych rozkład nieskończonych populacji jest mieszaniną rozkładów gaussowskich, których parametry określono. Rozkład sumaryczny mieszaniny jest przesunięty w stronę optimum globalnego. Jego wartość oczekiwana leży w pobliżu optimum globalnego, w odległości zależnej od wariancji mutacji. Symulacyjnie zbadano wpływ postaci funkcji celu oraz parametrów procesu ewolucyjnego na rozkłady populacji.

Analizę dwuelementowych populacji przeprowadzono w przestrzeni stanów populacji, w której każdy punkt odpowiada całej populacji. Obliczono wartości oczekiwane stanów populacji, które generują dwuwymiarowy dyskretny układ dynamiczny. Głównym przedmiotem zainteresowań były trajektorie oraz zachowania asymptotyczne układu. Analizowano liczbę i położenie punktów stałych oraz określono warunki ich stabilności. Badano obszary przyciągania punktów stałych oraz szybkość zbieżności do nich. Liczba punktów stałych układu zależy od modalności funkcji celu, a także od wartości odchylenia standardowego mutacji. Punkty stałe są zwykle położone w pobliżu optimumów funkcji celu i siodeł pomiędzy optimumami. Udowodniono, że punkty stałe leżące w pobliżu siodeł są niestabilne. Stabilne punkty stałe położone w pobliżu optimumów mogą stracić stabilność dla dużych wartości odchylenia standardowego mutacji. Pojawiają się wtedy orbity periodyczne. Dla niesymetrycznych funkcji celu obserwowano ciąg bifurkacji podwajania okresu prowadzący do zachowań chaotycznych. Zachowania chaotyczne badano następującymi metodami: wykresów fazowych, widma mocy, funkcji autokorelacji oraz wykładników Lapunowa.

Uzyskane rezultaty badań dynamiki nieskończonych i dwuelementowych populacji wykorzystano do analizy innych aspektów działania metod ewolucyjnych. Badano wpływ różnych mechanizmów ewolucyjnych: selekcji, rekombinacji i mutacji, na różnorodność populacji. Wyniki symulacyjne potwierdziły wcześniejsze obserwacje dotyczące zmian różnorodności populacji w trakcie procesu ewolucyjnego. Początkowo znacznie rozproszone populacje szybko się skupiają, a początkowo jednorodne populacje zwiększają swą różnorodność. W efekcie, populacja w zaledwie kilku generacjach tworzy zwarty klaster osobników o zbliżonych typach. Największy wpływ na zwiększanie lub podtrzymywanie różnorodności populacji ma mutacja. Selekcja i w większości przypadków rekombinacja powodują zmniejszenie rozproszenia. Parametry krytyczne układu dynamicznego, dla których punkty stałe zmieniają stabilność, posłużyły do opracowania heurystyk dotyczących sposobów doboru parametrów algorytmu ewolucyjnego oraz identyfikacji nieznannej jednowymiarowej funkcji celu.

W pracy dokonano także przeglądu znanych w literaturze, a związanych z tematem pracy, sposobów analizy metod ewolucyjnych: modele algorytmów fenotypowych oraz modele wykorzystujące łańcuchy Markowa i teorię układów dynamicznych. Przegląd uzupełniono krótkim opisem innych popularnych sposobów analizy metod ewolucyjnych.

Wykaz ważniejszych skrótów i oznaczeń

Skalary oznaczone są czcionką normalną, wektory i macierze – pogrubioną. Indeks górny oznacza zwykle numer iteracji (z wyjątkiem analizy nieskończonych populacji, gdzie numer populacji jest oznaczany przez indeks dolny), indeks dolny – numer osobnika.

- a, ζ, h, d – parametry gaussowskich funkcji celu: nachylenie zbocza, szerokość funkcji, wysokość i położenie optimum
- A, B, C – parametry trójkątnych funkcji celu
- AG – algorytm genetyczny
- NFLT – twierdzenie *no free lunch*
- SGA – prosty algorytm genetyczny
- SE – strategie ewolucyjne
- $B(x^s)$ – obszar przyciągania punktu stałego x^s
- div* – miara rozproszenia populacji
- $E[x], D^2[x]$ – wartość oczekiwana, wariancja zmiennej x
- FGP – funkcja gęstości prawdopodobieństwa
- $f(\mathbf{x}|P^i)$ – FGP rozkładu warunkowego typu x
- F_1, F_2 – odwzorowania definiujące dyskretny układ dynamiczny
- $g(\mathbf{x}, \mathbf{y})$ – rozkład mutacji osobnika $\mathbf{y}$
- $\mathcal{G}$ – iterowany operator rozkładu prawdopodobieństwa w modelu Vose’a, $\mathcal{G}^0(\mathbf{p}) = \mathbf{p}$, $\mathcal{G}^{i+1}(\mathbf{p}) = \mathcal{G}^i(\mathcal{G}(\mathbf{p}))$, $i = 1, 2, \dots$
- l – długość wektora binarnego
- m – liczebność populacji
- n – rozmiar przestrzeni poszukiwań
- $\mathcal{N}(\bar{\mathbf{x}}, \sigma)$ – rozkład normalny o wartości oczekiwanej $\bar{\mathbf{x}}$ i odchyleniu standardowym σ
- q – funkcja celu
- $\mathbf{p}$ – wektor populacji w modelu Vose’a
- Ω – dyskretna przestrzeń stanów w modelu Vose’a
- p_m – prawdopodobieństwo mutacji
- $p(x)$ – funkcja gęstości prawdopodobieństwa osobników z cechą x dla nieskończonej populacji
- P – populacja
- $P(\omega), C(\tau)$ – widmo mocy, funkcja autokorelacji
- Pr – prawdopodobieństwo
- R^n – n -wymiarowa przestrzeń rzeczywistoliczbowa

s	–	stan populacji
S	–	$n \cdot m$ -wymiarowa przestrzeń stanów
S_U	–	ilorazowa przestrzeń stanów
T	–	n -wymiarowa przestrzeń typów
U	–	relacja równoważności
x	–	wektor cech, typ osobnika
w, z	–	nowe współrzędne w ilorazowej przestrzeni stanów
$\alpha(x)$	–	prawdopodobieństwo wyboru elementu x zgodnie z selekcją proporcjonalną
β	–	wagi rozkładów składowych nieskończonej populacji
Δ_q	–	przesunięcie pomiędzy funkcjami celu
η	–	wartość oczekiwana rozkładów składowych nieskończonych populacji
λ_i	–	i -ta wartość własna
λ	–	wykładnik Lapunowa
μ	–	wartość oczekiwana rozkładu mieszaniny nieskończonej populacji
ν^2	–	wariancja rozkładów składowych nieskończonych populacji
ξ	–	zmienna losowa o zadanym rozkładzie
σ, σ^2	–	odchylenie standardowe, wariancja mutacji
v^2	–	wariancja rozkładu mieszaniny nieskończonej populacji
Ψ, Θ	–	funkcje pomocnicze w przestrzeni stanów
ω^s	–	punkt stały układu dynamicznego
$\binom{n}{k}$	–	symbol Newtona

1. Wstęp

1.1. Tło historyczne i motywacja

W ostatnich latach obserwuje się znaczną popularność metod inspirowanych ewolucją naturalną, które są stosowane głównie do rozwiązywania problemów optymalizacyjnych. Odpowiednio zaprojektowane algorytmy ewolucyjne znajdują rozwiązania (dokładne lub przybliżone) wielu skomplikowanych zadań optymalizacyjnych, w tym NP-trudnych; z wielowymiarowymi przestrzeniami poszukiwań; o funkcjach celu wielomodalnych, zaszumionych, nieciągłych, nieróżniczkowalnych czy nie zadanych w postaci analitycznej; z niepełnymi danymi czy ukrytymi zmiennymi; z dynamicznie zmieniającym się środowiskiem; z nieliniowymi ograniczeniami. O swoistej „modzie na ewolucję” świadczyć może lawinowy wzrost liczby zastosowań metod ewolucyjnych, rosnąca liczba książek, publikacji, czasopism oraz konferencji poświęconych metodom ewolucyjnym.

Początki obliczeń ewolucyjnych datuje się na lata 60. ubiegłego wieku, gdy na fali fascynacji teorią ewolucji (a szczególnie genetyką) i wiary w „moc ewolucji” zaczęto przenosić idee biologiczne do techniki, zwłaszcza do metod obliczeniowych. Za prekursora zastosowania mechanizmów ewolucyjnych w teorii optymalizacji uważa się H.J. Bremermanna [21]. Niewiele później L. Fogel wraz ze współpracownikami [61] zastosowali idee ewolucji do pewnego problemu sztucznej inteligencji, w którym symulowano ewolucję maszyn o skończonej liczbie stanów do prognozowania sekwencji symboli spełniających zadane kryteria. Zaproponowana technika została rozwinięta przez D. Fogla [57] i nazwana *programowaniem ewolucyjnym* (ang. *evolutionary programming*, EP)¹. Na początku lat 70. XX w. I. Rechenberg [170] używał metody inspirowanej ewolucją do wyznaczania optymalnych parametrów podczas eksperymentów laboratoryjnych. Rozwinięcie jego pomysłów dało początek *strategiom ewolucyjnym* (niem. *Evolutionsstrategie*, ang. *evolution strategy*, ES) [187]. J. Holland [91] zaproponował wykorzystanie mechanizmów genetycznych do rozwiązywania zadań optymalizacji, zapoczątkowując najbardziej popularną i rozwiniętą gałąź metod ewolucyjnych określaną jako *algorytmy genetyczne* (ang. *genetic algorithms*, GA). Do polskich twórców metod ewolucyjnych należy R. Galar [62], który zaproponował oryginalne modele *poszu-*

¹ Ponieważ nazewnictwo w języku polskim nie jest do końca ustalone, w pracy do terminów polskich dodano ich angielskie odpowiedniki, a w niektórych przypadkach podano jedynie nazwy angielskie.

kiwań ewolucyjnych z miękką selekcją, badał ich właściwości oraz zastosowania w optymalizacji.

Rozwój tej dziedziny nauki w latach 80. i 90. związany był z upowszechnieniem komputerów, ich wzrastającą mocą obliczeniową i zasobami pamięci. Metody optymalizacyjne inspirowane ewolucją znane są obecnie pod wspólną nazwą *algorytmów* lub *obliczeń ewolucyjnych* (ang. *evolutionary algorithms/computation*) i obejmują, oprócz wymienionych uprzednio algorytmów genetycznych, strategii ewolucyjnych i programowania ewolucyjnego, także nowsze techniki, takie jak *programowanie genetyczne* (ang. *genetic programming*) czy *programy ewolucyjne* (ang. *evolution programs*).

Duża łatwość implementacji metod ewolucyjnych, a przede wszystkim ich wysoka efektywność w zadaniach optymalizacji globalnej spowodowała, że nawet typowo biologiczna terminologia nie odstrasza rzeszy potencjalnych użytkowników tych metod, przeważnie o technicznym wykształceniu. Sukces technik ewolucyjnych wynika z ich prostoty, łatwości zastosowań dla szerokiego spektrum problemów inżynierskich oraz dużej skuteczności w rozwiązywaniu złożonych zadań optymalizacyjnych. Chociaż metody ewolucyjne nie gwarantują znalezienia najlepszego rozwiązania², ich działanie zależy od odpowiedniego wyboru parametrów, operatorów i reprezentacji, często wymagają więcej obliczeń wartości funkcji celu niż metody klasyczne³, to jednak sprawnie (a w wielu przypadkach jako jedyne z dotychczas stosowanych) znajdują rozwiązania akceptowalne. Łatwo je także łączyć z innymi, zwykle lokalnymi technikami, tworząc metody hybrydowe. Istnieje możliwość włączenia wiedzy o specyfice dziedziny czy konkretnego zadania do reprezentacji rozwiązywanego zadania, wykorzystywanych procedur lub wartości inicjujących obliczenia. Komputerowe wspomaganie projektowania (projektowanie kształtów, rozplanowanie), rozpoznawanie obrazów [141], zarządzanie produkcją (szeregowanie zadań, planowanie procesów, planowanie eksperymentów [99], harmonogramowanie), przetwarzanie sygnałów, chemia (analiza strukturalna), fizyka (elektromagnetyzm, mechanika płynów), robotyka (planowanie ścieżek i ruchu robotów, koordynacja działania robotów [42, 43, 45]), elektronika (projektowanie układów [4]), telekomunikacja (projektowanie sieci [4]), uczenie maszynowe (systemy uczące i klasyfikujące, automatyczne programowanie), sieci neuronowe (optymalizacja topologii [153], algorytmy uczące [127]), energetyka, a nawet sztuka – to tylko kilka wybranych dziedzin, w których metody ewolucyjne znalazły zastosowanie. Stosuje się je jednak nie tylko do rozwiązywania zadań optymalizacji, lecz także w zadaniach sztucznej inteligencji (szczególnie w systemach uczących się) lub jako narzędzie symulacyjne do analizy pewnych

² Są to raczej metody adaptacyjne, a nie optymalizacyjne *per se* [35], chociaż stosowane są głównie jako te drugie.

³ Generujące deterministyczny ciąg rozwiązań na podstawie obliczeń np. gradientu funkcji celu.

mechanizmów i zjawisk, np. w teorii ewolucji naturalnej, ekologii, ewolucyjnej ekonomii, socjologii. W niniejszej monografii obliczenia ewolucyjne rozpatrywane będą w aspekcie metod adaptacyjnych, aczkolwiek badany model może być także rozważany w kategoriach modeli ewolucji biologicznej. Ponieważ metody adaptacyjne nie zatrzymują się po zidentyfikowaniu optimum, mogą zlokalizować wiele optimów lokalnych, a wśród nich globalne, lub adaptować się do ewentualnych zmian (w dynamicznym środowisku) znalezionej już optimum. Poszukiwania kolejnych optimów prowadzone są w taki sposób, aby straty w jakości rozwiązań pośrednich były jak najmniejsze, co ma duże znaczenie przy zadaniach rozwiązywanych *on-line*. Kryterium adaptacji zadaje dla tych metod funkcja celu.

Wielości aplikacji metod ewolucyjnych nie towarzyszy, niestety, rozwój ich teorii. Stochastyczna i heurystyczna natura metod ewolucyjnych utrudnia, lub wręcz uniemożliwia, opracowanie podstaw teoretycznych wielu z nich. Dodatkową komplikację powoduje różnorodność algorytmów, reprezentacji i operatorów, struktura populacyjna, różnorodność optymalizowanych funkcji, ich nieliniowość lub brak jawnej postaci. Do naukowego traktowania metod ewolucyjnych konieczna jest analiza teoretyczna, aby w pełni lub choćby częściowo wyjaśnić istotę procesu ewolucji bądź jego poszczególnych mechanizmów. Zrozumienie działania metod inspirowanych ewolucją, czy tylko zbadanie ich właściwości, umożliwiłoby lepsze projektowanie algorytmów, zwiększenie ich efektywności i wykorzystanie odpowiednich operatorów. Pozwoliłoby również na określenie jak zmiany parametrów algorytmów ewolucyjnych wpływają na ich działanie i mogłoby być pomocne w optymalnym doborze parametrów. Ułatwiłoby także opracowanie praktycznych zasad implementacji i dało odpowiedź na nurtujące praktyków pytanie: czy stosować algorytm ewolucyjny do rozwiązania konkretnego zadania? Jeśli tak – to który, z jaką reprezentacją, z jakimi operatorami i jakie powinny być wartości parametrów? Modele matematyczne metod ewolucyjnych mogą być także pomocne dla biologów-teoretyków w badaniach różnorodnych cech i zachowań osobników oraz populacji. Trudność materii oddaje trafnie znane powiedzenie Johna von Neumanna *If people do not believe that mathematics is simple, it is only because they do not realize how complicated life is*, które można odnieść zarówno do nauk biologicznych, jak i społecznych oraz traktować zarówno jak przestrożę, lecz także jako wyzwanie.

W ostatnich latach obserwuje się wzrost zainteresowań ewolucyjnymi metodami fenotypowymi, które działają bezpośrednio na przestrzeniach poszukiwań określonych przez funkcje celu, bez konieczności ich kodowania i dekodowania do innych przestrzeni (głównie binarnych). Jedną z metod fenotypowych z kodowaniem rzeczywistoliczbowym jest *metoda poszukiwań ewolucyjnych z miękką selekcją* [62], będąca przedmiotem zainteresowań dysertacji. Podstawowa wersja algorytmu implementującego metodę wykorzystuje jedynie selekcję proporcjonalną (zwaną w metodzie miękką) oraz mutację z rozkładem normalnym. Metoda, charaktery-

zująca się niewielką liczbą parametrów, jest efektywna w optymalizacji globalnej dla wielowymiarowych przestrzeni poszukiwań. Wobec trudności jakie nastręcza analiza teoretyczna metod ewolucyjnych (szczególnie w przestrzeniach ciągłych) wydaje się metodologicznie uzasadnione podjąć próbę analizy dla modeli skrajnie nieskomplikowanych, a uzyskane wyniki rozszerzać na coraz bardziej złożone.

1.2. Cel i zakres pracy

W literaturze zaproponowano wiele modeli i teorii dla metod ewolucyjnych. Podstawowe prace w zakresie algorytmów genetycznych [91, 203] czy strategii ewolucyjnych [14] nie są jednak kompletne. Przyjmują wiele założeń upraszczających, dotyczą często wersji kanonicznych i bardzo prostych funkcji celu, istnieje więc potrzeba dalszych badań w tym kierunku. Niniejsza monografia stanowi próbę wzbogacenia studiów teoretycznych nad fenotypowymi metodami ewolucyjnymi działającymi w przestrzeniach rzeczywistych [14, 163, 164] o syntezę wyników teoretycznych i symulacyjnych otrzymanych w wyniku badań modelu ewolucji z miękką selekcją [62]. Przedmiotem zainteresowania jest analiza dynamiki populacji ewoluujących zgodnie z regułami tego modelu.

Głównym wyróżnikiem prezentowanych badań jest liczebność populacji. Analizowano dwa skrajne przypadki: populacje dwuelementowe oraz nieskończenie liczne. Wyniki teoretyczne uzyskano dla jednowymiarowej przestrzeni poszukiwań.

Nieskończone populacje opisano ich rozkładem w przestrzeni poszukiwań, co pozwoliło na analizę dynamiki w optymalizacji lokalnej oraz globalnej. Udowodniono, że rozkład populacji znajdującej się w stanie równowagi selekcyjno-mutacyjnej wokół optimum lokalnego jest normalny i określono jego parametry. Dla wielomodalnych funkcji celu pokazano, że populacja ma rozkład będący mieszaniną rozkładów gaussowskich. Opisano sposób konstrukcji rozkładu mieszaniny oraz obliczania parametrów rozkładów składowych. Przedstawiono wyniki badań symulacyjnych, w których rozważano wpływ parametrów algorytmu ewolucyjnego oraz funkcji celu na rozkłady populacji. Analiza wartości oczekiwanych pozwoliła na pokazanie fraktalnych właściwości modelu nieskończonych populacji (zbioru Cantora generowanego przez te wartości).

Stosując zaproponowany sposób opisu procesu ewolucji w postaci rozkładu prawdopodobieństwa położenia elementów w kolejnej generacji, studiowano dynamikę dla drugiego skrajnego przypadku – populacji dwuelementowych. Ewolucję badano w przestrzeni wszystkich możliwych populacji (przestrzeni stanów) i analizowano charakterystyki sekwencji populacji, generowanych przez proces. Ograniczenie analizy do jednowymiarowych funkcji celu umożliwiło teoretyczne badania wartości oczekiwanych stanów populacji i ich trajektorii. Pokazano, że

wartości oczekiwane stanów populacji generują dwuwymiarowy dyskretny układ dynamiczny. Zbadano zachowania asymptotyczne tego układu, określono położenia i charakter punktów stałych oraz wyspecyfikowano warunki ich stabilności. Szczegółowe badania przeprowadzono dla wybranego zbioru funkcji celu (jedno- i dwumodalnych, symetrycznych i asymetrycznych). Dla asymetrycznych funkcji celu zaobserwowano zachowania chaotyczne, które dokładniej rozważono stosując, znane z teorii układów dynamicznych, metody badania zjawisk chaotycznych. Dokonano analizy oczekiwanego czasu zbieżności wartości oczekiwanych stanów populacji do optimum. Przedstawiono także wyniki badań symulacyjnych uogólniających model na populację kilkuelementową.

Badania teoretyczne uzupełniono wynikami badań symulacyjnych dotyczących różnorodności oraz wielkości populacji. Badano wpływ różnych rodzajów operatorów selekcji, rekombinacji i mutacji na rozproszenie populacji. Pokazano, że mutacja jest mechanizmem podtrzymującym lub zwiększającym różnorodność populacji, a pozostałe operatory powodują jej zmniejszenie. Zaproponowano wykorzystanie wyników analizy zachowań asymptotycznych układu dynamicznego do określenia parametrów procesu ewolucyjnego oraz identyfikacji nieznanego, jednowymiarowego funkcji celu.

Uzyskane wyniki pozwalają na opisanie dynamiki analizowanego procesu ewolucyjnego. Symulowana ewolucja jest procesem o dwóch „prędkościach”: populacja bardzo szybko tworzy klastery elementów o podobnych cechach, który następnie wolno dryfuje w kierunku obszarów o wyższych przystosowaniach. Utworzenie przez populację klastra oznacza osiągnięcie pewnego charakterystycznego poziomu różnorodności: początkowo rozproszone populacje ujednolicają się, a początkowo jednorodne – zwiększają swą różnorodność. Zmniejszanie różnorodności początkowo znacznie zróżnicowanej populacji (o rozproszeniu rzędu kilku odchyleń standardowych mutacji) jest nieodłączną cechą ewolucji i zachodzi nawet wtedy, gdy presja selekcyjna nie działa. Zwarta populacja wędruje wzdłuż grzbietów funkcji celu, a lokalizacja optimum jest zwykle niedokładna. W przypadku analizowanego modelu, nie należy przeciwdziałać utracie różnorodności populacji, gdyż poszukiwanie optimum funkcji celu nie polega na „rozrzuceniu” populacji tak, aby z dużym prawdopodobieństwem pewien element znalazł się w otoczeniu ekstremum, lecz na eksploracji przestrzeni poszukiwań przez zwartą populację w procesie zależnym od ścieżki. Bardzo ważnym czynnikiem mającym wpływ na dynamikę populacji jest wielkość odchylenia standardowego mutacji. Skupienie populacji („promień” klastra) jest rzędu wielkości tego parametru. Im mniejsze odchylenie standardowe mutacji, tym lepsza identyfikacja optimum. Zbyt duża wartość parametru powoduje, że w procesie ewolucji pojawiają się zachowania chaotyczne i populacja nie jest w stanie się ustabilizować. Badania pokazują, że wysoką efektywność w eksploracji krajobrazu adaptacyjnego w poszukiwaniu optimum globalnego funkcji celu wykazują małe (kilkuelementowe) populacje.

1.3. Układ pracy

Praca składa się z siedmiu rozdziałów uzupełnionych dodatkiem i spisem literatury. W rozdziale drugim opisano pokrótce metody ewolucyjne, przedstawiając ich podstawowe idee, algorytmy i mechanizmy. Szczegółowo przedstawiono metodę poszukiwań ewolucyjnych z miękką selekcją, jako podstawowy algorytm badany w dysertacji. Omówiono także dynamikę poszukiwań ewolucyjnych, wykorzystując ideę krajobrazu adaptacyjnego funkcji przystosowania.

Rozdział trzeci zawiera przegląd sposobów teoretycznej analizy metod ewolucyjnych. Omówiono teorie związane z tematem monografii: modele fenotypowe z kodowaniem rzeczywistoliczbowym, modele wykorzystujące łańcuchy Markowa oraz modele systemów dynamicznych. Pokrótce opisano inne popularne podejścia: teorię schematów, modele wykorzystujące fizykę statystyczną, modele oparte na teorii genetyki populacyjnej.

W kolejnych rozdziałach przedstawiono analizę metody poszukiwań ewolucyjnych z miękką selekcją. Dynamikę nieskończonych populacji w optymalizacji lokalnej i globalnej opisano w rozdziale czwartym. Przedstawiono wyniki badań symulacyjnych rozkładów oraz właściwości fraktalne wartości oczekiwanych rozkładów. Rozdział piąty zawiera analizę dynamiki bardzo małych populacji. Zaproponowano metodę analizy dwuelementowej populacji w przestrzeni stanów i przedstawiono badania wartości oczekiwanych stanów populacji. Omówiono dyskretny układ dynamiczny generowany przez wartości oczekiwane stanów populacji oraz przedstawiono badania punktów stałych i ich stabilności. Podrozdziały zawierają wyniki symulacyjne zachowań chaotycznych oraz oczekiwanego czasu zbieżności do optimum. Rozdział zamykają uwagi dotyczące kilkuelementowych populacji.

W rozdziale szóstym opisano wykorzystanie rezultatów analizy dynamiki populacji do badań innych aspektów działania metod ewolucyjnych. Rozważano wpływ operatorów rekombinacji i mutacji na różnorodność populacji. Zaproponowano sposoby zastosowania wyników analizy układu dynamicznego do określania parametrów procesu ewolucyjnego i funkcji celu. Przedstawiono dyskusję dotyczącą doboru wielkości populacji.

Rozdział siódmy stanowi podsumowanie pracy. Zebrano w nim najważniejsze rezultaty teoretyczne i wnioski wynikające z symulacji. Całości dopełnia rozdział, w którym opisano funkcje testowe wykorzystywane w pracy.

2. Metody ewolucyjne

Ze względu na powszechność literatury dotyczącej metod ewolucyjnych (także w języku polskim [3, 79, 139, 140]) w rozdziale zaprezentowano ogólny schemat działania metod ewolucyjnych, uwypuklając jedynie zagadnienia, do których nawiązano w kolejnych częściach rozprawy. Z bogatego zbioru technik ewolucyjnych opisano pokrótce *algorytmy genetyczne* – jako najbardziej popularne i najlepiej przebadane, oraz *strategie ewolucyjne* – jako przykład metody fenotypowej z kodowaniem rzeczywistoliczbowym. Omówiono jedynie operatory genetyczne najczęściej stosowane w literaturze oraz wykorzystywane w pracy. Bardziej szczegółowo przedstawiono metody fenotypowe, a przede wszystkim *metodę poszukiwań ewolucyjnych z miękką selekcją*, która stanowi podstawę prezentowanych dalej badań. Rozdział zamyka ogólny opis dynamiki poszukiwań ewolucyjnych wraz z definicją krajobrazu adaptacyjnego, będący wstępem do analizy teoretycznej ewolucji fenotypowej.

2.1. Podstawowe idee

2.1.1. Ogólny schemat algorytmu ewolucyjnego

W klasycznych metodach optymalizacyjnych przeszukiwanie przestrzeni w celu znalezienia *optimum*¹ skalarnej *funkcji celu* $q(\mathbf{x})$ odbywa się przez odpowiednie generowanie sekwencji *punktów bazowych* $\mathbf{x}^0, \mathbf{x}^1, \dots$, stanowiących potencjalne rozwiązania. Zwykle punkt bazowy w iteracji $(i+1)$ jest pewną modyfikacją punktu bazowego z iteracji i -tej oraz ma lepszą jakość $q(\mathbf{x}^{i+1}) > q(\mathbf{x}^i)$. Metody ewolucyjne nie operują na pojedynczym punkcie bazowym, lecz na m -elementowym zbiorze zwanym *populacją* $P = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\}$. Element zbioru utożsamiany jest w biologii z *osobnikiem*, a w optymalizacji z *rozwiązaniem*. Osobnik $\mathbf{x}$ opisywany jest przez n -wymiarowy wektor² cech zwany *typem* i przyjmujący wartości z *przestrzeni typów* T . W zależności od metody, typ może odpowiadać biologicznemu *genotypowi*³ lub *fenotypowi*⁴, a przestrzeń typów może być ciągła lub dyskretna. Każdemu typowi przypisany jest *wskaźnik przystosowania*, czyli wartość liczbowa

¹ W pracy zadanie poszukiwania optimum jest utożsamiane z poszukiwaniem maksimum.

² W ogólnym przypadku może to być także macierz, drzewo czy graf.

³ Zespół genów danego organizmu warunkujących dziedziczenie.

⁴ Zbiór cech organizmu uwarunkowany przez genotyp oraz środowisko.

określająca jakość osobnika, obliczany na podstawie nieujemnej *funkcji przystosowania* (ang. *fitness function*). W ogólnym przypadku funkcja przystosowania jest odpowiednikiem funkcji celu, jednak gdy fenotyp jest kodowany przez genotyp, wtedy wartości przystosowania (określanej dla genotypów) nie musi odpowiadać wartość funkcji celu (ocenie fenotypu)⁵.

Pomimo szczegółowych różnic implementacyjnych można przedstawić ogólny schemat działania algorytmu ewolucyjnego [9, 139]. Ewolucja (proces przeszukiwania przestrzeni) następuje poprzez tworzenie kolejnych *generacji* rozwiązań wykorzystując, inspirowane procesami ewolucji naturalnej, mechanizmy selekcji, rekombinacji i mutacji (zob. schemat na rys. 2.1). Zwykle populację początkową

procedure ALGORYTM EWOLUCYJNY

begin

$i := 0$

inicjalizacja $P^0 = \{\mathbf{x}_1^0, \mathbf{x}_2^0, \dots, \mathbf{x}_m^0\}$

ocena $P^0 \rightarrow \{q_1^0, q_2^0, \dots, q_m^0\}$

while (warunek stopu niespełniony) do

begin

selekcja P' z P^i

operacje ewolucyjne (rekombinacja, mutacja) P'' z P'

ocena $P'' \rightarrow \{q_1'', q_2'', \dots, q_m''\}$

zastępowanie $(P^i, P'') \rightarrow P^{i+1}$

$i := i + 1$

end

end

Rys. 2.1. Schemat algorytmu ewolucyjnego [3, 139]

Fig. 2.1. Evolutionary algorithm according to [3, 139]

tworzy się losowo, korzystając z rozkładu równomiernego w (ograniczonej) przestrzeni poszukiwań, choć w pewnych metodach preferuje się populacje z elementami o jednakowych lub zbliżonych typach. Każdy osobnik poddawany jest ocenie: na podstawie funkcji celu obliczany jest jego wskaźnik przystosowania. Osobniki w generacji $(i + 1)$ -szej są potomkami elementów populacji P^i z i -tej generacji. Proces *reprodukcji* polega na *selekcji* z P^i *bazowej populacji rodzicielskiej*, która poddana działaniu *stochastycznych operatorów ewolucyjnych* tworzy nową generację⁶. Proces selekcji jest zwykle procesem stochastycznym, choć w niektórych

⁵ Funkcja celu (o przeciwdziałaniu niekoniecznie nieujemnej) jest powszechnie stosowana w optymalizacji. Funkcja przystosowania, definiowana w metodach ewolucyjnych, jest nieujemna. Dalej oba terminy są stosowane zamiennie.

⁶ Za proces reprodukcji uważamy wszystkie operacje przeprowadzane na generacji bieżącej, które w rezultacie tworzą kolejną generację osobników. Niekiedy reprodukcją zwany jest jedynie proces selekcji rodziców z populacji bieżącej [3, 58].

metodach stosuje się selekcję deterministyczną. Do przekształcenia typów rodzicielskich w potomne używa się różnych operatorów, głównie *rekombinacji* i *mutacji*. Operatory działają na pojedynczych elementach (mutacja) lub grupach, najczęściej parach, osobników (rekombinacje). Nowo utworzeni potomkowie stanowią kolejną generację (ang. *generational model*). Inne pomysły to zastępowanie przez nowo utworzonych potomków nie całej populacji bieżącej, lecz tylko jej części (ang. *generation gap model*) lub tylko jednego, zwykle najgorszego osobnika (ang. *steady-state model*) [37, 196]. W niektórych metodach (najczęściej strategiach ewolucyjnych) zastępowanie (zwane także sukcesją) określa sposób wyboru elementów do nowej populacji z większego niż liczebność populacji bazowej zbioru potomków lub łącznego zbioru rodziców i potomków. Ten ogólny cykl ewolucyjny przebiegający według reguł:

$$\begin{aligned} \text{bieżąca generacja} &\rightarrow \text{reprodukcja} = \text{selekcja} + \text{operacje ewolucyjne} \rightarrow \\ &\rightarrow \text{nowa generacja} \end{aligned}$$

powtarza się we wszystkich typach metod ewolucyjnych. Różnice polegają na szczegółowych implementacjach metod selekcji, operatorów i różnych reprezentacjach typów osobników.

2.1.2. Algorytmy genetyczne (AG)

Zaproponowana przez J. Hollanda [91] metoda rozwiązywania zadań optymalizacyjnych, bazująca na mechanizmach znanych z genetyki, dała początek najpopularniejszej gałęzi metod ewolucyjnych zwanej *algorytmami genetycznymi* (AG). Podstawowymi strukturami, na których operują AG, są wektory zwane *chromosomami*. Zasadniczą jednostką chromosomu są *geny*, które mogą przyjmować ograniczony zbiór wartości zwanych *allelami*. Najczęściej geny mają *reprezentację binarną*, czyli allele mogą przyjmować wartości 0 lub 1. Reprezentacja binarna cechująca się prostotą implementacyjną uważana była za najbardziej efektywną (szczegóły w rozdziale 2.3). Pozycja genu w chromosomie nazywana jest *locusem*. Aktualna struktura genetyczna osobnika odpowiada jego *genotypowi*, a obserwowane przystosowanie wynikające z danego genotypu nazywane jest *fenotypem*. Zbiór chromosomów tworzy *populację*. Proces reprodukcji polega na *selekcji* elementów, które poddane operacjom *rekombinacji* i/lub *mutacji* stanowią nową generację.

Klasyczna wersja algorytmu genetycznego, zwana też prostą lub kanoniczną (ang. *simple genetic algorithm*, SGA), działa na wektorach binarnych o stałej długości, wykorzystując selekcję proporcjonalną, krzyżowanie jednopunktowe i mutację zmieniającą, z zadanyim prawdopodobieństwem mutacji p_m , każdy allel na przeciwny. Obecnie częściej stosuje się inne rodzaje selekcji (przykładowo turniejową, rankingową czy progową), a także inne wersje krzyżowania (np. wie-

lopunktowe, równomierne). W teorii AG podkreśla się znaczenie rekombinacji realizowanej za pomocą operatora krzyżowania.

2.1.3. Strategie ewolucyjne (SE)

Grupa algorytmów zwana *strategiami ewolucyjnymi* (SE) wykorzystywana jest głównie jako narzędzie optymalizacji parametrycznej w R^n . Rozwinęła się z prostej metody zaproponowanej przez I. Rechenberga [170] nazywanej *strategią* $(1 + 1)$, która za pomocą mutacji i selekcji przekształcała tylko jednego osobnika w jego potomka. Typ elementu opisywany był wektorem liczb rzeczywistych. Najważniejszą rolę w tej metodzie odgrywała mutacja z rozkładem normalnym, w której do każdej współrzędnej wektora cech dodawano zmienną losową o rozkładzie normalnym. Zasięg mutacji, mierzony odchyleniem standardowym σ , dobierano według *reguły 1/5 sukcesów*, czyli iloraz wszystkich zakończonych sukcesem mutacji (dających lepsze przystosowania) i liczby wykonanych mutacji powinien wynosić w przybliżeniu 1/5. Jeżeli iloraz był większy – należało zwiększyć odchylenie standardowe, jeżeli był mniejszy – zmniejszano ten parametr.

W kolejnych wersjach metody znanych jako $(\mu + \lambda)$ i (μ, λ) , z μ -elementowych populacji bazowych generuje się λ -elementowe populacje potomne [187]. Rodzice losowani są z jednakowym prawdopodobieństwem (ze zwracaniem) z populacji bazowej, a ich typy poddaje się nie tylko mutacji, lecz często także rekombinacji. Następnie ze zbioru rodziców i potomków, dla strategii $(\mu + \lambda)$, lub tylko samych potomków, dla strategii (μ, λ) , wybiera się μ najlepszych osobników (zastępowanie deterministyczne) do nowej populacji. Wymienione wersje algorytmów wzbogacono o *samoadaptację* parametrów sterujących, które ewoluują wraz z osobnikami. Elementy populacji są wtedy kodowane za pomocą dwóch wektorów o wartościach rzeczywistych. W jednym wektorze przechowywane są typy osobników, a w drugim wartości parametrów sterujących procesu, najczęściej odchylenie standardowe mutacji. Samoadaptacja zasięgu mutacji, która zastąpiła regułę 1/5 sukcesów, powoduje zmianę wielkości w drugim chromosomie tak, aby minimalizować wartość oczekiwaną czasu (liczby iteracji) osiągnięcia optimum. Dostrajane parametry wykorzystywane są do transformacji typów rodzicielskich w potomne.

2.2. Mechanizmy ewolucji

W metodach ewolucyjnych stosuje się bardzo różne sposoby selekcji, mutacji i rekombinacji. W podrozdziale przedstawiono kilka powszechnie stosowanych operatorów uzupełnionych o te wykorzystywane w dalszej części pracy. Ponieważ przedmiotem zainteresowania są metody oparte na ewolucji fenotypowej w R^n , akcent położono na mechanizmy stosowane dla reprezentacji rzeczywistoliczbowej.

2.2.1. Metody selekcji

W zależności od metody, selekcja może być procesem deterministycznym lub stochastycznym, a jej najczęstsze postacie to:

Selekcja twarda (ang. *hard*) – jest selekcją deterministyczną, która polega na wyborze rozwiązania do tej pory najlepszego na jedyne go rodzica m osobników kolejnej generacji.

Selekcja progowa (ang. *truncation*) – może być traktowana jako pewna modyfikacja selekcji twardej. W selekcji progowej z jednakowym prawdopodobieństwem reprodukuje się jedynie pewna część populacji bazowej, ta najlepsza, o rangach powyżej ustalonego poziomu progowego Tr (ang. *truncation threshold*)⁷.

Selekcja proporcjonalna/ruletkowa (ang. *proportional/roulette wheel*) – jest to prosta wersja selekcji stochastycznej. Prawdopodobieństwo wyboru elementu $\mathbf{x}_k$ na rodzica jest proporcjonalne do jego przystosowania: $Pr(\mathbf{x}_k) = q(\mathbf{x}_k) / \sum_{j=1}^m q(\mathbf{x}_j)$. Intensywność selekcji proporcjonalnej zmniejsza się w kolejnych generacjach oraz na płaskich obszarach funkcji celu, ponieważ wartości przystosowania osobników stają się zbliżone.

Selekcja turniejowa (ang. *tournament*) – należy do grupy selekcji stochastycznych. Z populacji bieżącej wybierana jest losowo (z jednakowym prawdopodobieństwem, jako wariacja z powtórzeniami) k -elementowa grupa osobników zwana *turniejem*. Wśród elementów wybranej grupy rozgrywany jest turniej, którego zwycięzcą jest element o najwyższej jakości. Zwycięzca lokuje swego potomka w nowej generacji. Turnieje powtarzają się m razy, by wypełnić generację potomną. Najczęściej stosowany wariant selekcji turniejowej, zwany turniejem binarnym, zakłada współzawodnictwo tylko dwóch osobników ($k = 2$).

Selekcja rankingowa (rangowa) (ang. *ranking*) – elementy bieżącej populacji są uszeregowane według ich jakości i każdemu z nich nadawana jest *ranga* odpowiadająca numerowi porządkowemu w szeregu. Prawdopodobieństwo wyboru osobnika na rodzica jest zależne od jego rangi, a nie od przystosowania. W selekcji rankingowej prawdopodobieństwo selekcji osobnika jest określone w jawnej postaci funkcyjnej⁸.

Selekcja elitarna (ang. *elitist*) – elitaryzm, w którym najlepsze rozwiązanie jest automatycznie kopiowane do kolejnej generacji, jest często dodawany do różnych schematów selekcji⁹. Wykorzystywana w pracy selekcja elitarna polega na

⁷ Jeżeli potraktować populację potomną strategii ewolucyjnej (μ, λ) jako populację bazową, to sposób selekcji w algorytmie (μ, λ) odpowiada selekcji progowej.

⁸ Najczęściej są to liniowe lub wykładnicze funkcje prawdopodobieństwa, zależne od rangi [3, 77, 139].

⁹ W niektórych pracach elitaryzm traktowany jest jako jeden ze schematów zastępowania (sukcesji) (patrz schemat na rys. 2.1) [3], a do populacji potomnej przechodzi nie tylko najlepszy osobnik, lecz pewien zbiór osobników o najwyższych przystosowaniach.

dobudowaniu najlepszego bieżącego osobnika do populacji bazowej złożonej z $(m - 1)$ -elementów wybranych w procesie selekcji proporcjonalnej.

2.2.2. Mutacja

Mutacja jest operatorem, który działa na typie pojedynczego osobnika i polega na losowej zmianie wektora cech. W przypadku metod, w których osobniki mają reprezentację binarną, operator mutacji z założonym *prawdopodobieństwem mutacji* p_m neguje każdy bit z wektora cech. W przestrzeniach rzeczywistych mutacja polega na dodaniu do każdej współrzędnej wektora cech realizacji zmiennej losowej o zadanym rozkładzie. Osobnik $\mathbf{x}$ w generacji $(i + 1)$, będący potomkiem osobnika $\tilde{\mathbf{x}}^i$, będzie miał zmodyfikowany typ rodzica o cechach: $x_j^{i+1} = \tilde{x}_j^i + \xi_j$, $j = 1, 2, \dots, n$, gdzie ξ_j jest realizacją jednowymiarowej zmiennej losowej o zadanym rozkładzie.

Dodawane zmienne mają często rozkład normalny $\mathcal{N}(0, \sigma)$, a odpowiadający im operator zwany jest *mutacją gaussowską*, z jedynym parametrem – odchyleniem standardowym mutacji σ . W niektórych algorytmach (np. strategiach ewolucyjnych) odchylenie standardowe mutacji jest różne dla kolejnych współrzędnych przestrzeni poszukiwań $\sigma = (\sigma_1, \sigma_2, \dots, \sigma_n)$ i, dodatkowo, może podlegać zmianom lub samoadaptacji w trakcie procesu poszukiwań. Główną właściwością mutacji gaussowskiej jest generowanie potomków w niewielkiej odległości od rodzica w przestrzeniach poszukiwań o niskiej wymiarowości. Odległość ta zwiększa się jednak wraz ze wzrostem wymiarowości przestrzeni [153]. W ostatnich latach wykorzystuje się także mutację z rozkładem *Cauchy’ego* [153, 216], który ma bardziej spłaszczone i szersze „ogony” niż rozkład gaussowski, przez co zwiększa się prawdopodobieństwo dużych zmian cech, czyli makromutacji.

2.2.3. Operatory rekombinacji (krzyżowania)

Do przeprowadzenia operacji rekombinacji (krzyżowania)¹⁰ potrzebny jest więcej niż jeden (najczęściej dwa) element rodzicielski, wybrany w procesie selekcji. Pierwsze dwa opisane poniżej rodzaje krzyżowania dotyczą zarówno kodowania binarnego, jak i rzeczywistoliczbowego (choć głównie stosowane są w tym pierwszym przypadku). Rekombinacje uśredniające stosuje się dla wektorów rzeczywistych.

Krzyżowanie jedno- i wielopunktowe (ang. *one- and multi-point crossover*). Typy potomne składane są z cech (genów) rodziców. Typ jednego lub dwóch potomków jest wynikiem wymiany fragmentów wektorów cech rodzicielskich rozpoczynających się w losowo wybranym punkcie (jednopunktowe) lub punktach (wielopunktowe) wektora.

¹⁰ W terminologii metod ewolucyjnych terminy *rekombinacja* i *krzyżowanie* stosuje się zamiennie, choć czasem termin *krzyżowanie* ogranicza się do metod z kodowaniem binarnym.

Krzyżowanie równomierne/jednorodne (ang. *uniform crossover*). Każda cecha potomka jest wylosowana, z jednakowym prawdopodobieństwem, cechą jednego z jego rodziców.

Krzyżowanie uśredniające/arytmetyczne (ang. *intermediate crossover*). Ogólna formuła rekombinująca rzeczywistoliczbowe typy rodziców $\mathbf{x}_1$ i $\mathbf{x}_2$ w typ potomka $\mathbf{x}_\bullet = (x_{\bullet,1}, x_{\bullet,2}, \dots, x_{\bullet,n})$ zadana jest zależnością $x_{\bullet,k} = \alpha x_{1,k} + (1-\alpha)x_{2,k}$, ($k = 1, 2, \dots, n$) i wymaga przyjęcia wartości parametru $\alpha \in [0, 1]$, który czasami jest losowany z rozkładem jednostajnym na tym przedziale. Typy potomków leżą więc na odcinku łączącym pary typów rodzicielskich lub we wnętrzu hiperprostopadłościanu generowanego przez typy rodzicielskie, gdy parametr α przyjmuje potencjalnie różne wartości dla kolejnych współrzędnych wektora $\mathbf{x}_\bullet$. W przypadku najczęściej przyjmowanej wartości $\alpha = 1/2$, cechy potomka są średnią arytmetyczną cech jego rodziców.

Operatory rekombinacji często uważane są za najważniejszy czynnik decydujący o sukcesie metod ewolucyjnych, zwłaszcza algorytmów genetycznych z kodowaniem binarnym¹¹.

2.3. Metody fenotypowe. Metoda poszukiwań ewolucyjnych z miękką selekcją

2.3.1. Modele ewolucji fenotypowej

W zastosowaniach praktycznych oraz w analizie teoretycznej często stosowano metody ewolucyjne (przede wszystkim algorytmy genetyczne), w których osobniki opisywane są przez kodowane binarnie genotypy. Genotypy przekształcane są następnie w fenotypy, które podlegają ocenie środowiska. Na podstawie fenotypów określana jest jakość osobnika. Wybór przestrzeni genotypów jako podstawowej w algorytmach genetycznych spowodowany był początkowo analogią do ewolucji biologicznej, a później przekonaniem, że kodowanie binarne jest najlepsze, zgodnie z tak zwaną *zasadą minimalnego alfabetu* (ang. *principle of minimal alphabets*) [91]. Panowała opinia, że reprezentacja binarna jest w pewnym sensie najlepsza, gdyż używa minimalnego alfabetu, w którym liczba podzbiorów podobieństwa (tzw. schematów, zob. rozdział 3.4) przypadająca na bit informacji jest największa, i który w największym stopniu maksymalizuje liczbę przetwarzanych równolegle schematów (tzw. *implicit parallelism*) [77]¹². Uważano, że typy kodowane przez inny niż

¹¹ Rolę operatorów rekombinacji w metodach ewolucyjnych, analizując głównie ich wpływ na różnorodność populacji, omówiono w rozdz. 6.1.

¹² Jednak Antonisse [2], inaczej niż klasycznie interpretując, schematy pokazał, że jest przeciwnie, i że alfabet o dużej liczebności symboli zawiera więcej schematów niż binarne.

binarny alfabet nie będą odpowiednio reprezentowane w skończonej populacji i konieczne będzie wykorzystanie bardzo licznych populacji [173]¹³. Argumentowano także, że w praktyce problemy rozwiązywane są przez komputery, w których obowiązuje reprezentacja binarna, czyli zadania z innych dziedzin i tak będą w końcu kodowane binarnie. W reprezentacji binarnej łatwiej jest dokonywać niektórych operacji genetycznych, głównie krzyżowania, z którym wiąże się efektywność metod ewolucyjnych. Kodowanie binarne ułatwia też analizę matematyczną tych metod.

Obecnie powszechną praktyką staje się wybór reprezentacji bardziej adekwatny do rozwiązywanego zadania (ang. *problem specific representation*), np. rzeczywistoliczbową, która wydaje się naturalna w zadaniach zdefiniowanych w ciągłych przestrzeniach rzeczywistych [32, 52, 89, 139] czy strojeniu sieci neuronowych [127, 176]; całkowitoliczbową w zadaniach optymalizacji kombinatorycznej [56, 82, 133]; uporządkowanej listy w problemach szeregowania [196, 198]; macierzową w problemach transportowych [138].

W metodach zwanych *fenotypowymi* operuje się bezpośrednio na fenotypach, pomijając etap kodowania. Genotyp jest utożsamiany z fenotypem, który służy bezpośrednio do obliczenia przystosowania osobnika. Za bezpośrednim wykorzystaniem fenotypów przemawia wiele względów:

- przyspieszenie procesu obliczeniowego, nie ma etapu kodowania i dekodowania;
- zwiększona dokładność uzyskiwanych wyników. Dyskretne kodowanie ciągłych dziedzin nigdy nie jest dokładne, co może być kłopotliwe, szczególnie dla wielowymiarowych problemów, w których potrzebna jest duża precyzja obliczeń. Zwiększanie liczby bitów w celu zwiększenia dokładności obliczeń powoduje eksponencjalny (z liczbą bitów) wzrost mocy przestrzeni poszukiwań. Wystarczająco duża dokładność w reprezentowaniu liczb rzeczywistych przez współczesne komputery powoduje, że skończoność zbioru liczb maszynowych nie zmniejsza precyzji obliczeń;
- zmniejszenie wymiarowości przestrzeni poszukiwań w stosunku do zadania kodowanego binarnie;
- zapobieganie redundancji, która może występować, gdy binarnie kodowane są liczby należące do skończonego dyskretnego zbioru o mocy różnej od potęg dwójki;
- często mniejsza liczba generacji potrzebnych do znalezienia rozwiązania o porównywalnej jakości;
- przestrzeń poszukiwań jest określona jedynie przez funkcję celu, podczas gdy w przypadkach kodowania (dyskretnego) topologia przestrzeni poszukiwań jest także zależna od stosowanej metody kodowania (sąsiedztwa indukowanego przez operatory [95]);

¹³ Co wynika z przyjmowanego założenia, że dla efektywności AG każdy możliwy punkt przestrzeni poszukiwań powinien być osiągalny z populacji początkowej tylko za pomocą operatora krzyżowania.

- kodowanie może zmienić strukturę krajobrazu adaptacyjnego, np. mogą pojawić się nowe optima, nieliniowości czy tzw. klify Hamminga¹⁴;
- rozszerzony zbiór dostępnych operatorów mutacji i rekombinacji, opartych na operacjach arytmetycznych i rozkładach prawdopodobieństw;
- możliwość wykorzystania apriorycznej wiedzy o problemie;
- łatwiejsza implementacja niektórych zadań z ograniczeniami;
- większa kompatybilność z innymi metodami optymalizacyjnymi, co ułatwia ich hybrydyzację;
- możliwość strojenia parametrów metody oraz stopniowego ulepszania rozwiązań bieżących dla ciągłych funkcji przystosowania z R^n , gdy małe zmiany zmiennych powodują małe zmiany wartości funkcji.

W metodach fenotypowych dużą grupę stanowią zastosowania, w których funkcja celu jest definiowana explicite w R^n [32, 89, 139, 167, 177, 209]. Porównując kodowanie binarne z rzeczywistoliczbowym, Michalewicz przyznał, że: *reprezentacja zmiennopozycyjna jest szybsza, daje bardziej zbliżone wyniki w różnych przebiegach i prowadzi do lepszej dokładności (w szczególności przy większych dziedzinach, gdzie kodowanie binarne wymaga zbyt długich reprezentacji)* [139]. Jedną z metod fenotypowych z kodowaniem rzeczywistoliczbowym, metoda poszukiwań ewolucyjnych z miękką selekcją, jest przedmiotem badań niniejszej pracy.

2.3.2. Poszukiwania ewolucyjne z miękką selekcją

Darwin w teorii ewolucji stwierdził, że kumulacja wyselekcjonowanych, lepiej dostosowanych osobników w populacji prowadzi w dłuższym czasie do zwiększenia średniego przystosowania populacji [31]. Wzorując się na teorii darwinowskiej, Galar zaproponował adaptacyjną metodę *poszukiwań ewolucyjnych z miękką selekcją*, która działa w przestrzeniach rzeczywistych bez ograniczeń [62, 63, 64]. Wyprowadzono ją z prostego modelu ewolucji fenotypowej, w której główną rolę odgrywają selekcja oraz mutacja. Inspiracją powstania metody stała się sprawność pokonywania pułapek ewolucyjnych i przejścia siodeł przez populacje ewoluujące zgodnie z tym modelem. Główne założenia modelu systemu ewolucyjnego przedstawiono następująco: *Populacja osobników jednego gatunku żyje w stałym, jednorodnym i izolowanym środowisku, zawierającym pewną liczbę identycznych nisz osobniczych. Reprodukacja jest aseksualna, kolejne generacje nie zachodzą na siebie, potencjał reprodukcyjny jest dostateczny, by wszystkie nisze były zapełnione. Typy osobników są określone wartościami kilku ciągłych cech fenotypowych. Od typu zależy stopień przystosowania (fitness) mierzony szansami reprodukcji osobnika: prawdopodobieństwo umieszczenia potomka osobnika danego typu w dowolnej niszy jest proporcjonalne do wartości wskaźnika przystosowania tego typu. Oczekiwane*

¹⁴ Obszary położone bardzo blisko optimum w przestrzeni fenotypowej, lecz bardzo daleko w terminach odległości Hamminga.

wartości cech potomków są takie same jak wartości cech rodzicielskich. Zrealizowane wartości cech potomków różnią się od wartości rodzicielskich o losowe modyfikacje – niezależne i normalnie rozłożone. Rozkład wskaźników przystosowania nad przestrzenią cech określa krajobraz adaptacyjny. W tym krajobrazie typy osobników w populacji tworzą grupę, która w kolejnych generacjach przemieszcza się – ewoluje – w stronę obszarów, które zapewniają wyższe średnie przystosowanie [64].

Model procesu ewolucji sformalizowano jako algorytm optymalizacyjny, który przedstawimy na podstawie schematu z rys. 2.1. Populacja P składa się z m elementów $P = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\}$. Każdy element opisany jest n -elementowym wektorem cech rzeczywistych – typem $\mathbf{x}_k = (x_{k,1}, x_{k,2}, \dots, x_{k,n})$, $k = 1, \dots, m$. Populacja ewoluje w ciągłej i nieograniczonej przestrzeni poszukiwań (typów) $T = R^n$. Jakość k -tego elementu określona jest przez ewaluację rzeczywistej nieujemnej funkcji przystosowania, zadanej na przestrzeni poszukiwań $q: R^n \rightarrow R^+ \cup \{0\}$, dla danego typu: $q_k = q(\mathbf{x}_k)$. Proces ewolucji, przekształcający populację P^i w i -tej generacji w populację potomną P^{i+1} , przebiega w dwóch krokach:

selekcja proporcjonalna (zwana w tym modelu także *miękką*¹⁵) – określa prawdopodobieństwo wylosowania elementu $\mathbf{x}_k^i$ z bieżącej m -elementowej populacji P^i jako rodzica do kolejnej generacji

$$Pr(\mathbf{x}_k^i) = \frac{q(\mathbf{x}_k^i)}{\sum_{j=1}^m q(\mathbf{x}_j^i)}, \quad (2.1)$$

mutacja polegająca na dodaniu do każdej składowej wektora cech rodzica $\mathbf{x}_k^i$ realizacji jednowymiarowej zmiennej losowej ξ o rozkładzie normalnym $\mathcal{N}(0, \sigma)$

$$\mathbf{x}_{\bullet,j}^{i+1} = \mathbf{x}_{k,j}^i + \xi_j, \quad j = 1, \dots, n. \quad (2.2)$$

Cała populacja P^i jest zastępowana, otrzymaną w procesie selekcji i mutacji, populacją potomną. W przedstawionej metodzie są jedynie dwa parametry: rozmiar populacji m oraz odchylenie standardowe mutacji σ . Ponieważ stan generacji kolejnej zależy tylko od generacji bieżącej, generowany proces jest jednorodnym łańcuchem Markowa o nieskończonej liczbie stanów.

Dwutaktowy model opisany prawdopodobieństwami selekcji i mutacji można zapisać łącznie w postaci rozkładu prawdopodobieństwa położenia elementu $\mathbf{x}$ należącego do kolejnej generacji, przy założeniu, że bieżącą generacją jest P^i [97], o funkcji gęstości prawdopodobieństwa (FGP) danej wzorem

$$f_{i+1}(\mathbf{x}|P^i) = \sum_{k=1}^m \left(\frac{q(\mathbf{x}_k^i)}{\sum_{j=1}^m q(\mathbf{x}_j^i)} g(\mathbf{x}, \mathbf{x}_k^i) \right) = \sum_{k=1}^m \alpha(\mathbf{x}_k^i) g(\mathbf{x}, \mathbf{x}_k^i), \quad (2.3)$$

¹⁵ W oryginalnym modelu selekcję proporcjonalną nazwano *miękką* i takie określenie pozostało w nazwie metody. Obecnie za „miękkie” (ang. *soft*) uważa się wszystkie rodzaje selekcji, które z niezerowym prawdopodobieństwem pozwalają na wybór każdego z dopuszczalnych rozwiązań, czyli usytuowane pomiędzy selekcją twardą (ang. *hard*) a czysto losową (ang. *pure random search*).

gdzie $\alpha(\mathbf{x}_k^i)$ opisuje prawdopodobieństwo wyboru elementu $\mathbf{x}_k^i$, a $g(\mathbf{x}, \mathbf{x}_k^i)$ jest rozkładem mutacji k -tego osobnika, $g(\mathbf{x}, \mathbf{x}_k^i) = \mathcal{N}(\mathbf{x}_k^i, \sigma)$. W odróżnieniu od większości metod optymalizacji stochastycznej, gdzie rozkłady są zwykle jednostajne lub co najwyżej jednomodalne [217], rozkład (2.3) może być wielomodalny. Wielomodalność rozkładu może być przyczyną dużych („skokowych”) zmian położenia elementów w kolejnych generacjach.

Badania symulacyjne algorytmu wykazały jego dużą efektywność w pokonywaniu siodła oddzielających optima wielowymiarowych funkcji celu [25, 63, 64, 149]. Dzieje się tak głównie dzięki właściwościom selekcji proporcjonalnej, która daje szansę na reprodukcję każdemu, nawet najgorszemu elementowi. Populacje sprawnie przechodzą bardzo szerokie siodła (dziesiątki razy większe niż σ) oraz takie o wąskich grzbietach. Zwiększenie wymiarowości przestrzeni poszukiwań nie stanowi poważnego problemu dla algorytmu: populacja nadal znajduje optimum globalne, choć w nieco dłuższym czasie. Co więcej, zaobserwowano, że na bardzo szerokich siodłach sprawność przestaje zależeć od wymiaru przestrzeni poszukiwań. Najbardziej efektywne w pokonywaniu siodła są małe, kilkuelementowe populacje. Metoda jest mało wrażliwa na wybór długości kroku poszukiwań σ oraz rodzaj topografii, perturbacje i zmiany kształtu krajobrazu adaptacyjnego [27, 26].

Do algorytmu wprowadzano także rekombinację uśredniającą, gdzie typ potomka posiadał cechy będące średnią cech rodziców [64, 96]. Populacje, w których występuje rekombinacja cech, mają wyższe średnie przystosowanie w okolicy optimum i są bardziej skoncentrowane (różnorodność populacji jest mniejsza). Przechodzą siodła jako zwarta grupa o małym rozproszeniu, a przejście jest tym szybsze, im populacja jest mniejsza.

Metoda poszukiwań z miękką selekcją stanowi raczej algorytm adaptacyjny niż optymalizacyjny. Proces poszukiwań jest procesem zależnym od ścieżki: *Ewolucja nie jest asymptotycznie zbieżna do optimum – sprawność interpolacyjna miękkiej selekcji jest słaba. Ewolucja jest wrażliwa raczej na lokalne relacje bezwzględnych wysokości, niż na lokalne gradienty powierzchni odpowiedzi. Ewolucja prowadzi kolejne generacje prób raczej do wyniesionych obszarów powierzchni odpowiedzi niż do maksimów. Ewolucja zapoczątkowana na wysokim i smukłym wzgórzu jakości może się ustabilizować na wzgórzu niższym, ale szerszym. Ewolucyjna stabilizacja nie jest ostateczna; proces może niekiedy oscylować między sąsiednimi wierzchołkami* [64].

Praktyczne aplikacje metody poszukiwań z miękką selekcją wykazały jej skuteczność w zadaniach optymalizacji – strojenia sieci neuronowych [132, 150, 151], detekcji komórek nowotworowych na obrazach mikroskopowych pochodzących z mikroskopii kontrastowo-fazowej [141], a także modelowaniu procesów rozwoju [65, 67] i zarządzania [66, 68].

2.4. Dynamika poszukiwań ewolucyjnych

W celu lepszego zrozumienia sposobu w jaki działają metody ewolucyjne, opiszemy jakościowo dynamikę przeszukiwania przestrzeni przez ewoluujące populacje, wyjaśniając wcześniej pojęcie krajobrazu adaptacyjnego. Typowe zachowania ewoluujących populacji przedstawiono na podstawie badań symulacyjnych dotyczących metod fenotypowych z kodowaniem rzeczywistoliczbowym [63, 64, 153], a które obserwowano także dla algorytmów z kodowaniem binarnym [145, 146]. Analiza ta uzyska teoretyczne uzasadnienie w dalszej części pracy.

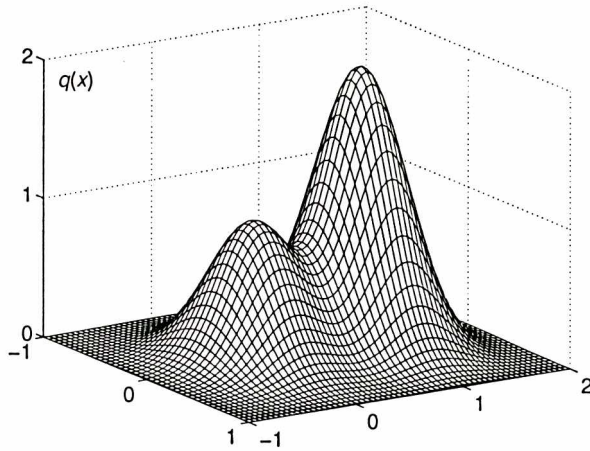
Znajomość dynamiki ewolucyjnej może ułatwić określenie klas funkcji (problemów), dla których wskazane jest ewolucyjne przeszukiwanie przestrzeni poszukiwań. Interesujące jest jednak pytanie o istnienie pewnego uniwersalnego algorytmu umożliwiającego efektywne rozwiązanie praktycznie wszystkich zadań. Jest, lub było, ono stawiane w wielu dziedzinach nauki (np. próba konstrukcji ogólnego rozwiązywacza problemów, ang. GPS – *general problem solver*, w sztucznej inteligencji). Odpowiedzi dostarcza podrozdział 2.4.3, w którym przedstawiono twierdzenie „No free lunch”. Jakkolwiek twierdzenie jest ogólne i dotyczy efektywności wszystkich metod optymalizacyjnych, to często jest cytowane w kontekście metod ewolucyjnych.

2.4.1. Krajobraz adaptacyjny funkcji przystosowania

Aby opisać relacje pomiędzy fenotypami (genotypami) a przystosowaniem Wright [214] zaproponował koncepcję *krajobrazu adaptacyjnego* lub *krajobrazu funkcji przystosowania* (ang. *adaptive topography*, *adaptive fitness landscape*). Krajobraz adaptacyjny jest $(n+1)$ -wymiarową przestrzenią otrzymaną przez rozpięcie funkcji przystosowania nad przestrzenią typów. Każdy typ z przestrzeni T ma swoją reprezentację w krajobrazie, a „wysokość” punktu na krajobrazie odpowiada jego jakości. W przestrzeni fenotypów niewielkie zmiany cech powodują niewielkie zmiany w wartości funkcji przystosowania. W przestrzeni genotypów – niekoniecznie. Definiuje się także pojęcie *sąsiedztwa*, jako zbioru punktów, w których mogą się znaleźć typy potomków danego osobnika w jednym (kilku) krokach działania metody. Krajobraz adaptacyjny wyznacza więc zbiór typów ze zdefiniowaną odległością i sąsiedztwem oraz odpowiadające im wartości przystosowania¹⁶.

W krajobrazach adaptacyjnych można wyróżnić kilka podstawowych „formacji”: *wzgórza* o *grzbietach* różnej stromości, których *wierzchołki* odpowiadają optymalnemu (lokalnie) fenotypowi; *siodła* pomiędzy wzgórzami; *pasma* wzgórz; *plateau* (*plateau*), czyli płaskie rozległe obszary o jednakowej lub podobnej wartości przystosowania. Najbardziej nośna jest chyba metafora Eigena, który po-

¹⁶ Pojęcie krajobrazu jest także wykorzystywane przez fizyków do badania potencjałów lub powierzchni energii, ale w tym przypadku poszukiwane jest minimum funkcji.



Rys. 2.2. Krajobraz adaptacyjny gaussowskiej dwumodalnej funkcji przystosowania. Typowi osobnika opisanego przez współrzędne $\mathbf{x} = (x_1, x_2)$ odpowiada przystosowanie $q(\mathbf{x}) = \exp[-5(x_1^2 + x_2^2)] + 2 \exp[-5(x_1^2 + (x_2 - 1)^2)]$

Fig. 2.2. Evolutionary landscape for a bimodal Gaussian function. The fitness value $q(\mathbf{x}) = \exp[-5(x_1^2 + x_2^2)] + 2 \exp[-5(x_1^2 + (x_2 - 1)^2)]$ corresponds to type of individual described by coordinates $\mathbf{x} = (x_1, x_2)$

wiedział, że taki krajobraz przypomina łańcuchy gór na Ziemi [48]¹⁷. Analizując krajobrazy adaptacyjne, można rozpatrywać różne właściwości: liczbę optimów, ich wysokości i rozłożenie, średnie różnice w jakości pomiędzy poszczególnymi punktami zwane strukturą korelacyjną krajobrazu. Na rysunku 2.2 przedstawiono przykładowy krajobraz adaptacyjny dla funkcji przystosowania będącej sumą dwóch funkcji Gaussa w dwuwymiarowej przestrzeni poszukiwań.

Efektywność poszukiwań ewolucyjnych zależy od topografii krajobrazu adaptacyjnego. Płaskowizne mogą stanowić problem dla metod ewolucyjnych, ponieważ taki krajobraz nie zawiera żadnej informacji, która ukierunkowałaby poszukiwania. Nie działa wtedy selekcja, a procedury ewolucyjne zachowują się podobnie jak błądzenie losowe (ang. *random walk*). Ewolucja preferuje raczej obszary o wysokich przystosowaniach, niż pojedyncze punkty o bardzo wysokiej jakości (np. funkcje typu „igła w stogu siana”, ang. *the needle-in-the-haystack*), gdzie niewielka zmiana typów wywołana na przykład mutacją nie powoduje drastycznej zmiany w jakości rozwiązania. Krajobrazy adaptacyjne różnią się też gładkością. Gładkie krajobrazy (ang. *smooth*) o niewielkiej liczbie szczytów i niewielkich średnich różnicach w jakości sąsiednich punktów są zadaniami łatwymi dla metod ewolucyjnych. Efektywność ewolucji w krajobrazach, gdzie wiele lokalnych szczytów jest otoczo-

¹⁷ Obserwacje te potwierdzają analizy krajobrazów w złożonych zadaniach typu problem komiwojażera [19], podziały grafów [137] czy szeregowanie [174]. Optima są tam zwykle rozmieszczone w skupiskach zwanych *central massif*.

nych przez głębokie doliny i różnice w jakości sąsiednich punktów są duże (ang. *rugged*), zależy od stopnia pofałdowania i jest stosunkowo duża dla nieznacznych nierówności a mniejsza dla funkcji, w których lokalne szczyty przypominają funkcje „igłowe” [93]. Stosunkowo duża odporność (ang. *robustness*) metod ewolucyjnych na niedokładności określenia funkcji celu może być wielką zaletą w praktycznych zastosowaniach, gdzie wartości przystosowania są zaszumione lub mierzone z niewielką dokładnością [20].

Przedstawiane zwykle jedno- lub dwuwymiarowe krajobrazy funkcji przystosowania, to przykłady akademickie, gdyż w rzeczywistości przestrzenie poszukiwań są wielowymiarowe, a funkcja celu często nie jest dana w postaci analitycznej, lecz może być wynikiem pomiarów, dodatkowych wyliczeń numerycznych lub zmienną procesu technologicznego. Idea krajobrazu adaptacyjnego jest jednak bardzo pomocna w intuicyjnym zrozumieniu działania zarówno ewolucji, jak i metod optymalizacyjnych, a szczególnie dynamiki populacji.

2.4.2. Dynamika populacji

Zdefiniowanie pojęcia krajobrazu adaptacyjnego ułatwia opis dynamiki ewoluującej populacji. Rozpatrzmy populację znajdującą się na zboczu wzgórza w krajobrazie adaptacyjnym, dla którego funkcja przystosowania posiada jedno optimum. Mutacja i rekombinacja produkują pulę nowych typów, z których selekcja wybiera zwykle te lepiej przystosowane. W rezultacie populacja pnie się ku lokalnemu optimum, przy czym średnia jakość populacji rośnie¹⁸. Kiedy populacja osiągnie otoczenie szczytu, znajduje się wokół niego w stanie tzw. *quasi-równowagi* pomiędzy selekcją, utrzymującą populację blisko optimum, a losowymi perturbacjami spowodowanymi mutacją czy rekombinacją. Oczekuje się, że rozkład populacji wokół szczytu jest w przybliżeniu normalny [96, 169]¹⁹. Populacja przebywa w stanie quasi-równowagi zwykle znacznie dłużej niż trwało osiągnięcie szczytu. Selekcja, oprócz wzrostu średniego przystosowania, przyczynia się także do zmniejszenia różnorodności populacji, która staje się zbiorem punktów o bliskich typach. Mutacja, wprowadzając błędy w reprodukcji, zapobiega ujednolicaniu się populacji. W rezultacie działania tych mechanizmów osobniki tworzą zwartą grupę – *klaster*, która porusza się w krajobrazie adaptacyjnym. Taki zbiór blisko powiązanych mutantów, zwykle znajdujących się wokół optymalnego typu, zwany jest często *quasi-species* (quasi-gatunkiem) [47, 49].

W krajobrazach z wieloma wzgórzami znalezienie innego optimum może być wynikiem serii kolejnych drobnych zmian typów lub „skoku” w obszar przycią-

¹⁸ Proces wzrostu średniej jakości populacji niekoniecznie jest monotoniczny. Mutacje i rekombinacje mogą powodować chwilowe zastoje lub obniżenie przystosowania.

¹⁹ Potwierdzają to wyniki badań doświadczalnych i analitycznych w genetyce populacyjnej [128, 135].

gania innego szczytu, spowodowanego rekombinacją genów lub makromutacją. Jeżeli presja selekcji jest zbyt duża (np. selekcja twarda), osobniki o niższym niż optymalne przystosowaniu nie będą przeżywały, a populacja znajdzie się w tzw. *pułapce ewolucyjnej*. W tym przypadku znalezienie kolejnego optimum może się zdarzyć tylko w wyniku skoku. Jeżeli selekcja jest łagodniejsza, mutacja może spowodować obniżenie jakości osobników i przesunięcie populacji z optimum w dół wzgórza. Ciąg „niekorzystnych” lokalnie kroków, zmniejszających bieżące średnie przystosowanie populacji, może doprowadzić osobniki w pobliże siodła pomiędzy wzgórzami. Dzięki topologii siodła (wyższe przystosowania w kierunku kolejnego optimum) populacja może przejść jego grzbietem na inne wzgórze. Możliwe jest więc opuszczenie lokalnego wierzchołka (wyjście z pułapki), przekroczenie siodła i znalezienie kolejnego wzniesienia, być może optimum globalnego [58, 62, 63, 64]. Czasy przejść pomiędzy optimami są znacznie krótsze niż czasy fluktuacji wokół szczytu.

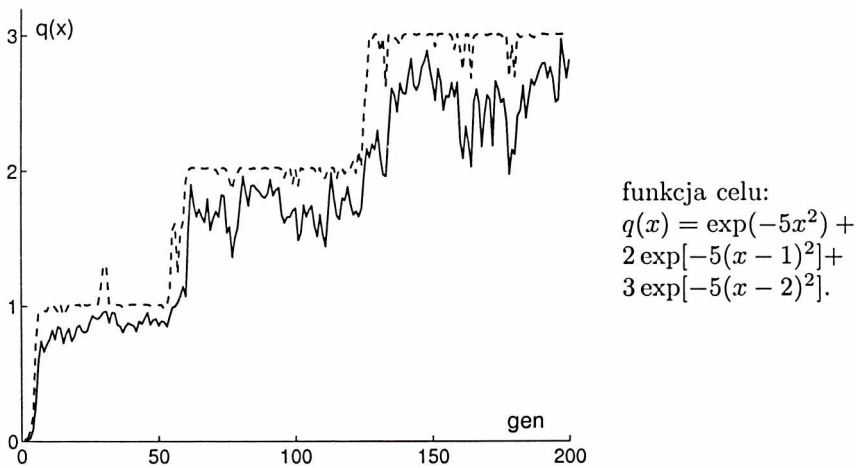
Takie dłuższe okresy zastoju i krótkie okresy szybkich zmian można zauważyć, analizując zachowanie się średniego przystosowania populacji ewoluującej w wielomodalnych krajobrazach adaptacyjnych (tzw. *epochal evolution*) [84, 130]. Na rysunku 2.3 przedstawiono przykładowe wykresy średniego oraz najlepszego w generacji przystosowania dla 8-elementowej populacji ewoluującej z selekcją proporcjonalną i mutacją gaussowską w krajobrazie adaptacyjnym dla funkcji celu będącej sumą trzech funkcji gaussowskich o różnych wysokościach²⁰. Wartość przystosowania niewiele się zmienia w okresach *quasi-stabilnego* trwania populacji wokół znajdującego optimum, zwanych też *epokami fitness* lub *fazą latentną*. Faza latentna może trwać bardzo długo w porównaniu z okresami nagłych i szybkich wzrostów średniego przystosowania zwanych *innowacjami* lub *fazą aktywną* [64, 146, 153]. W czasie innowacji populacja wspina się na nowo odkryte wzgórze i jej średnie przystosowanie rośnie²¹. Jeżeli w krajobrazie adaptacyjnym znajdują się inne optima, to cykl *faza latentna* → *faza aktywna* może się powtarzać. Długość trwania okresów, czas pojawienia się innowacji jest zjawiskiem losowym i zależy między innymi od parametrów procesu ewolucji. W biologii zjawisko ewolucji zachodzącej w cyklach stabilności następujących po okresach nagłych zmian znane jest jako *shifting balance evolution* [215] lub *punctuated equilibria* [50].

W dynamice poszukiwań ewolucyjnych wyróżnia się więc dwie przeciwstawne tendencje:

eksploracja oznaczająca przeszukiwanie przestrzeni poszukiwań w celu znalezienia obszarów o lepszej jakości, czyli wychodzenie na zewnątrz obszaru przyciągania zlokalizowanego już optimum, przeszukiwanie globalne;

²⁰ Funkcje celu wykorzystywane w pracy opisano w Dodatku.

²¹ Może się także zdarzyć, że kolejne znalezione przez populację wzgórze jest niższe niż poprzednie i wtedy średnie przystosowanie populacji położonej w nowym optimum jest mniejsze niż przystosowanie na poprzednio znalezionym wzgórzu.



Rys. 2.3. Średnie (linia ciągła) oraz najlepsze (linia przerywana) przystosowania populacji ewoluującej z selekcją proporcjonalną i mutacją gaussowską ($\sigma = 0,2$, $m = 8$)
 Fig. 2.3. Average (solid line) and the best (dashed line) fitness of a population evolving with proportional selection and Gaussian mutation ($\sigma = 0.2$, $m = 8$)

eksploatacja – poszukiwanie najlepszej wartości w zlokalizowanym już obszarze przyciągania optimum, czyli wspinanie się na wzgórze, przeszukiwanie lokalne. Znalezienie równowagi między eksploracją a eksploatacją jest ważne dla zapewnienia efektywnego działania metod ewolucyjnych. Eksploacyjny nacisk selekcji jest równoważony przez eksploracyjny nacisk mutacji i rekombinacji, a ich wielkości (ustawiane przez odpowiedni dobór rodzaju selekcji i parametrów operatorów reprodukcji) mogą regulować wpływy poszczególnych tendencji. Szczególnie istotny wydaje się być wybór wielkości mutacji, czyli odpowiedniego zasięgu mutacji dla metod z kodowaniem rzeczywistoliczbowym i prawdopodobieństwa mutacji p_m dla kodowania binarnego. Mutacja zapewnia, że każdy punkt w przestrzeni poszukiwań może być przeszukany. Dostatecznie duże mutacje są potrzebne dla podtrzymania lub wprowadzenia różnorodności do populacji. Jednak zbyt duże mutacje mogą przekształcić proces w czysto przypadkowy, przypominający błędzenie losowe. Poszukiwanie z bardzo małym krokiem (prawdopodobieństwem) mutacji upodabnia proces optymalizacyjny do metody lokalnego wspinania, powoduje jego większą wrażliwość na zakłócenia funkcji celu i uniemożliwia adaptację w przypadku dynamicznie zmieniających się funkcji celu. Znalezienie optymalnej wielkości mutacji jest trudne, ponieważ procesy te zależą także od rodzaju selekcji, reprezentacji, innych operatorów oraz funkcji przystosowania. Najczęściej proponuje się wartości otrzymane empirycznie [7] lub heurystyki [90]. Jednym z pomysłów może być zmiana zasięgu mutacji w trakcie poszukiwań (np. samoadaptacja w strategiach ewolucyjnych [8, 46]).

2.4.3. Twierdzenie *No free lunch*

Twierdzenie znane pod nazwą *No free lunch theorem* (NFLT) [207] wskazuje, że w ogólnym przypadku, dla wszystkich możliwych funkcji celu rozpiętych na przestrzeni poszukiwań (dyskretnej i skończonej), nie można wskazać metody, która jest lepsza od wszystkich innych metod. Średnie jakości rozwiązań uzyskane daną metodą, uśrednione po zbiorze wszystkich możliwych funkcji celu i dla każdej miary efektywności metody, która nawiązuje do liczby różnych odwiedzonych punktów przestrzeni, są identyczne dla każdego algorytmu. Jeżeli procedura optymalizacyjna jest efektywna dla ustalonego problemu, to prawdopodobnie będzie nieodpowiednia dla innego. Nie ma uniwersalnie efektywnego algorytmu, a metodę należy dobierać uwzględniając specyfikę rozwiązywanego problemu. Procedury optymalizacyjne są zwykle kierowane do pewnej szczególnej klasy zadań i dla niej mogą być najlepsze, ale mogą nie sprawdzić się dla innych problemów. Wolpert i Macready [207] zaproponowali nawet miarę dopasowania algorytmu do danego problemu.

NFLT wywołało wiele dyskusji w środowisku [40, 166]. Adwersarze argumentowali, że zbiór wszystkich możliwych funkcji celu dla dyskretnej i skończonej przestrzeni poszukiwań składa się głównie z całkowicie „chaotycznych” funkcji, w których wartości przystosowania sąsiadujących punktów nie są skorelowane. Tak więc problem, z którym dany algorytm sobie nie radzi, może być bardzo abstrakcyjny i niespotykany w rzeczywistości. Zarzucano, że twierdzenia opracowane zostały dla algorytmów, które mogą tylko raz odwiedzić dany punkt w przestrzeni poszukiwań, co nie zachodzi w przypadku metod ewolucyjnych. Ponadto twierdzono, że w praktyce nie są potrzebne metody, które działają efektywnie dla wszystkich możliwych funkcji, a jedynie na pewnym ich podzbiórze, który wynika z rzeczywistych potrzeb. W odpowiedzi na krytykę zaprezentowano generalizację twierdzenia dla procedur, które mogą wielokrotnie odwiedzać te same punkty [166]. Dowiedziono także wersji twierdzenia, mówiącej, że NFLT zachodzi dla każdego podzbioru rozpatrywanych funkcji celu, jeśli tylko zbiór ów jest zamknięty ze względu na permutacje (tzw. *block uniform functions*) [51].

Z twierdzenia NFLT wynika, że funkcje przystosowania, które nie mają żadnych regularności, są trudne do optymalizacji dla każdego algorytmu. Przedstawiona w poprzednim podrozdziale dynamika ewoluujących populacji może pomóc w określeniu (bardzo ogólnym) klasy funkcji przystosowania, dla których algorytmy ewolucyjne mogą działać stosunkowo efektywnie. Generalnie, są nimi funkcje przypominające, opisane wcześniej, „eigenowskie” łańcuchy gór, z wyraźnymi zboczami, wierzchołkami i siodłami.

3. Teoria metod ewolucyjnych

Do analizy teoretycznej metod ewolucyjnych stosuje się klasyczne teorie matematyczne (np. łańcuchy Markowa, układy dynamiczne), metody fizyki statystycznej, genetyki populacyjnej, a także specjalnie opracowane teorie (np. schematów). Większość modeli teoretycznych dotyczy algorytmów genetycznych z kodowaniem binarnym, głównie ze względu na ich popularność oraz łatwiejszą analizę przestrzeni dyskretnych i kodowania binarnego. Wyniki uzyskane dla AG przenosi się na inne algorytmy ewolucyjne (np. programowanie genetyczne), a niektóre rozszerza na ogólne metody heurystyczne losowego przeszukiwania przestrzeni.

Pomimo wielu podejmowanych wysiłków, uzyskane dotychczas rezultaty analizy teoretycznej metod ewolucyjnych nie są w pełni zadowalające. Stochastyczny charakter metod, skomplikowane zależności i wielowymiarowość zadań powodują, że otrzymane wyniki są zwykle niepełne. Do tej pory nie opracowano efektywnych teorii matematycznych dla tak złożonych zagadnień (np. teorii procesów Markowa z ciągłą i nieskończoną przestrzenią stanów). Najczęściej buduje się modele aproksymacyjne, dotyczące ograniczonej klasy algorytmów, operatorów, funkcji celu, czy obłożone restrykcyjnymi lub nierealnymi założeniami (np. nieskończone populacje). Istniejące dokładne modele markowskie można stosować tylko w odniesieniu do problemów o małych rozmiarach i szczególnych parametrach i operatorach. Teorie opisujące zachowanie algorytmu w jednym kroku, takie jak teoria schematów, nie dają wiedzy o zbieżności, zachowaniach asymptotycznych czy prawdopodobieństwie sukcesu. Próbowano także opracować wspólną teorię obliczeń ewolucyjnych w ogólności, co dodatkowo utrudnia obszerność i różnorodność dziedziny [143, 167].

W rozdziale przedstawiono teorie powiązane z tematem dysertacji: sposoby analizy metod z kodowaniem rzeczywistoliczbowym, modele wykorzystujące łańcuchy Markowa, teorię układów dynamicznych. Inne modele, ważne ze względu na rozwój dziedziny, opisano pokrótce w podrozdziale 3.4.

3.1. Analiza metod z kodowaniem rzeczywistoliczbowym

Analiza matematyczna metod fenotypowych z kodowaniem rzeczywistoliczbowym jest trudniejsza niż metod z kodowaniem binarnym i przestrzeniami dyskretnymi. Wykorzystywane w tych algorytmach sposoby mutacji opisywane są ciągłymi rozkładami prawdopodobieństwa, algorytm jest zatem modelowany jako proces losowy, w którym zmienne losowe mają określony rozkład prawdopodobieństwa. Do analizy ewolucji rozkładów w czasie używa się np. modelu

dyskretnego procesu stochastycznego [4, 163, 164] lub zakłada się, że mutacja generowana jest przez dyskretne zmienne losowe, dzięki czemu rozkład populacji może przyjmować skończoną liczbę stanów w każdej generacji [87].

Analizując ewolucję jako proces stochastyczny, którego stan należy do wielowymiarowej przestrzeni euklidesowej, Qi i Palmeri rozpatrywali bardzo duże (dążące do nieskończoności) populacje, co ułatwiło analizę matematyczną zachowania zespołowego populacji [163, 164]. Populacje opisywano sekwencją rozkładów prawdopodobieństwa, charakteryzujących rozkład całej populacji. Populacje nieskończone pokrywają całą przestrzeń poszukiwań, a rozkład gęstości charakteryzuje częstość występowania elementów przestrzeni poszukiwań w populacji. Badano zmiany funkcji gęstości nieskończonych populacji poddanych działaniu operatorów genetycznych. Pokazano, że selekcja powoduje koncentrację osobników w regionie o wyższym niż średnie przystosowanie¹. Badając zbieżność udowodniono, że selekcja wraz z mutacją mogą znaleźć maksimum funkcji celu, jeżeli zasięg mutacji jest dostatecznie mały. Podano warunki konieczne dla monotonicznej zbieżności średniego przystosowania dla klasy funkcji lipszycowskich, która nie gwarantuje jednak zbieżności do optimum globalnego. Zasugerowano schemat adaptacji mutacji w trakcie procesu poszukiwań oparty na danych z populacji bieżącej. Pokazano, że rekombinacja równomierna nie wpływa na wartość oczekiwaną i wariancję rozkładu populacji, lecz prowadzi do zaniku korelacji pomiędzy współrzędnymi przestrzeni poszukiwań, pozostawiając niezmiennione rozkłady brzegowe każdej współrzędnej. Rekombinacja powoduje także wzrost entropii populacji, utrzymując wartość wariancji populacji na poziomie wariancji populacji początkowej. Badając łączny efekt operatorów selekcji, mutacji i rekombinacji, podano formalne uzasadnienie roli rekombinacji w poszukiwaniach globalnych: dany obszar przestrzeni poszukiwań jest próbkowany częściej, jeżeli jakkolwiek hiperpłaszczyzna zawierająca ten rejon daje wyższe średnie przystosowanie.

Metody z kodowaniem rzeczywistoliczbowym były analizowane także za pomocą teorii schematów (patrz rozdz. 3.4). Wright [209] podał twierdzenie o schematach dla algorytmów genetycznych z chromosomami kodowanymi liczbami rzeczywistymi, krzyżowania jednopunktowego i selekcji proporcjonalnej, korzystając z ciągłej wersji schematów, tzw. *connected schemata*. Pokazał, że w takim przypadku prawdopodobieństwo mutacji powinno być znacznie większe niż dla kodowania binarnego. Do badań problemów z parametrami ciągłymi posłużyły także schematy uogólnione przez Radcliffa [168], tzw. *formae*. Przedstawiono zasady konstrukcji operatorów rekombinacji efektywnych dla problemów określonych w ciągłych dziedzinach rzeczywistych. Zaproponowano nową wersję mutacji zwaną *end point mutation*, pozwalającą na ponowne pojawienie się w populacji

¹ Rezultat ten może być analogiem dla przestrzeni ciągłych twierdzenia o schematach, które mówi, że oczekiwana wartość polepszenia rozwiązania rośnie z generacji na generację (zobacz rozdz. 3.4).

typów o ekstremalnych wartościach, które operatory rekombinacji gubią. Inny typ schematów dla kodowania rzeczywistoliczbowego, tzw. *interval-schemata*, wykorzystano w pracy [52]. Badano operatory krzyżowania zachowujące i eksplorujące ten rodzaj schematów, oraz pokazano, że np. rekombinacja arytmetyczna i BLX² produkują potomków należących do tych samych schematów co rodzice.

Dla strategii ewolucyjnych analizowano głównie asymptotyczne zachowania czasowo-lokalne (w jednym kroku algorytmu) modeli $(1, \lambda)$ i (μ, λ) [14, 170]. Dynamikę ewolucyjną badano, korzystając ze *statystyk pozycyjnych* prostych (także wielowymiarowych) funkcji celu: liniowych, kwadratowych (tzw. model sferyczny) oraz funkcji barierowych (tzw. model korytarzowy). Estymowano zachowania populacji pod działaniem operatorów genetycznych za pomocą dwóch lokalnych miar: *tempa postępu* (ang. *progress rate*), definiowanego jako oczekiwana zmiana odległości centrum masy populacji od optimum globalnego w funkcji czasu, oraz *wzrostu jakości* (ang. *quality gain*) jako wartości oczekiwanej polepszenia jakości w kolejnych generacjach. Analityczną postać współczynnika tempa postępu otrzymano dla wypukłych funkcji celu i wykorzystano w analizie zbieżności. Dla algorytmu (μ, λ) w krajobrazie funkcji kwadratowej pokazano, że populacja nie śledzi lokalnego gradientu, jak wcześniej sądzono, lecz dokonuje przeszukiwania losowego przestrzeni. Badania rozszerzono na funkcje przystosowania z zakłóceniami, algorytmy z rekombinacją oraz samoadaptacją [13, 15].

Metody wykorzystujące samoadaptację mutacji są jeszcze trudniejsze do analizy, ponieważ generują niejednorodne łańcuchy Markowa, a ewolucja procesu opisywana jest za pomocą równań Chapmana–Kolmogorowa. Dokonuje się więc pewnych uproszczeń i rozwiązuje przybliżone równania dynamiki. Rudolph [180] pokazał, że dla funkcji niewypukłych każda metoda samoadaptacji prowadzi do zbieżności globalnej, pod warunkiem, że selekcja jest elitarna, a reguły adaptacyjne nie naruszają ergodycznych właściwości procesu z mutacją.

3.2. Modele, w których stosowane są łańcuchy Markowa

Do analizy teoretycznej algorytmów genetycznych z kodowaniem binarnym zaproponowano zastosowanie teorii procesów Markowa, co pozwoliło na modelowanie zachowań metod, bez zbędnych uproszczeń i aproksymacji. Nix i Vose [148] przedstawili kompletny (z krzyżowaniem i mutacją) i dokładny (bez uproszczeń) model markowski prostego AG³. Później opracowano modele dla innych wersji algo-

² Rekombinacja opracowana dla kodowania rzeczywistoliczbowego, gdzie cechy potomków są losowo wybierane z odcinka utworzonego przez cechy rodziców.

³ Łańcuchy Markowa wykorzystywano także wcześniej do analizy niektórych aspektów działania algorytmów genetycznych [33, 83].

rytmów [34, 181, 182, 195, 208] i rozszerzono model podstawowy dla reprezentacji całkowitoliczbowej z alfabetu o określonej mocy [123].

Realizacje algorytmów ewolucyjnych w czasie są procesami stochastycznymi. Wszystkie możliwe wartości, jakie mogą przyjmować zmienne losowe opisujące kolejne generacje, nazywamy *stanami*. Prawdopodobieństwo zmiany stanów w kolejnych iteracjach procesu jest opisywane przez odpowiednie rozkłady prawdopodobieństwa, które mogą ewoluować w czasie⁴. Ponieważ kolejna generacja zależy od generacji bieżącej, jest to proces Markowa, a sekwencja populacji w czasie tworzy jednorodny łańcuch Markowa [94]. Korzystając z teorii łańcuchów Markowa, można określić długoterminowe zachowanie procesu. Analizuje się trajektorie procesu, bada ich zachowania asymptotyczne oraz rozkłady w stanie ustalonym.

Rozpatrując algorytmy ewolucyjne w terminach łańcuchów Markowa, przestrzeń stanów definiuje się jako zbiór wszystkich możliwych populacji. Ponieważ dla klasy algorytmów genetycznych z kodowaniem binarnym, które są najczęściej analizowane w terminach łańcuchów Markowa, przestrzeń poszukiwań jest dyskretna i skończona, zwykle populacje reprezentowane są przez wektor częstości występowania w niej poszczególnych typów elementów w przestrzeni. Wymiary wektora częstości jest równa rozmiarowi przestrzeni poszukiwań, a reprezentacja populacji jest niezależna od jej liczności. Poszczególne mechanizmy reprodukcji (selekcja, mutacja i rekombinacja) definiują prawdopodobieństwa przejścia w przestrzeni stanów (jądro markowowskie). Ułatwieniem w badaniach jądra jest możliwość jego dekompozycji na niezależne jądra odpowiadające poszczególnym operatorom. Dla skończonych przestrzeni stanów jądro markowowskie jest opisane *macierzą przejścia* [181], która zawiera prawdopodobieństwa różnych możliwych zmian dyskretnych stanów. Analiza macierzy przejścia pokazuje, że dla algorytmów wykorzystujących tylko selekcję, wszystkie jednorodne populacje, posiadające kopie tylko jednego typu, są stanami *absorbującymi*. Tak więc kodowany binarnie AG (bez mutacji i krzyżowania) jest zbieżny do populacji jednorodnej, nie zawsze zawierającej typy globalnie, a czasem nawet lokalnie, optymalne. Graniczne zachowanie algorytmu jest zależne od stanu początkowego, a zbieżność do optimum globalnego nie jest gwarantowana.

Dodanie do selekcji mutacji powoduje, że macierz przejścia staje się *regularna*⁵, nie ma już stanów absorbujących, a łańcuch Markowa jest ergodyczny [94]. Dla regularnych macierzy przejścia, z niezerowym prawdopodobieństwem, algorytm z populacją o skończonej liczebności odwiedzi wszystkie stany. Zgodnie z twierdzeniem ergodycznym [54] asymptotyczne zachowanie łańcucha ergodycznego jest opisane przez stacjonarny rozkład graniczny nad wszystkimi możliwymi stanami

⁴ Na przykład gdy stosuje się zmienne w czasie prawdopodobieństwa mutacji, rekombinacji lub niejednorodny sposób selekcji [34].

⁵ Regularność macierzy powoduje, że wszystkie stany mogą się ze sobą komunikować w skończonej liczbie kroków.

(ang. *steady state distribution*). W granicy prawdopodobieństwo, że populacja jest w pewnym stanie, nie zależy więc od jej stanu początkowego, a jedynie od funkcji celu i parametrów badanego algorytmu. Pokazano [204], że dla wzrastającej liczebności, populacja przez większość czasu (mierzonego liczbą iteracji) znajduje się w stanie odpowiadającym punktowi stałemu modelu markowskiego, jeżeli jest jedynym takim punktem. Jeżeli punktów stałych jest więcej, to można określić gdzie, z dodatnim prawdopodobieństwem, znajdują się punkty graniczne stacjonarnego rozkładu granicznego. W przypadku nieskończonych populacji i algorytmu działającego tylko z selekcją i mutacją, istnieje niezależny od stanu początkowego jednoznaczny rozkład graniczny wektora częstości, do którego proces jest zbieżny [85]. Ergodyczność łańcucha, dla dodatniego prawdopodobieństwa mutacji, determinuje asymptotyczną gwarancję sukcesu oraz zbieżność globalną [183]. Krzyżowanie nie zmienia typu macierzy przejścia, która nadal jest regularna, bez stanów pochłaniających, a więc nie przynosi znaczących korzyści w zwiększaniu efektywności algorytmu.

Kanoniczny AG nie jest zbieżny do optimum, będzie odwiedzał nieskończenie wiele razy wszystkie stany. Zbieżność algorytmu z prawdopodobieństwem jeden może zagwarantować wykorzystanie elitarnego sposobu selekcji [180, 194] lub powolne zmniejszanie do zera prawdopodobieństwa mutacji i krzyżowania oraz zwiększanie presji selekcyjnej [34, 195]. Zmniejszanie prawdopodobieństwa mutacji nie gwarantuje jednak zbieżności do optimum, ale do pewnej homogenicznej populacji. Także wzrost liczebności populacji może zwiększyć prawdopodobieństwo zbieżności procesu do globalnie optymalnej populacji [34].

Modele algorytmów ewolucyjnych wykorzystujące łańcuchy Markowa są dokładne, lecz obliczenia wykonywane na ich podstawie są złożone obliczeniowo [190]. Macierze przejścia są w praktycznych przypadkach bardzo duże, ponieważ liczba stanów rośnie wraz z liczbą cech oraz iteracji. Dodatkowo, większość rezultatów otrzymano przy założeniu nieskończonej populacji [33, 179, 194]. Pewnym rozwiązaniem problemu jest ograniczanie wielkości macierzy przejść przez grupowanie podobnych stanów i definiowanie prawdopodobieństw przejścia pomiędzy tymi grupami [191] lub zastosowanie teorii punktów stałych modeli układów dynamicznych [142]. Proponowano też operowanie na wektorach reprezentujących rozkład funkcji przystosowania, co redukuje rozmiary macierzy przejścia [146].

Oprócz badania zachowań asymptotycznych – szybkości zbieżności do stanów ustalonych czy momentu pierwszego dojścia do stanu [33, 34, 148, 179, 195, 208], markowskie modele algorytmów ewolucyjnych stosowano także do analizy zachowań w stanie nieustalonym [36], badania różnych schematów selekcji [24], analizy rozproszenia populacji oraz zjawiska przedwczesnej zbieżności [129], poszukiwań kryterium zatrzymania algorytmów [6], estymacji parametrów procesu (np. prawdopodobieństwa krzyżowania i mutacji) [171], oraz modelowania specyficznych właściwości AG, takich jak selekcja, dryf genetyczny, niszowanie [37, 83, 92].

Łańcuchy Markowa wykorzystywano także do analizy strategii ewolucyjnych [180] oraz programowania genetycznego [157].

3.3. Modele, w których stosuje się teorię układów dynamicznych

Rozszerzeniem markowowskich modeli algorytmów genetycznych z kodowaniem binarnym (szczególnie [148]) są modele wykorzystujące teorię układów dynamicznych. Nie analizuje się w nich zachowania populacji, lecz ich wartości oczekiwane, podając dokładne rozkłady prawdopodobieństwa wykorzystywane w wyborze osobników do kolejnej generacji. Vose wraz ze współpracownikami zaproponowali model populacji ewoluującej w skończonej dyskretnej przestrzeni poszukiwań Ω o n stanach [203, 204, 205]. Przestrzeń Ω zawiera wszystkie możliwe ciągi kodowe o określonej długości. Podobnie jak w modelu markowowskim, populacja reprezentowana jest przez n -wymiarowy *wektor populacji* $\mathbf{p}$ opisujący rozkład częstości występowania poszczególnych elementów przestrzeni Ω w populacji. Wektory populacji odpowiadają punktom n -wymiarowego simpleksu jednostkowego⁶. Każdy wektor $\mathbf{p}$ dla skończonej populacji leży w simpleksie, ale nie wszystkie punkty w simpleksie odpowiadają skończonym populacjom⁷. Ewolucję populacji można rozważać jako trajektorię wektora $\mathbf{p}$ w simpleksie.

Proces ewolucji jest definiowany przez kolejne iteracje *operatora heurystycznego* $\mathcal{G}$. Operator $\mathcal{G}(\mathbf{p})$ opisuje rozkład prawdopodobieństwa nad wszystkimi populacjami kolejnej generacji i reprezentuje wartość oczekiwaną kolejnego wektora populacji. Następną generacją będzie miała wektor populacji równy rzeczywiście $\mathbf{p}^{i+1} = \mathcal{G}(\mathbf{p}^i)$, jeśli populacja będzie nieskończona ($r \rightarrow \infty$), a selekcja nieelitarna. Dla nieskończonych populacji proces staje się deterministyczny i jego trajektoria $\mathbf{p}, \mathcal{G}(\mathbf{p}), \mathcal{G}^2(\mathbf{p}), \dots$ w simpleksie może być analizowana jako klasyczny układ dynamiczny. Dla rzeczywistych, skończonych populacji obserwowane będą stochastyczne odchylenia od tej trajektorii. Opisany model nazywany jest w literaturze *modelem nieskończonej populacji* (ang. *infinite population model*) i dla niego uzyskano większość rezultatów [203].

Dla nieskończonych populacji odwzorowanie $\mathcal{G}(\mathbf{p})$ operujące na wektorze $\mathbf{p}$ definiuje *dyskretny układ dynamiczny* w simpleksie. Szczegółowa interpretacja operatora heurystycznego zależy od przyjętych mechanizmów ewolucji i przejawia się w złożeniu funkcji definiujących selekcję, mutację oraz krzyżowanie. Badając dynamikę operatora heurystycznego, można określić asymptotyczne zachowanie procesu – punkty stałe i ich stabilność. Populacje nieskończone reprezentowane są jako zbiór

⁶ Simpleks jest $(n-1)$ -wymiarowym hiperczworościanem, zawierającym się w n -wymiarowej przestrzeni otaczającej, czyli generalizacją czworościanu do n wymiarów.

⁷ Na przykład punktom z niewymiernymi współrzędnymi nie odpowiada żadna populacja.

gęsty w simpleksie, a stabilne punkty stałe dla modelu bez mutacji znajdują się w jego narożnikach. Punkty stałe nie muszą odpowiadać rozwiązaniom optymalnym, lecz reprezentują oczekiwany, w długim czasie, rozkład dużej populacji. Gdy założymy mutację o niezerowym prawdopodobieństwie, populacja w punkcie stałym nie jest jednorodna [146], zawiera pewną część optymalnych typów, ale także inne, położone w sąsiedztwie optimum.

Włączenie rekombinacji powoduje dalszą komplikację modelu. Operator $\mathcal{G}$, który zdefiniowano dla mutacji oraz krzyżowania jedno-, dwupunktowego i równomiernego, jest trudny do zanalizowania, a określenie jego punktów stałych i ich stabilności staje się niełatwym zadaniem. Dla krzyżowania jednopunktowego i mutacji wykazano, że gdy prawdopodobieństwo mutacji jest mniejsze niż 0,5, to każdy punkt stały operatora jest atraktorem [122]. W przypadku liniowych funkcji celu udowodniono, że pod działaniem selekcji i krzyżowania średnie przystosowanie populacji rośnie, iterowanie operatora $\mathcal{G}$ zawsze prowadzi do punktu stałego, a wszystkie punkty stałe odpowiadają populacjom składającym się z osobników jednego typu [205]. Wyniki te uogólniono później dla innych funkcji celu [208]. Większość wersji algorytmu genetycznego z samą rekombinacją prowadzi do wielokrotnych stanów absorbujących, którymi są homogeniczne populacje. Rekombinacja powoduje, że w stanie ustalonym zmniejsza się liczba optymalnych osobników, a zwiększa liczba prawie optymalnych.

Pokazano także bardziej skomplikowane zachowania układu dynamicznego generowanego przez algorytm genetyczny z kodowaniem binarnym. Dla selekcji proporcjonalnej, szczególnych postaci mutacji i krzyżowania pojawiają się orbity periodyczne o okresie 2 i większych, gdy prawdopodobieństwo mutacji rośnie [211]. Dla selekcji progowej i mutacji zależnej od gęstości (ang. *density-dependent mutation*)⁸ wykryto atraktory periodyczne oraz chaos [210]. Pokazano, że dla modelu z selekcją proporcjonalną, mutacją oraz krzyżowaniem równomiernym [212] lub wybierającym poszczególne allele potomka z całej puli genowej bieżącej populacji (ang. *gene pool crossover*) [213], w krajobrazie „igłowej” funkcji przystosowania (ang. *needle-in-the haystack*) istnieją dwa stabilne punkty stałe (ang. *bistability*), jeden w pobliżu optimum, a drugi w centrum simpleksu, któremu odpowiada populacja, w której wszystkie typy są jednakowo często reprezentowane. Populacje ewoluujące w przestrzeni funkcji mających dwa igłowe wierzchołki (ang. *bineedle-in-the haystack*) mogą mieć trzy punkty stałe, po jednym dla każdego maksimum i trzeci w centrum simpleksu [212]. Do analizy niektórych aspektów modelu stosowano także funkcje ortogonalne przekształcane transformacjami Fouriera i Walsha, co uelastyczyło i uogólniło teorię [201, 202, 203].

W przypadku skończonych populacji model Vose’a prowadzi do markowskiego modelu opisanego w poprzednim podrozdziale. Operator $\mathcal{G}$ wykorzystuje

⁸ Operator, w którym prawdopodobieństwo mutacji jest różne dla każdej współrzędnej wektora typów i zależne od częstości cech w populacji dla danej współrzędnej.

się do obliczania prawdopodobieństw przejścia łańcucha Markowa. Przedstawiono *twierdzenie asymptotyczne* [148, 203], w którym udowodniono, że dla bardzo dużych populacji i długiego czasu działania algorytmu trajektorie populacji są zbieżne do punktów stałych operatora $\mathcal{G}$. Twierdzenie to działa przy pewnych założeniach dotyczących operatora $\mathcal{G}$ [125]. Model nieskończonej populacji można wtedy traktować jako aproksymację modelu rzeczywistych skończonych populacji.

Przedstawiony model stosuje się nie tylko do algorytmów genetycznych z kodowaniem binarnym. Operator $\mathcal{G}$ jest równoważny z heurystyką zwaną przez Vose'a *Random Heuristic Search* RHS, która definiuje szeroką klasę algorytmów przeszukiwania heurystycznego, zawierającą także inne procedury optymalizacyjne, np. programowanie genetyczne, strategie ewolucyjne ($\mu + \lambda$), symulowane wyżarzanie czy algorytmy wspinaczkowe (ang. *hill-climbing*). Model ten zastosowano w algorytmach z kodowaniem rzeczywistoliczbowym [124, 182] oraz w analizie miar probabilistycznych algorytmu genetycznego [184, 185].

Model Vose'a dla nieskończonych populacji charakteryzuje się minimalną liczbą założeń. W większości przypadków pozwala na obliczenie wartości własnych jakobianu w punktach stałych, a więc na określenie stabilności tych punktów. Chociaż prezentowane podejście jest dokładne dla nieskończonych populacji, to jego wadą jest ograniczenie w długości analizowanych wektorów binarnych oraz komplikacja operatora $\mathcal{G}$ dla niektórych wersji algorytmu, która uniemożliwia analityczne obliczanie punktów stałych (obliczane są wtedy numerycznie). Większość wyników otrzymano dla ograniczonej klasy funkcji celu (tzw. unitarnych, ang. *unitation function*). Twierdzenie asymptotyczne, mówiące o zbieżności populacji do punktów stałych operatora $\mathcal{G}$ jest spełnione tylko dla algorytmu genetycznego z reprezentacją binarną o ustalonej długości wektora cech, z selekcją proporcjonalną, mutacją bitową o bardzo małym prawdopodobieństwie p_m i bez rekombinacji. Jednak nawet tak restrykcyjne założenia nie gwarantują, że model nieskończonej populacji będzie odpowiednim przybliżeniem do modelu populacji skończonych [125].

3.4. Inne modele

Teoria schematów. Pierwszym historycznie modelem teoretycznym opisującym działanie algorytmów ewolucyjnych była *teoria schematów* (ang. *schema theory*), którą Holland [91] przedstawił dla algorytmów genetycznych z kodowaniem binarnym. Główną ideą tej teorii jest dekompozycja skomplikowanych problemów na pewną hierarchię łatwiejszych zadań. Przypuszczano, że AG rozwiązują najpierw prostsze (pod)zadania i te częściowo dobre odpowiedzi przekształcają równolegle (głównie przez rekombinację) w rozwiązanie złożonego problemu. Na potrzeby modelu zdefiniowano *schematy* (ang. *schema*) – zestaw wzorców podobieństwa, które mają szczególny zbiór zdefiniowanych wartości, lub inaczej – mają te same allele

dla określonych loci (miejsc). Każdy schemat odpowiada pewnemu podzbiorowi przestrzeni poszukiwań.

Korzystając ze schematów, oszacowano oczekiwaną wartość polepszenia rozwiązania z generacji na generację uzyskiwaną przez algorytm z selekcją proporcjonalną, krzyżowaniem jednopunktowym i mutacją bitową. Obliczenia dokonano dla dużych populacji pokrywających przestrzeń poszukiwań tak, że znajduje się w niej wiele egzemplarzy różnych podzbiorów. W *twierdzeniu o schematach* udowodniono, że liczba schematów o ponadprzeciętnej jakości (wyższej niż średnie przystosowanie wszystkich wektorów binarnych w przestrzeni poszukiwań) wzrasta wykładniczo w kolejnych generacjach, a liczba schematów o niskiej jakości maleje wykładniczo [77, 91]. Z twierdzenia o schematach wyciągnięto wniosek, znany jako *hipoteza bloków tworzących* (ang. *building block hypothesis*) [77], mówiący o tym, że *bloki tworzące*, czyli krótkie schematy o wysokiej jakości (zwane także cegiełkami), które znajdują się w różnych rodzicach, mogą być rekombinowane i tworzyć schematy o jeszcze wyższym przystosowaniu. Hipoteza przedstawia krzyżowanie jako główne źródło efektywnego działania AG.

Oryginalna teoria schematów Hollanda została znacznie rozbudowana. Zaproponowano wiele uogólnień dotyczących bardziej ogólnych reprezentacji (tzw. *forma analysis*) [166, 167], dodano inne typy operatorów [53, 80], a także uszczegółowiono, podając dokładną wartość wartości oczekiwanej szacowanej w twierdzeniu o schematach [157, 203]. Teoria schematów była wielokrotnie krytykowana [1, 175]. Jej użyteczność jest ograniczona, głównie ze względu na stosunkowo restrykcyjne założenia oraz postać nierówności. Była jednak pierwszą próbą formalnego opisu algorytmów ewolucyjnych i mimo swoich wad wniosła ważny wkład w rozwój teorii. Nadal stosowana jest w badaniach metod ewolucyjnych, szczególnie w programowaniu genetycznym [158].

Modele, w których stosuje się fizykę statystyczną. Do analizy dynamiki algorytmów genetycznych z kodowaniem binarnym wykorzystuje się także metody stosowane w fizyce statystycznej [161, 162]. Ewoluuujące populacje, zwłaszcza bardzo liczne, mogą przypominać systemy cząstek w cieczech lub gazach, które są badane w fizyce. Aproksymacyjne metody mechaniki statystycznej pozwalają na badanie zbioru cząstek poprzez jego wartości średnie – kumulanty. Kumulanty są wielkościami związanymi z momentami rozkładów prawdopodobieństwa. Pierwsza kumulanta odpowiada wartości oczekiwanej, druga – wariancji, a kumulanty wyższego rzędu mierzą jak bardzo analizowany rozkład różni się od rozkładu normalnego. Trzecią i czwartą kumulantę można traktować jako odpowiedniki skośności i kurtozy rozkładu. Rozpatrując wartości średnie pewnych cech populacji, można analizować makroskopowe zachowanie systemu. Jeżeli przyjmie się założenie, że populacje są nieskończone, otrzymane wyniki będą statystycznie poprawne. Przenosząc rezultaty otrzymane dla nieskończonych populacji na skończone, należy uwzględnić fluktuacje wynikające z próbkowania danego rozkładu skończoną liczbę

razy (ang. *sampling effect*) i wprowadzić odpowiednie korekty [159]. Dla bardzo małych populacji korekty te mogą być znaczne.

W modelach mechaniki statystycznej analizuje się rozkłady funkcji celu (jako odpowiednika energii), a szczególnie wartości średnie ich kilku pierwszych kumulant. Kumulanty traktowane są jako makroskopowe zmienne poddane działaniu selekcji, mutacji i rekombinacji. Iterując deterministyczne równania opisujące kumulanty dla danych operatorów, można analizować ich trajektorie i badać dynamikę układu. Analiza modeli fizyki statystycznej pozwala na bardziej dokładną, niż modele układów dynamicznych, predykcję zachowań krótkoterminowych skończonych populacji. Niestety, stosuje się ją dla ograniczonej klasy funkcji celu typu one-max (D.7) i jej modyfikacji np. funkcji typu „basen z barierą”.

Badano różne rodzaje selekcji: proporcjonalną, turniejową, rankingową, Boltzmannna. Pokazano, że selekcja powoduje wzrost średniego przystosowania i zmniejszenie rozproszenia populacji. Wpływ mutacji i rekombinacji na rozkłady funkcji celu i ich kumulanty zależy od rozwiązywanego problemu i jego reprezentacji. Rekombinacja powoduje redukcję kumulant wyższego rzędu, co daje lepszą dokładność modelu. W zadaniach typu one-max (D.7) kombinacja selekcji i rekombinacji powoduje wzrost średniego przystosowania i lepszą zbieżność populacji do optimum [162]. Rekombinacja jest bardziej efektywna od mutacji w „skoku przez barierę”, przy optymalizacji funkcji typu „basen z barierą” [178]. Otrzymane estymaty kumulant pozwalają także na określenie optymalnych wartości parametrów procesu ewolucyjnego.

Modele bazujące na teorii genetyki populacyjnej. Inspiracje biologiczne w metodach ewolucyjnych nie ograniczają się jedynie do aplikacji mechanizmów naturalnych, ale także do wykorzystania metod analizy teoretycznej stosowanej w *genetyce populacyjnej* (ilościowej) (ang. *population/quantitative genetics*) [23]. Genetyka populacyjna koncentruje się na analizie częstości występowania poszczególnych alleli w populacjach naturalnych i bada różne czynniki wpływające na te częstości (np. selekcja, mutacja, dryf genetyczny).

Do badań procesów, dla których ewolucja średniego przystosowania następuje epokowo (ang. *epochal evolution*), wykorzystano konstrukcje matematyczne zaczerpnięte z genetyki populacyjnej, molekularnej teorii ewolucji oraz mechaniki statystycznej [145, 146, 147]. Makroskopowy stan populacji opisywano wektorem częstości występowania osobników o danym przystosowaniu, który w analogii do teorii Eigena [47], nazwano fenotypową quasi-populacją (ang. *quasi-species*). Analizowano dynamikę ewolucji i badano mechanizmy jej kontroli. Podobnie jak w modelu Vose’a, zdefiniowano operator $\mathcal{G}$ będący tutaj złożeniem operatorów selekcji i mutacji, który działając na bieżącym rozkładzie przystosowania daje oczekiwany rozkład jakości populacji w następnej generacji. Określono właściwości operatora wyjaśniające epokowość ewolucji średniego przystosowania. Pokazano, że rola rekombinacji w ewolucji epokowej jest marginalna. Oszacowano analitycznie liczbę

koniecznych obliczeń funkcji celu, aby osiągnąć optimum globalne w zależności od tempa mutacji, rozmiaru populacji oraz funkcji celu, co pozwoliło na rozwiązanie zadania optymalnego doboru parametrów dla danych klas zadań. Najbardziej efektywne okazały się wartości parametrów, dla których populacje w fazie latentnej są stabilne brzegowo.

4. Dynamika modelu nieskończonej populacji

Trudności, jakie napotykają badacze podejmujący próby analizy teoretycznej algorytmów ewolucyjnych zmuszają do przyjmowania założeń uproszczających, a w przypadkach modeli dokładnych (np. markowowskich) do analizy prostych przykładów o niewielkiej liczbie wymiarów. Jedną z najczęściej stosowanych aproksymacji jest założenie nieskończonej liczebności populacji. Warunek ten leży u podstaw zarówno modelu Vose'a [203, 204], jak i modeli wykorzystujących fizykę statystyczną [162], genetykę populacyjną [146] oraz innych [75, 200]. Założenie jest w pewnym sensie rozszerzeniem powszechnej praktyki stosowania w algorytmach genetycznych z kodowaniem binarnym licznych populacji, tak aby, z dużym prawdopodobieństwem, w populacji początkowej egzemplarze zer i jedynek były reprezentowane dla każdej cechy [76, 172]. Nieskończone populacje są także analizowane w biologii, np. w teorii quasi-species Eigena [47, 49], modelach cech ilościowych [128] czy rozwoju populacji wirusów [17].

Dzięki analizie nieskończonych populacji, losowe fluktuacje z generacji na generację, występujące w ewolucji populacji skończonych, mogą być aproksymowane nieliniowymi procesami deterministycznymi z ciągłą przestrzenią stanów. Warunek ten umożliwia opisanie w sposób makroskopowy zbiorowego zachowania populacji jako całości przez podanie kilku parametrów, takich jak: wartości oczekiwane, wariancje i wyższe kumulanty [162]. Oczekuje się, że trajektorie nieskończonych populacji są zbieżne do oczekiwanych trajektorii populacji mniejszych. W rzeczywistości ewolucja modelu aproksymacyjnego może być istotnie różna od poszczególnych realizacji procesów rzeczywistych [159, 199]. Jednakże modele nieskończonych populacji (czasem przy pewnych dodatkowych założeniach) pomagają w zrozumieniu dynamiki algorytmów ewolucyjnych z populacjami skończonymi [147].

Modele nieskończonych populacji dotyczące metod z kodowaniem binarnym rozpatrywane są w skończonych przestrzeniach poszukiwań. Nieskończone populacje w nieskończonych, rzeczywistych przestrzeniach poszukiwań, które są przedmiotem tego rozdziału, były analizowane przez Qi i Palmeriego [163]. Proces stochastyczny z czasem dyskretnym wykorzystano do modelowania zmian funkcji gęstości prawdopodobieństwa opisującego rozkład całej populacji poddanej działaniu operatorów genetycznych (zobacz rozdział 3.1). Analizę rozkładów próbkowania nieskończonych populacji działających w R^n dokonano także w rozprawie [4]. W niniejszym rozdziale przeanalizowano metodę poszukiwań ewolucyjnych z miękką selekcją dla nieskończonych populacji w rzeczywistych, nieograniczonych przestrzeniach poszukiwań [98, 100, 101, 102, 116]. Założenie o nieskończonej liczebności populacji

umożliwia rozpatrywanie rozkładów populacji w kolejnych generacjach w krajo-
brazach jedno- oraz wielomodalnych funkcji celu.

4.1. Model poszukiwań ewolucyjnych z miękką selekcją dla nieskończonych populacji

Metoda poszukiwań ewolucyjnych z miękką selekcją, przedstawiona w rozdziale 2.3.2, jest w ogólnym przypadku (m osobników i n -wymiarowa przestrzeń poszukiwań) trudna do analizy teoretycznej. Aproksymacja modelem nieskończonej populacji w jednowymiarowej przestrzeni poszukiwań ($n = 1$) przynosi korzyści w postaci opisu populacji poprzez funkcję rozkładu cech (proporcji osobników z cechą o danej wartości) w populacji. Zakładamy, że gdy liczebność populacji rośnie, różnice pomiędzy możliwymi rzeczywistymi przebiegami ewolucji a ich oczekiwaną aproksymacją dla nieskończonych populacji będą maleć¹.

W generacji i -tej rozkład cechy opisany jest przez funkcję gęstości prawdopodobieństwa $p_i(x)$ osobników z cechą x . Zgodnie z zasadami selekcji proporcjonalnej, FGP po selekcji $p'_i(x)$ jest wyrażona jako iloczyn rozkładu FGP populacji bieżącej oraz funkcji celu, który jest znormalizowany wartością oczekiwaną funkcji celu względem rozkładu prawdopodobieństwa $p_i(x)$ populacji w i -tej generacji

$$p'_i(x) = \frac{q(x)p_i(x)}{\int_{-\infty}^{\infty} q(z)p_i(z) dz} \quad (4.1)$$

oraz $\int_{-\infty}^{\infty} p'_i(z) dz = 1$. Wartości FGP po selekcji $p'_i(x)$ wzrastają dla cech, których przystosowania osobników są duże. Jednak także dla cech o niewielkiej wartości funkcji celu FGP przyjmuje wartości różne od zera. Łatwo zauważyć jedną z właściwości selekcji proporcjonalnej: lepsze (o większym przystosowaniu) osobniki są preferowane, ale również gorsze mają szansę na „wybranie”. Operacja mutacji jest opisana splotem funkcji gęstości po selekcji $p'_i(x)$ oraz funkcji rozkładu mutacji $g(x, y)$, która definiuje prawdopodobieństwo zmiany cech osobnika z y na x . Rozkład populacji w generacji $i + 1$ jest więc równy

$$p_{i+1}(x) = \int_{-\infty}^{\infty} p'_i(y)g(x, y) dy = \frac{\int_{-\infty}^{\infty} q(y)p_i(y)g(x, y) dy}{\int_{-\infty}^{\infty} q(z)p_i(z) dz}. \quad (4.2)$$

Otrzymano rozkład znany z deterministycznych modeli genetycznych cech ilościowych jako rozkład nieskończonej liczby alleli w nieskończonej populacji [119]. Zgod-

¹ Badania symulacyjne z rozdz. 4.3.3 potwierdzają prawdziwość tego założenia.

nie z założeniami modelu poszukiwań z miękką selekcją przyjmiemy, że $g(x, y)$ ma rozkład normalny wokół cechy y z wartością wariancji σ^2 jako parametrem:

$$g(x, y) = \mathcal{N}(y, \sigma^2)(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-y)^2}{2\sigma^2}\right). \quad (4.3)$$

Założenie, że mutacja powinna wprowadzać niewielkie zmiany do rozkładu po selekcji oraz prawdopodobieństwo zmiany tych cech wskutek mutacji powinno maleć ze wzrostem odległości między x i y , usprawiedliwia postać operatora mutacji. W literaturze spotyka się także zastosowanie rozkładów Cauchy'ego jako gęstości mutacji, które mają znacznie bardziej płaskie „ogony” niż rozkłady gaussowskie, co zwiększa prawdopodobieństwo mutacji o większym zasięgu (makromutacji) i wyjścia z pułapki ewolucyjnej [152, 216].

Funkcje celu². W niniejszej dysertacji, dynamikę lokalną nieskończonych populacji badano w krajobrazie jednomodalnej funkcji celu zadanej równaniem $q(x) = \exp(-ax^2)$ (patrz Dodatek, funkcja (D.2)), z maksimum w zerze i parametrem a opisującym nachylenie zbocza, a odpowiedzialnym za intensywność selekcji. Na potrzeby analizy wygodnie jest także opisać funkcję gaussowską parametrem szerokości wzgórza ζ^2 , tak aby odchylenie standardowe mutacji i rozległość wzgórze miały taką samą miarę odległości. Funkcja celu jest wtedy zapisana w następującej postaci: $q(x) = \exp[-x^2/(2\zeta^2)]$, gdzie $\zeta^2 = 1/(2a)$.

Dynamikę globalną analizowano dla rodziny dwumodalnych funkcji jakości będących sumą dwóch krzywych Gaussa, zadanych zależnością $q(x) = \exp(-ax^2) + h \exp[-a(x-d)^2]$ (patrz Dodatek, funkcja (D.4)), z maksimum lokalnym o wartości jeden w zerze oraz maksimum globalnym o wartości h , ($h \geq 1$) położonym w punkcie d . Wybrano krzywe gaussowskie, ponieważ mogą one aproksymować szerokie klasy funkcji, są opisane przez niewielką liczbę parametrów oraz tworzą zamkniętą klasę ze względu na operacje, którym poddawane są populacje (wyniki tych przekształceń także należą do klasy krzywych gaussowskich). Gaussowskie funkcje przystosowania stosuje się także w genetyce populacyjnej do analizy modeli ewolucji fenotypowej [120, 128]. Do celów porównawczych będą stosowane dodatkowo funkcje trójkątne.

4.2. Dynamika lokalna

Nawet dla jednomodalnej gaussowskiej funkcji przystosowania (D.2) nie jest możliwe podanie dokładnej formy $p_i(x)$ dla dowolnego rozkładu początkowego populacji, ponieważ równania (4.1) oraz (4.2) nie są rozwiązywalne analitycznie w przypadku ogólnym. Gdy jednak założymy, że początkowa populacja ma rozkład normalny $p_0(x) = \mathcal{N}(\eta_0, \nu_0^2)$, to możliwe jest wyznaczenie rozkładu w kolejnej

² Szczegółowe opisy funkcji celu wraz z ich reprezentacją graficzną zebrano w Dodatku.

generacji. Wybór rozkładu normalnego podyktowany jest jego pożądanymi właściwościami analitycznymi i statystycznymi³. Ponieważ funkcja przystosowania jest także gaussowska, więc całkowanie w nieskończonych granicach równania (4.2) daje w wyniku funkcję należącą do klasy funkcji gaussowskich. Ponadto, dowolną funkcję (ograniczoną i ciągłą w przedziale $-\infty < a \leq x \leq b < \infty$) można, z małym błędem, aproksymować w tym przedziale przez liniową kombinację funkcji gaussowskich. Przyjmując nawet, że rozkład populacji początkowej jest innego typu, np. jednostajny w zadanym przedziale, to, ze względu na normalną postać mutacji, w kolejnych generacjach zbliża się do normalnego⁴ (rys. 4.11) lub co najmniej jest nierównomierny i staje się normalny po pewnym czasie [4]. Rozkłady normalne stosowano również do analizy rozkładów przystosowania w modelach algorytmów genetycznych wykorzystujących fizykę statystyczną, co ułatwiło znalezienie wartości średnich i wariancji przystosowania w kolejnych generacjach [161]. Są też stosowane w biologii, w analizie rozkładów populacji i wpływów na nie różnych mechanizmów ewolucyjnych [119, 128].

Rozkład populacji w $(i + 1)$ -generacji, przy założeniu normalnego rozkładu populacji w i -tej generacji oraz jednomodalnej gaussowskiej funkcji przystosowania (D.2), obliczony na podstawie rozkładu (4.3), jest normalny $p_{i+1}(x) = \mathcal{N}(\eta_{i+1}, \nu_{i+1}^2)$ ze średnią η_{i+1} oraz wariancją ν_{i+1}^2 danymi wzorami

$$\eta_{i+1} = \frac{\eta_i}{1 + 2a\nu_i^2}, \quad (4.4)$$

$$\nu_{i+1}^2 = \sigma^2 + \frac{\nu_i^2}{1 + 2a\nu_i^2}. \quad (4.5)$$

Można zauważyć, że $\eta_{i+1} < \eta_i$, czyli rozkłady populacji przesuwają się w kierunku zera. Wariancja ν_{i+1}^2 jest większa od wariancji modyfikacji σ^2 . W granicy otrzymujemy rozkład stacjonarny (dynamiczny stan równowagi) $p_i(x) = p_{i+1}(x)$, którego parametry opisuje następujący lemat:

Lemat 4.2.1. *Nieskończona populacja w stanie równowagi selekcyjno-mutacyjnej ma rozkład normalny $\mathcal{N}(\eta_s, \nu_s^2)$ wokół optimum jednomodalnej funkcji przystosowania z wartością oczekiwaną w ekstremum $\eta_s = 0$ oraz wariancją opisaną wzorem*

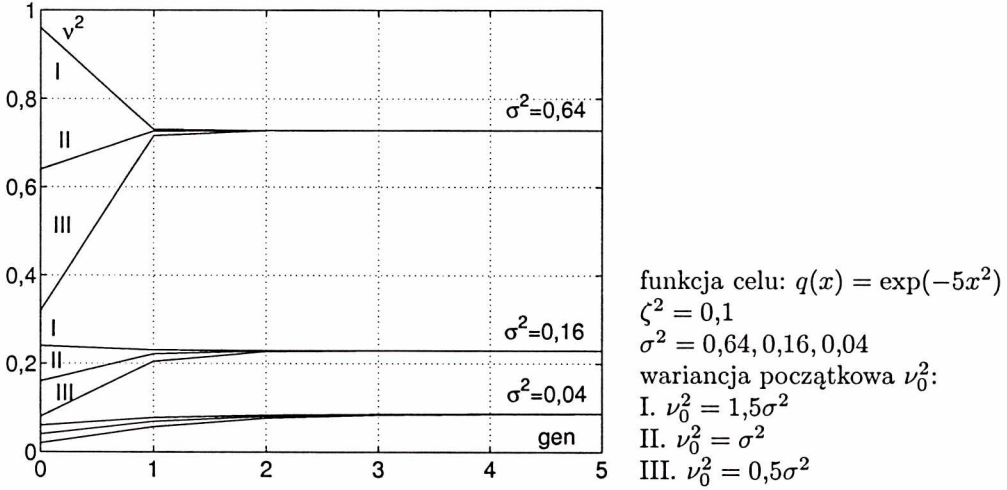
$$\nu_s^2 = \frac{\sigma^2}{2} \left(1 + \sqrt{1 + 2/(a\sigma^2)} \right) = \frac{\sigma^2}{2} \left(1 + \sqrt{1 + \zeta^2/\sigma^2} \right). \quad (4.6)$$

Dowód: Dla $i \rightarrow +\infty$ wartości oczekiwane η_i tworzą ściśle monotoniczny ciąg liczb rzeczywistych ze współczynnikiem zawężenia $\eta_{i+1}/\eta_i = 1/(1+2a\nu_i^2) < 1$. Tak więc,

³ W pracy [16] podano szczegółową charakterystykę właściwości rozkładów gaussowskich, a także ich mieszanin.

⁴ Wariancja tego rozkładu zależy od wariancji mutacji oraz nachylenia a funkcji celu.

$\eta_i \rightarrow 0$, gdy $i \rightarrow +\infty$. Wartość oczekiwana rozkładu stacjonarnego jest równa zero, czyli jest położona w punkcie, dla którego funkcja przystosowania osiąga optimum. Stacjonarną wartość wariancji ν_s^2 zadanej przez wzór (4.6) otrzymano z warunku $\nu_i^2 = \nu_{i+1}^2$. $\square$



Rys. 4.1. Wariancja rozkładu ν^2 dla różnych parametrów σ^2 i ν_0^2 w kolejnych generacjach

Fig. 4.1. Distribution variance ν^2 for various parameters σ^2 and ν_0^2 in consecutive generations

Lemat 4.2.1 wskazuje na zbieżność rozkładów nieskończonych populacji według średniej oraz z prawdopodobieństwem do optimum. Wariancja ν_{i+1}^2 opisana równaniem (4.5) jest zawsze większa niż wariancja rozkładu mutacji σ^2 i niezależnie od rozkładu początkowego bardzo szybko zbiega do wartości stacjonarnej ν_s^2 (rys. 4.1). Wartość wariancji rozkładu stacjonarnego zależy zarówno od wariancji mutacji σ^2 , jak i od szerokości wzgórza ζ^2 funkcji przystosowania. Rozproszenie populacji jest mniejsze dla małych σ^2 . Jeżeli wpływ ζ^2 dominuje nad σ^2 ($\zeta^2 \gg \sigma^2$), to wariancję ν_s^2 można oszacować jako $\nu_s^2 \simeq \sigma^2(1 + \zeta)/2$. Gdy wariancja mutacji jest znacznie większa od szerokości wzgórza $\sigma^2 \gg \zeta^2$, wartość ν_s^2 można oszacować jako $\nu_s^2 \simeq \sigma^2$. Wynik ten jest zbieżny z wnioskiem Qi i Palmeriego [163] dotyczącym niekorzystnego wpływu zaburzeń wywołanych mutacją na jakość populacji. Podobne rezultaty przyniosła analiza pierwszych dwóch momentów dla jednomodalnej gaussowskiej funkcji przystosowania w modelu wykorzystującym fizykę statystyczną [161] oraz badania dynamiki rozkładu próbkowania nieskończonych populacji [4]. Wariancja szybko osiąga wartość stacjonarną, a zbieżność wartości oczekiwanej jest wolniejsza.

4.3. Dynamika globalna

4.3.1. Rozkłady populacji

Dla efektywności poszukiwań globalnych ważne są eksploracyjne właściwości metody, pozwalające na opuszczenie obszaru optimum lokalnego i penetrację innych obszarów krajobrazu w poszukiwaniu ekstremum globalnego. Zbadamy więc możliwości eksploracyjne analizowanego modelu, w szczególności przejście siodła, w krajobrazie dwumodalnej funkcji przystosowania (D.4), gdy populacja początkowa skoncentrowana jest na niższym jej wzgórzu. Stanem wyjściowym będzie populacja w stanie równowagi selekcyjno-mutacyjnej, rozłożona normalnie wokół optimum lokalnego, z wariancją ν_0^2 (rys. 4.2a.). Funkcja gęstości prawdopodobieństwa rozkładu populacji jest wtedy zadana zależnością

$$p_0(x) = \frac{1}{\sqrt{2\pi\nu_0^2}} \exp\left(-\frac{x^2}{2\nu_0^2}\right). \quad (4.7)$$

Rozkład populacji pierwszej obliczony na podstawie równań (4.1), (4.2), z wykorzystaniem danych (4.3), (4.7), jest wyrażony wzorem

$$p_1(x) = \frac{1}{1 + h\kappa_{1,1}} \mathcal{N}(0, \nu_1^2) + \frac{h\kappa_{1,1}}{1 + h\kappa_{1,1}} \mathcal{N}(\gamma_{1,1}, \nu_1^2), \quad (4.8)$$

gdzie:

$$\nu_1^2 = \sigma^2 + \frac{\nu_0^2}{1 + 2a\nu_0^2}, \quad \kappa_{1,1} = \exp\left(-\frac{ad^2}{1 + 2a\nu_0^2}\right), \quad \gamma_{1,1} = \frac{2a\nu_0^2}{1 + 2a\nu_0^2}d.$$

Wartość oczekiwana rozkładu o FGP $p_1(x)$ wynosi

$$\mu_1 = \frac{h\kappa_{1,1}}{1 + h\kappa_{1,1}} \gamma_{1,1}. \quad (4.9)$$

Otrzymany rozkład (4.8) jest sumą dwóch rozkładów normalnych (rys. 4.2b): jednego o wartości oczekiwanej równej zeru i wariancji ν_1^2 (skupionego na szczycie niższego wzgórza) i drugiego o wartości oczekiwanej równej $\gamma_{1,1}$ i wariancji ν_1^2 (przesuniętego w kierunku wyższego wzgórza). Ponieważ $\mu_1 > \mu_0$, populacja nie pozostaje w optimum lokalnym, lecz przemieszcza się w kierunku optimum globalnego. Wielkość przesunięcia wartości oczekiwanej (4.9) zależy od odległości między wzgórzami d , wariancji rozkładu początkowego ν_0^2 i parametrów funkcji celu h, a .

Rozkład populacji po generacji drugiej jest sumą czterech rozkładów normalnych

$$p_2(x) = \frac{\mathcal{N}(0, \nu_2^2) + h\kappa_{2,1}\mathcal{N}(\gamma_{2,1}, \nu_2^2) + h\kappa_{2,2}\mathcal{N}(\gamma_{2,2}, \nu_2^2) + h^2\kappa_{2,3}\mathcal{N}(\gamma_{2,1} + \gamma_{2,2}, \nu_2^2)}{1 + h\kappa_{2,1} + h\kappa_{2,2} + h^2\kappa_{2,3}} \quad (4.10)$$

z parametrami

$$\begin{aligned}\nu_2^2 &= \sigma^2 + \frac{\nu_1^2}{1 + 2a\nu_1^2}, & \gamma_{2,2} &= \frac{\gamma_{1,1}}{1 + 2a\nu_1^2}, & \gamma_{2,1} &= \frac{2a\nu_1^2}{1 + 2a\nu_1^2}d, \\ \kappa_{2,1} &= \exp\left(-\frac{ad^2}{1 + 2a\nu_1^2}\right), & \kappa_{2,2} &= \kappa_{1,1} \exp\left(-\frac{a\gamma_{1,1}^2}{1 + 2a\nu_1^2}\right), \\ \kappa_{2,3} &= \kappa_{1,1} \exp\left(-\frac{a(\gamma_{1,1} - d)^2}{1 + 2a\nu_1^2}\right).\end{aligned}$$

Ponownie jeden z rozkładów składowych jest skoncentrowany wokół optimum lokalnego, a pozostałe są przesunięte w kierunku optimum globalnego (rys. 4.2c). Kontynuując, można obliczyć rozkłady populacji w kolejnych generacjach, aczkolwiek obliczenia stają się coraz bardziej skomplikowane. W generacji trzeciej rozkład jest sumą ośmiu (rys. 4.2d)^{5,6}, a w i -tej – 2^i rozkładów normalnych.

Modele, w których występują mieszaniny rozkładów o danej funkcji gęstości prawdopodobieństwa (szczególnie rozkłady normalne, ang. *Gaussian Mixture Models* GMMs [41]) są znane i często stosowane w statystyce do klasyfikacji danych oraz w określaniu wielowymiarowych rozkładów zmiennych losowych. Mieszaniny gaussowskie są popularne, ponieważ umożliwiają wierną reprezentację złożonych FGP i skomplikowanych wzajemnych powiązań w zbiorach danych⁷. Zwykle mieszaniny rozkładów otrzymuje się wtedy, gdy próbki pochodzą z różnorodnych populacji lub wielowymiarowych niejednorodnych danych, gdzie dane z każdej klasy są modelowane przez pojedynczy rozkład [136] (np. analiza częstości występowania ryb o danej długości w rybołówstwie [131]), podczas gdy w badanym w pracy przypadku populacja jest jednorodna⁸. Również w dziedzinie algorytmów ewolucyjnych wykorzystywane są mieszaniny rozkładów normalnych jako estymatory kolejnych rozkładów prawdopodobieństwa dla metod, które nie stosują w badaniach populacji osobników, lecz populacje rozkładów [73]. Oprócz mieszanin gaussowskich zaproponowano inne estymatory, przykładowo rozkłady Erlanga [126]. Ponieważ rozkład mieszaniny rozkładów gaussowskich nie jest normalny, jego analiza w ogólnym przypadku jest utrudniona.

4.3.2. Parametry rozkładu mieszaniny

Analizując kolejne generacje osobników w krajobrazie dwumodalnej funkcji przystosowania, można określić parametry rozkładów składowych i -tej generacji,

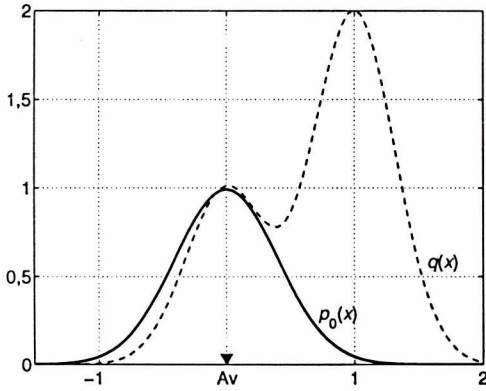
⁵ Cztery z ośmiu rozkładów nie są widoczne na rysunku z powodu bardzo małej wagi.

⁶ Trójkąt Av na rys. 4.2, 4.4–4.8, 4.10, 4.11 wskazuje wartość oczekiwaną mieszaniny.

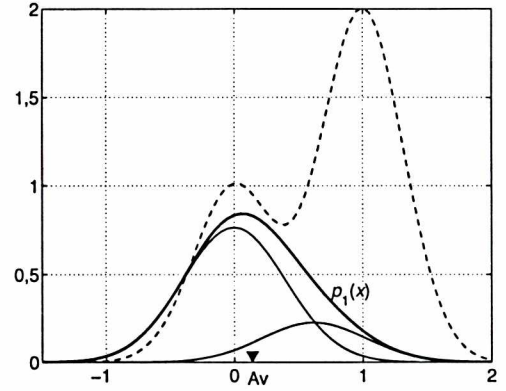
⁷ Opis wykorzystania mieszanin rozkładów gaussowskich w różnorodnych dziedzinach, od psychologii przez geologię po astrofizykę, można znaleźć w [197].

⁸ Wszystkie cechy należą do jednej klasy.

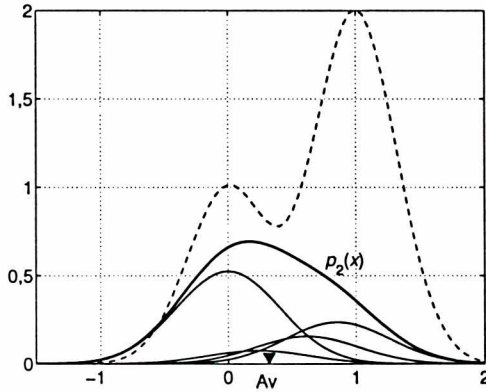
a) generacja 0



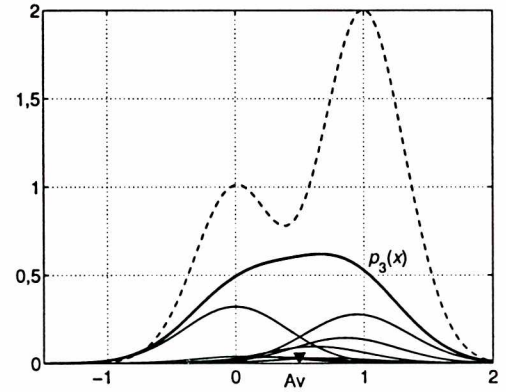
b) generacja 1



c) generacja 2



d) generacja 3



Rys. 4.2. Rozkłady nieskończonych populacji w krajobrazie dwumodalnej funkcji (D.4), $a = 5$, $h = 2$, $d = 1$, $\sigma^2 = 0,1$. Rozkład początkowy $p_0(x) = \mathcal{N}(0, \nu_0^2)$, $\nu_0^2 = 0,162$

Fig. 4.2. Distributions of trait in the landscape of bimodal fitness functions (D.4), $a = 5$, $h = 2$, $d = 1$, $\sigma^2 = 0.1$. Initial distribution: $p_0(x) = \mathcal{N}(0, \nu_0^2)$, $\nu_0^2 = 0.162$

dla populacji początkowej położonej w optimum lokalnym. Poniżej przedstawiono sposób obliczenia wartości oczekiwanych rozkładów składowych $\eta_{i,j}$, oznaczając przez l_k ($k = 0, 1, \dots, i$; $\sum_{k=0}^i l_k = 2^i$) liczbę rozkładów składowych, dla których wartości oczekiwane są sumą k współczynników $\gamma_{i,j}$. Rozkład mieszaniny w i -tej generacji jest więc sumą:

- $l_0 = \binom{i}{0} = 1$ rozkładu o wartości oczekiwanej równej zero $\eta_{i,0} = 0$, który pozostaje w optimum lokalnym;

- $l_1 = \binom{i}{1} = i$ rozkładów, dla których wartości oczekiwane są równe współczynnikom $\gamma_{i,j}$: $\eta_{i,1} = \gamma_{i,1}, \dots, \eta_{i,i} = \gamma_{i,i}$

– $l_2 = \binom{i}{2}$ rozkładów, dla których wartości oczekiwane są równe sumie dwóch współczynników $\gamma_{i,j}$: $\eta_{i,r} = \gamma_{i,p} + \gamma_{i,s}$, gdzie $r = i+1, \dots, i+l_2$, a p i s ($p, s = 1, \dots, i$) przebiegają zbiór wszystkich par (p, s) ,

– $l_3 = \binom{i}{3}$ rozkładów, dla których wartości oczekiwane są równe sumie trzech współczynników $\gamma_{i,j}$: $\eta_{i,r} = \gamma_{i,p} + \gamma_{i,s} + \gamma_{i,t}$, gdzie $r = i+l_2+1, \dots, i+l_2+l_3$, a p, s, t ($p, s, t = 1, \dots, i$) przebiegają zbiór wszystkich trójek (p, s, t) ,

– ...

– $l_i = \binom{i}{i} = 1$ rozkładu, dla którego wartość oczekiwana jest równa sumie wszystkich współczynników $\gamma_{i,j}$: $\eta_{i,(2^i-1)} = \gamma_{i,1} + \gamma_{i,2} + \dots + \gamma_{i,i}$.

Ponieważ powyższy sposób określania wartości oczekiwanych rozkładów składowych jest dość skomplikowany, zaproponowano prostszą metodę wyznaczania parametrów mieszaniny. Rozkład populacji w i -tej generacji jest sumą 2^i składowych rozkładów normalnych opisanych ogólnie jako

$$p_i(x) = \sum_{j=0}^{2^i-1} \beta_i^{(j)} \mathcal{N}(\eta_i^{(j)}, \nu_i^2), \quad (4.11)$$

gdzie współczynniki $\beta_i^{(j)}$, $j = 0, \dots, 2^i - 1$ opisują wagi poszczególnych rozkładów. Rozkład (4.11) można uogólnić dla funkcji wielomodalnej o k maksimach (D.1), sumując k^i rozkładów normalnych w i -tej iteracji.

Z analizy rozkładów początkowych przedstawionych na rys. 4.2 wynika, że każdy składowy rozkład normalny generuje w kolejnej generacji dwa rozkłady „potomne”: jeden ulokowany w pobliżu optimum lokalnego, a drugi w okolicach optimum globalnego. Potomka położonego w pobliżu optimum lokalnego oznaczać będziemy indeksem 0, a potomka lokowanego w pobliżu ekstremum globalnego opatrzymy indeksem 1. Zmienną r indeksującą kolejność rozkładów w mieszaninie wygodnie jest zakodować binarnie. W takiej interpretacji binarny indeks kolejnego rozkładu (czytany od strony lewej ku prawej) wskazuje na rodowód danego rozkładu. Tak więc globalny rozkład składowy w generacji czwartej oznaczony (1011) jest potomkiem dwóch rozkładów globalnych („ojciec” i „pradziad”) oraz jednego lokalnego („dziad”). Rozkłady, których łańcuch rodowodu kończy 0 nazwane są *frakcją lokalną*, a te zakończone 1 – *frakcją globalną*. Wykorzystywany dalej zapis indeksu ($j0/1$) oznacza konkatencję binarnej reprezentacji j oraz 0/1 (na przykład dla $j = 5$: ($j0$) = (1010), a ($j1$) = (1011)). Zapis łańcucha rodowodu w postaci binarnej ułatwia określenie parametrów kolejnych rozkładów w postaci rekurencyjnej. Lemat 4.3.1 podaje wartości oczekiwane, wagi oraz wariancję rozkładu mieszaniny w i -tej generacji, gdy populacja początkowa położona jest w optimum lokalnym.

Lemat 4.3.1. *Dla dwumodalnej funkcji przystosowania (D.4) j -ta składowa rozkładu populacji w generacji i -tej danego wzorem (4.11), o wartości oczekiwanej*

$\eta_i^{(j)}$ i wadze $\beta_i^{(j)}$, generuje w kolejnej generacji dwa nowe rozkłady o wartościach oczekiwanych:

$$\eta_{i+1}^{(j0)} = \frac{\eta_i^{(j)}}{z_i}, \quad (4.12)$$

$$\eta_{i+1}^{(j1)} = \frac{\eta_i^{(j)}}{z_i} + d \left(1 - \frac{1}{z_i} \right), \quad (4.13)$$

wagach

$$\beta_{i+1}^{(j0)} = \beta_i^{(j)} \frac{\exp(-a \eta_i^{(j)2} / z_i)}{\sum_{r=0}^{2^i-1} \beta_i^{(r)} \left(\exp(-a \eta_i^{(r)2} / z_i) + h \exp(-a (d - \eta_i^{(r)})^2 / z_i) \right)}, \quad (4.14)$$

$$\beta_{i+1}^{(j1)} = \beta_i^{(j)} \frac{h \exp(-a (d - \eta_i^{(j)})^2 / z_i)}{\sum_{r=0}^{2^i-1} \beta_i^{(r)} \left(\exp(-a \eta_i^{(r)2} / z_i) + h \exp(-a (d - \eta_i^{(r)})^2 / z_i) \right)}, \quad (4.15)$$

i takiej samej wariancji

$$\nu_{i+1}^2 = \sigma^2 + \frac{\nu_i^2}{z_i}, \quad (4.16)$$

gdzie $z_i = 1 + 2a\nu_i^2$. Rozkład populacji w kolejnej generacji przedstawia sumę frakcji lokalnej i globalnej i wynosi

$$\begin{aligned} p_{i+1}(x) &= \sum_{j=0}^{2^i-1} \left(\beta_{i+1}^{(j0)} \mathcal{N}(\eta_{i+1}^{(j0)}, \nu_{i+1}^2) + \beta_{i+1}^{(j1)} \mathcal{N}(\eta_{i+1}^{(j1)}, \nu_{i+1}^2) \right) \\ &= \sum_{j=0}^{2^{i+1}-1} \beta_{i+1}^{(j)} \mathcal{N}(\eta_{i+1}^{(j)}, \nu_{i+1}^2). \end{aligned} \quad (4.17)$$

Dowód: Obliczenia kolejnych rozkładów mieszaniny (4.11), przeprowadzone według równań (4.1), (4.2), (4.3), dają w rezultacie wzory (4.12)–(4.16). $\square$

Zauważmy, że wariancja (4.16) jest równa wariancji rozkładu dla jednomodalnej funkcji przystosowania (4.5). Ponieważ wariancja szybko zbiega do swojej wartości stacjonarnej (4.6), można przyjąć wartość wariancji początkowej równą wariancji stacjonarnej $\nu_0 = \nu_s$. Ułatwi to analizę rozkładów, gdyż wariancja nie zmienia się w kolejnych generacjach, czyli wariancje rozkładów składowych (4.16) są niezależne od numeru generacji, podobnie jak parametr $z_i = z = 1 + 2a\nu_s^2$. Rozkład nieskończonej populacji w kolejnych generacjach można teraz wyrazić w bardziej zwartej formie.

Wniosek 4.3.1. Dla normalnego rozkładu nieskończonej populacji początkowej z wartością oczekiwaną w optimum lokalnym $\mathcal{N}(0, \nu_0^2 = \nu_s^2)$, rozkład populacji $p_{i+1}(x)$ w generacji $(i + 1)$ jest mieszaniną rozkładów normalnych daną wzorem

$$p_{i+1}(x) = \sum_{j=0}^{2^i-1} \left(\beta_{i+1}^{(j0)} \mathcal{N}(\eta_{i+1}^{(j0)}, \nu_s^2) + \beta_{i+1}^{(j1)} \mathcal{N}(\eta_{i+1}^{(j1)}, \nu_s^2) \right), \quad (4.18)$$

gdzie

$$\eta_{i+1}^{(j0)} = \frac{\eta_i^{(j)}}{z}, \quad (4.19)$$

$$\eta_{i+1}^{(j1)} = \frac{\eta_i^{(j)}}{z} + d \left(1 - \frac{1}{z} \right), \quad (4.20)$$

$$\beta_{i+1}^{(j0)} = \beta_i^{(j)} \frac{\exp(-a \eta_i^{(j)2} / z)}{\sum_{r=0}^{2^i-1} \beta_i^{(r)} \left(\exp(-a \eta_i^{(r)2} / z) + h \exp(-a (d - \eta_i^{(r)})^2 / z) \right)}, \quad (4.21)$$

$$\beta_{i+1}^{(j1)} = \beta_i^{(j)} \frac{h \exp(-a (d - \eta_i^{(j)})^2 / z)}{\sum_{r=0}^{2^i-1} \beta_i^{(r)} \left(\exp(-a \eta_i^{(r)2} / z) + h \exp(-a (d - \eta_i^{(r)})^2 / z) \right)}, \quad (4.22)$$

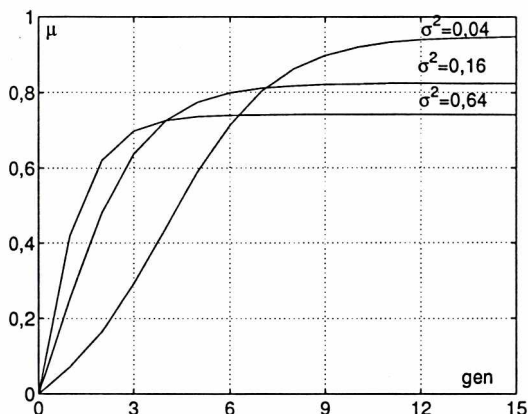
Wartość oczekiwana μ_{i+1} oraz wariancja ν_{i+1}^2 rozkładu $p_{i+1}(x)$ wynoszą

$$\mu_{i+1} = \sum_{j=0}^{2^{i+1}-1} \beta_{i+1}^{(j)} \eta_{i+1}^{(j)} = \sum_{j=0}^{2^i-1} \left(\beta_{i+1}^{(j0)} \eta_{i+1}^{(j0)} + \beta_{i+1}^{(j1)} \eta_{i+1}^{(j1)} \right), \quad (4.23)$$

$$\nu_{i+1}^2 = \nu_{i+1}^2 \sum_{j=0}^{2^{i+1}-1} \beta_{i+1}^{(j)2} = \nu_{i+1}^2 \sum_{j=0}^{2^i-1} \left(\beta_{i+1}^{(j0)2} + \beta_{i+1}^{(j1)2} \right). \quad (4.24)$$

Wariancja mieszaniny (4.24) jest mniejsza niż wariancja (4.16) rozkładów składowych, albowiem $\sum_{j=0}^{2^{i+1}-1} (\beta_{i+1}^{(j)})^2 < 1$. Ponieważ $0 \leq \eta_{i+1}^{(*0)} < \eta_i^{(*)} < \eta_{i+1}^{(*1)}$, jeden z rozkładów potomnych jest ulokowany bliżej optimum lokalnego, a drugi bliżej optimum globalnego niż rozkład rodzicielski. Odległość pomiędzy rozkładami potomnymi jest stała i wynosi $d - d/z$, jest więc mniejsza niż odległość pomiędzy wzgórzami funkcji przystosowania. Obie frakcje zawierają jednakową liczbę 2^{i-1} rozkładów, lecz „masa” frakcji lokalnej jest mniejsza⁹ i maleje w kolejnych generacjach. Składowy rozkład globalny opisany łańcuchem złożonym z samych jedynek szybko zaczyna mieć dominujące znaczenie, gdyż jego waga jest proporcjonalna do

⁹ Wagi frakcji globalnej są mnożone przez h .



Rys. 4.3. Wartości oczekiwane rozkładu populacji (4.23) dla różnych wartości σ^2 w kolejnych generacjach
 Fig. 4.3. Distribution expected values (4.23) parametrized with the mutation variance σ^2 in consecutive generations

i -tej potęgi wysokości wzgórza globalnego h^i . Wartość oczekiwaną tego rozkładu można obliczyć ze wzoru

$$\eta_i^{11\dots 11} = d + \frac{\eta_0 - d}{\prod_{j=0}^{i-1} z_j} = d + \frac{\eta_0 - d}{z^i},$$

a ponieważ $z > 1$, więc dąży ona do optimum globalnego $\lim_{i \rightarrow \infty} \eta_i^{11\dots 11} = d$.

Skomplikowana forma rozkładu populacji (4.18) uniemożliwia otrzymanie dokładnych rezultatów z analizy stanów ustalonych. Wartości oczekiwane rozkładu mogą być obliczone tylko numerycznie, a wyniki ich ewolucji przedstawiono na rys. 4.3. Wartości te stabilizują się już po kilku generacjach, przy czym im większa wariancja mutacji σ^2 , tym zbieżność jest szybsza. Wartości oczekiwane rozkładów stacjonarnych nie leżą dokładnie w optimum globalnym, lecz są przesunięte w stronę niższego wzgórza. Niedokładność lokalizacji zależy od wariancji mutacji i jest tym większa, im większe jest σ^2 .

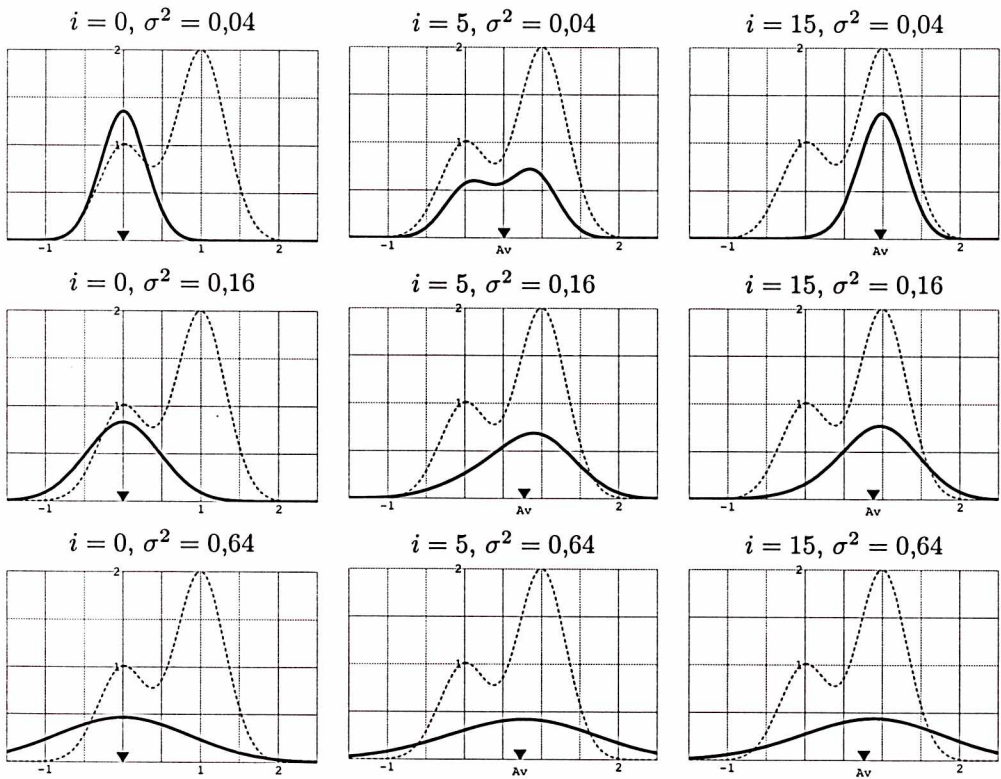
4.3.3. Badania symulacyjne rozkładów

Rozkłady nieskończonej populacji opisane równaniami (4.18)–(4.22) symulowano w krajobrazie dwumodalnej funkcji przystosowania dla różnych wartości parametrów procesu oraz krajobrazu:

- wariancji mutacji $\sigma^2 = 0,04; 0,16; 0,64$;
- nachylenia zboczy wzgórz funkcji przystosowania $a = 5, 10, 15$;
- wysokości optimum globalnego $h = 1, 2, 3$.

Praktycznie we wszystkich symulacjach początkowa populacja ma rozkład normalny $\mathcal{N}(0, \nu_s^2)$. Liczbę możliwych do wykonania generacji ograniczono do 15 ze względu na ograniczenia zasobowe komputera (w i -tej generacji obliczano 2^i rozkładów wraz z ich parametrami)¹⁰. Rozkład z generacji 15. może być traktowany jako rozkład w stanie równowagi, gdyż dla podanych wartości σ^2 wartości oczekiwane rozkładu mieszaniny są już stacjonarne (zob. rys. 4.3).

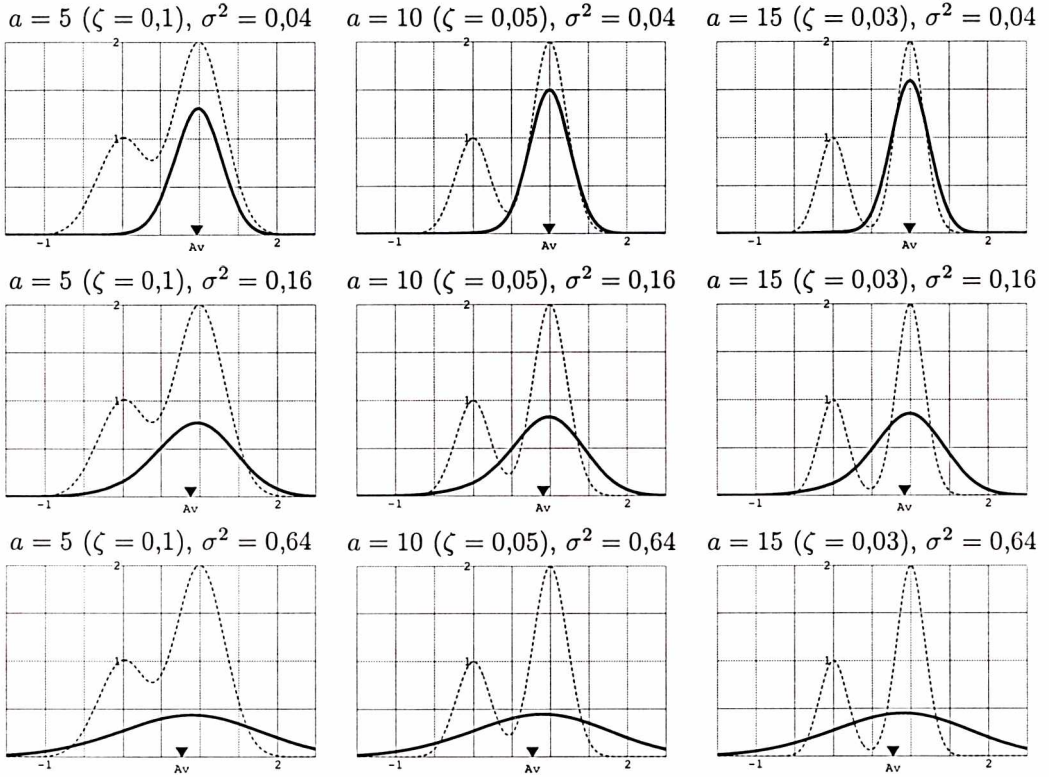
Na rysunku 4.4 przedstawiono rozkład początkowy oraz rozkłady po 5. i 15. generacji dla dwumodalnej funkcji przystosowania (D.4) z parametrami $h = 2$, $a = 5$ oraz różnych wartości wariancji mutacji. Populacja początkowo rozlokowana wokół optimum lokalnego szybko przesuwa się w kierunku wyższego wzniesienia i stabilizuje



Rys. 4.4. Rozkłady nieskończonej populacji (4.18) w generacjach 0, 5, 15 dla różnych wartości wariancji mutacji σ^2 , $a = 5$, $h = 2$

Fig. 4.4. Distributions of trait (4.18) in generations 0, 5, 15 for diverse values of the mutation variance σ^2 , $a = 5$, $h = 2$

¹⁰ Sposób aproksymacji rozkładów wyższych generacji przedstawiono w dalszej części rozdziału.

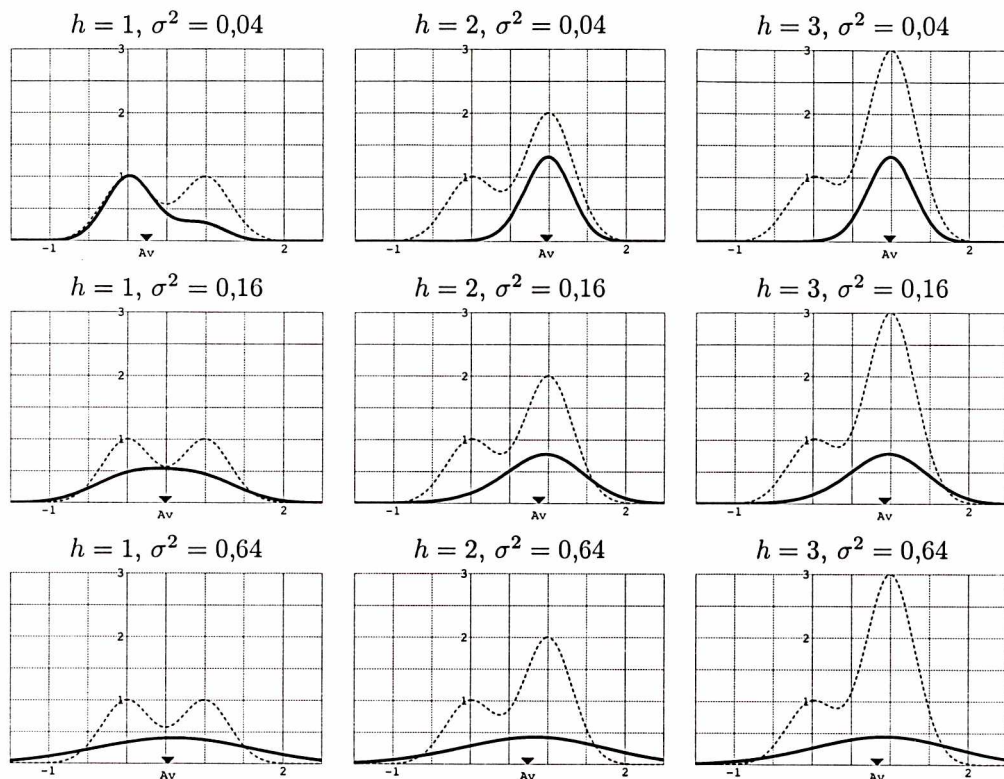


Rys. 4.5. Rozkłady nieskończonej populacji (4.18) w 15. generacji jako funkcja nachylenia zbocza wzgórz a , $h = 2$, $i = 15$

Fig. 4.5. Distributions of trait (4.18) in the 15th generation as a function of a hillside decline a , $h = 2$, $i = 15$

się w jego sąsiedztwie. Dla dużych wartości σ^2 stabilizacja jest szybsza. Rozkłady w generacji numer 15, choć podobne kształtem, nie są jednak rozkładami normalnymi, z powodu zaburzeń spowodowanych frakcją lokalną. Dla małych wariancji mutacji wartość oczekiwana rozkładu mieszaniny jest położona bardzo blisko ekstremum globalnego, a składowe lokalne mają niewielkie wagi. Dla większych σ^2 populacja jest szeroko rozłożona w krajobrazie funkcji przystosowania, a wartości oczekiwane rozkładów lokalizują wzgórze globalne, aczkolwiek dokładność lokalizacji optimum pogarsza się wraz ze wzrostem wartości σ^2 .

Wpływ nachylenia zboczy wzgórz funkcji przystosowania na rozkład mieszaniny w 15. generacji przedstawiono na rys. 4.5. Zmiana parametru a powoduje zmiany szerokości wzgórz (opisywanego parametrem ζ^2) oraz głębokości siodła pomiędzy optimami. Różnice w szerokości wzgórz mają wpływ na rozkład tylko dla bardzo małych σ^2 . Wtedy parametry σ^2 i ζ^2 mają podobne wielkości i nie można pomi-

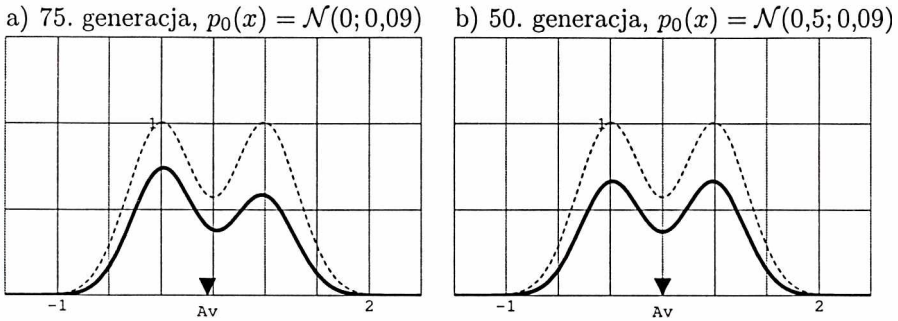


Rys. 4.6. Zależność rozkładów nieskończonej populacji (4.18) w 15. generacji od wysokości optimum globalnego h , $a = 5$, $i = 15$

Fig. 4.6. Distributions of trait (4.18) in the 15th generation as a function of the global optimum height h , $a = 5$, $i = 15$

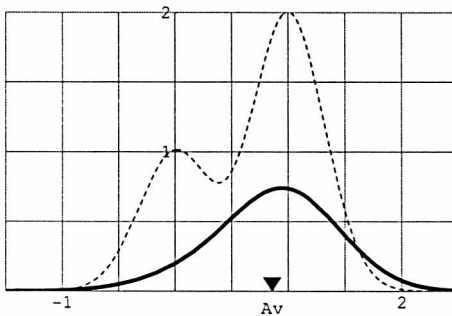
nąć wpływu żadnego z nich. Bardziej intensywna selekcja powoduje wzrost wagi frakcji lokalnej i przesunięcie rozkładu w stronę optimum lokalnego. Gdy $\sigma^2 \gg \zeta^2$, szerokość wzgórza nie ma znaczącego wpływu na rozkład populacji.

Bardziej zróżnicowane rozkłady obserwuje się wtedy, gdy analizujemy wpływ wysokości wzgórz (rys. 4.6). Wraz ze wzrostem wysokości optimum globalnego, wagi frakcji globalnej (mnożone przez h) rosną i dominują nad frakcją lokalną. Im wyższe wzgórze globalne, tym szybciej wartość oczekiwana rozkładu mieszaniny jest zbieżna do ekstremum, dokładniej go lokalizuje, a rozkład bardziej przypomina normalny. Gdy optima mają jednakową wysokość $h = 1$, wartość oczekiwana rozkładu leży na siodle dla większych wartości σ^2 . Punkt siodłowy jest punktem stacjonarnym rozkładu. Gdy wariancja mutacji jest bardzo mała, zbieżność rozkładów jest powolna i w generacji 15. (przedstawionej na rysunku) wartość oczekiwana mieszaniny jest niewiele przesunięta z optimum lokalnego. Można oczekiwać, że



Rys. 4.7. Dwumodalny rozkład nieskończonej populacji (4.18), dla jednakowych wysokości optimum ($h = 1$, $a = 5$, $\sigma^2 = 0,04$) i różnych rozkładów początkowych

Fig. 4.7. Bimodal distribution of trait (4.18) for equal optima height ($h = 1$, $a = 5$, $\sigma^2 = 0,04$) and various initial distributions



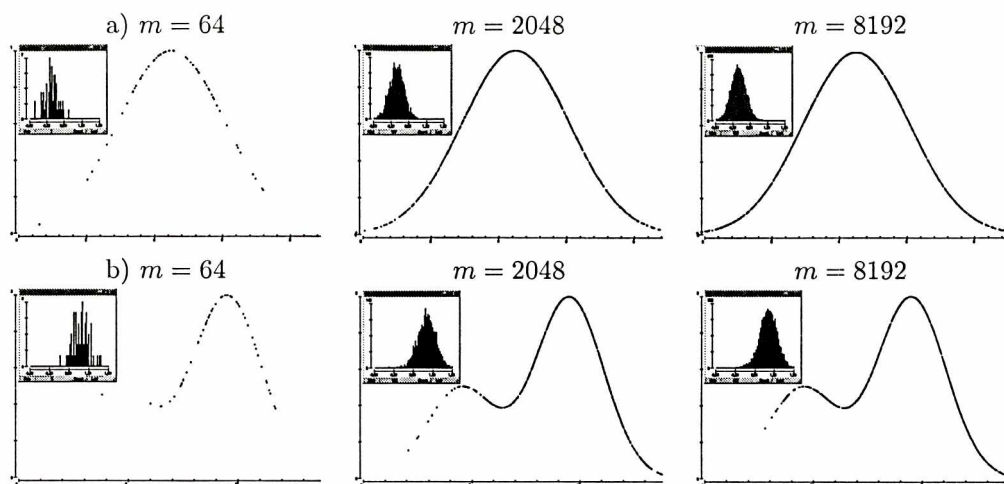
populacja początkowa rozłożona wokół optimum globalnego:
 $p_0(x) = \mathcal{N}(1; 0,23)$;
 $a = 5$, $h = 2$, $\sigma^2 = 0,16$

Rys. 4.8. Rozkład nieskończonej populacji (4.18) w 15. generacji
 Fig. 4.8. Distribution of trait (4.18) in the 15th generation

podobnie jak w przypadku większych σ^2 , oczekiwana wartość stacjonarna będzie leżała na siodle. Hipotezę tę sprawdzono, pomijając w mieszaniu rozkłady składowe o bardzo małej wadze, mniejszej od założonego parametru (rys. 4.7). Symulacje potwierdziły przypuszczenia: oczekiwana wartość stacjonarna leży na siodle $\mu_s = 0,5$, a otrzymany rozkład stacjonarny jest dwumodalny (rys. 4.7a). Rozkład dwumodalny otrzymamy także wtedy, gdy populację początkową rozłożymy wokół punktu siodłowego (rys. 4.7b), na co wskazują obliczenia parametrów zgodnie z równaniami (4.12)–(4.15).

Dla $\sigma^2 = 0,16$ sprawdzono, jak zachowa się populacja rozłożona początkowo w optimum globalnym $p_0(x) = \mathcal{N}(d, \nu_s^2)$ (rys. 4.8). Widoczny jest wpływ wzgórza lokalnego, który deformuje rozkład, przyciągając jego wartość oczekiwaną.

Wyniki przedstawionych symulacji pokazują, że nieskończone populacje potrafią zlokalizować optimum globalne badanej funkcji celu, nawet wtedy, gdy różnica wy-



Rys. 4.9. Symulowane rozkłady dużych populacji. Jednorodna populacja początkowa.

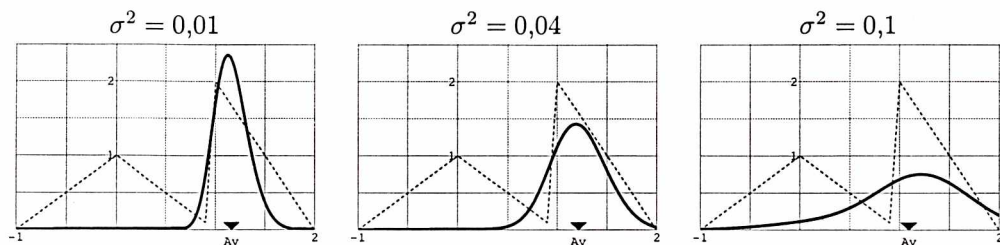
Punkty przedstawiają osobniki dla: a) jednomodalnej, b) dwumodalnej funkcji celu ($a = 5$, $h = 2$, $\sigma^2 = 0,04$). W lewym rogu podano histogramy populacji wynikowej

Fig. 4.9. Simulated distributions of big populations initiated with identical individuals.

Points denote individuals in the landscape of: unimodal (a), bimodal (b) fitness functions ($a = 5$, $h = 2$, $\sigma^2 = 0.04$). Upper-left corner – the histogram of a final population

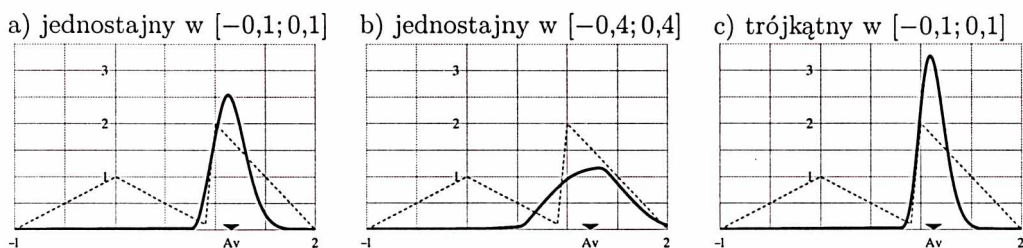
sokości szczytów jest znaczna czy siodło jest bardzo głębokie. Przyjęcie niewielkich wartości wariancji mutacji σ^2 spowalnia zbieżność, ale pozwala na dokładniejszą lokalizację ekstremum, a w przypadku jednakowych wzgórz prowadzi do otrzymania rozkładów dwumodalnych. Przytoczone badania rozkładów nieskończonych porównano z symulacją populacji bardzo dużych, lecz skończonych (rys. 4.9). Populacje te są rozłożone wokół optimów: lokalnego dla funkcji jednomodalnej i globalnego dla funkcji dwumodalnej. Im liczniejsza populacja, tym bardziej otrzymane rozkłady przypominają rozkład normalny przewidziany w analizie teoretycznej. Populacja mająca kilka tysięcy osobników wystarczy do dobrej aproksymacji modelu nieskończonego.

Analizowano także w jaki sposób założenie, że zarówno rozkład populacji początkowej, rozkład mutacji, jak i funkcje przystosowania są gaussowskie wpływa na zachowanie nieskończonej populacji. Rozkłady nieskończonych populacji zasymulowano dla innej funkcji celu – trójkątnej niesymetrycznej i dwumodalnej (patrz Dodatek, funkcja (D.6)) (rys. 4.10). Rozkłady stacjonarne przypominają rozkłady otrzymane dla gaussowskiej funkcji dwumodalnej (por. rys. 4.4). Dla tego samego krajobrazu przetestowano również inne rozkłady mutacji: jednostajny oraz trójkątny (Simpsona). Założono, że populacja początkowa jest prawie homogeniczna – zawiera tylko typy quasi-optymalne. Rozkład populacji początkowej jest mocno



Rys. 4.10. Rozkłady nieskończonej populacji (4.18) w 100. generacji, w krajobrazie trójkątnej funkcji (D.6), dla różnych variancji mutacji i $\nu_0^2 = 0,1$, $a = 5$, $h = 2$

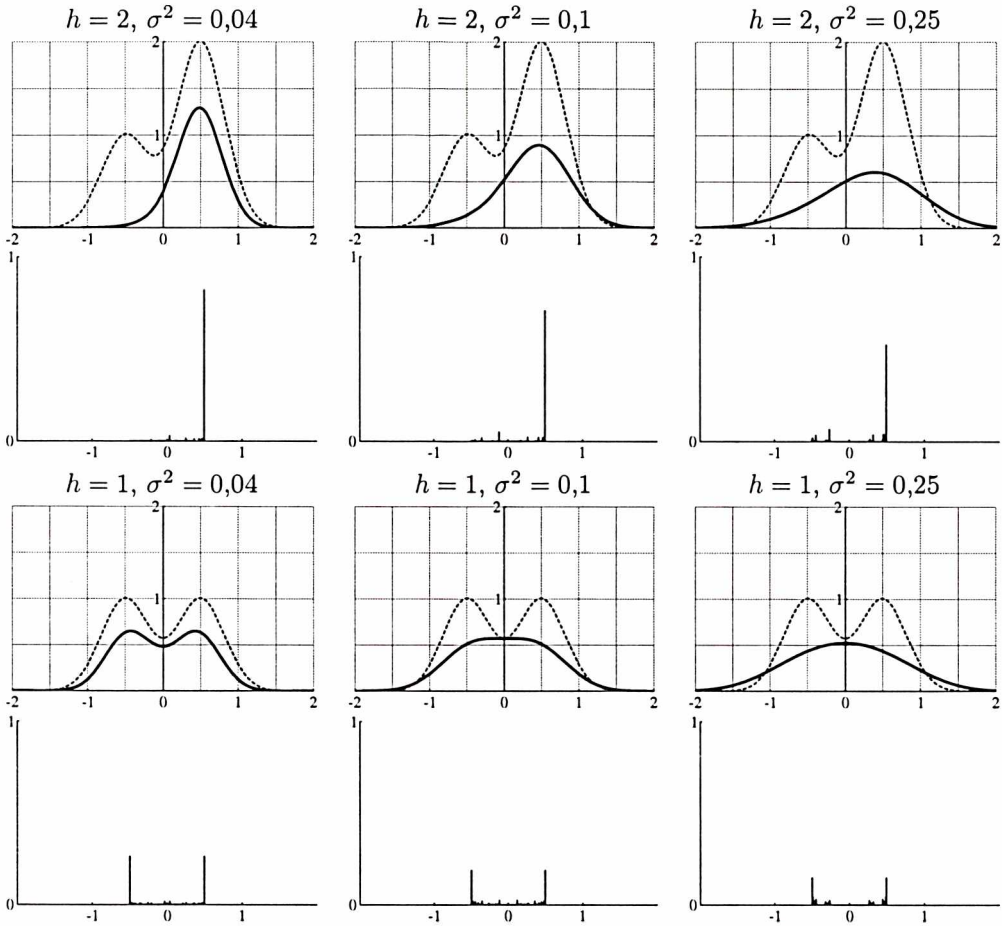
Fig. 4.10. Distributions of trait in the 100th generation in the landscape of the tent fitness function (D.6) for different mutation variances and $\nu_0^2 = 0.1$, $a = 5$, $h = 2$



Rys. 4.11. Rozkłady nieskończonej populacji (4.18) w 100. generacji, w krajobrazie trójkątnej funkcji (D.6), dla różnych rozkładów mutacji. Rozkład populacji początkowej jest jednostajny w przedziale $[-0,01;0,01]$, skoncentrowany na optimum lokalnym

Fig. 4.11. Distributions of trait (4.18) in the 100th generation in the landscape of the tent fitness function (D.6) for various distributions of mutation. Distribution of the initial population is uniform in the range $[-0.01;0.01]$, centered in the local optimum

skoncentrowany na optimum lokalnym i przypomina deltę Diraca. Przedstawione na rys. 4.11 rozkłady stacjonarne w 100. generacji przypominają rozkłady dla funkcji gaussowskich (por. rys. 4.5). Podobnie jak uprzednio, duży zasięg mutacji przyspiesza proces dochodzenia do stanu równowagi, ale powoduje jednocześnie większe niedokładności w lokalizacji ekstremum globalnego. Zaobserwowano ciekawe zachowania początkowego rozkładu aproksymującego deltę Diraca. Zanim główna masa rozkładu populacji zajęła wzgórze globalne, rozkład się spłaszczył, jego wartość maksymalna szybko zmalała, podczas gdy wartość oczekiwana praktycznie nie zmieniła swojego położenia w pobliżu optimum lokalnego. Kiedy fragment rozszerzającego się rozkładu przekroczył siodło, bardzo szybko, bo już po kilku generacjach rozkład przesunął się i ustabilizował na wzgórzu globalnym. Z przeprowadzonych analiz można więc wnioskować, że przyjęcie funkcji gaussowskich do badań nie ogranicza ogólności otrzymanych rezultatów.



Rys. 4.12. Rozkład graniczny nieskończonej populacji (wiersz górny) oraz wagi rozkładów składowych i granicznych (wiersz dolny) uzyskany za pomocą metody fraktalnej, dla różnych parametrów modelu, $a = 5$

Fig. 4.12. Limit distribution of trait (the first row); weights of component and limit distributions (the second row) obtained by applying the fractal method, for different parameters of the model, $a = 5$

Ponieważ parametry rozkładów składowych (4.19) i (4.20) opisują deterministyczny system iteracyjny [156], zaproponowano fraktalną metodę analizy rozkładów granicznych [116], dzięki której można z dużą dokładnością oszacować parametry kolejnych rozkładów bez względu na zasoby pamięciowe i obliczeniowe komputera. Pomija się w niej rozkłady generacji początkowych, pamiętając tylko K końcowych, gdzie K jest parametrem metody i zależy od zasobów obliczeniowych komputera. Oszacowane błędy aproksymacji metody są dostatecznie małe,

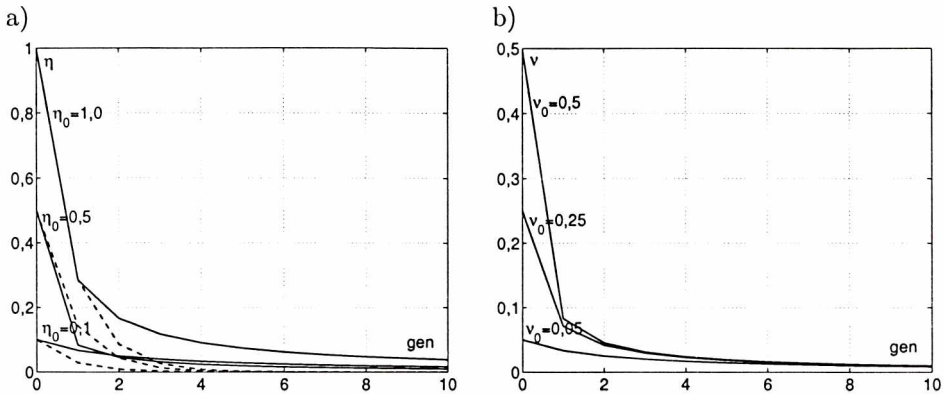
aby uznać metodę za wiarygodną. Na rysunku 4.12 przedstawiono uzyskane metodą fraktalną wykresy wag rozkładów składowych oraz granicznych dla różnych wartości wariancji modyfikacji. Otrzymane rezultaty potwierdzają obserwację, że ewolucja nieskończonych populacji z małymi wartościami σ^2 najdokładniej lokalizuje optimum globalne. Uzyskano także dwumodalny rozkład graniczny dla małych wariancji mutacji oraz wzgórz o jednakowej wysokości.

4.3.4. Dyskusja

Zbadajmy wpływ mutacji na zachowanie modelu nieskończonych populacji. Rozważmy przypadek, gdy mutacja nie występuje. W tym przypadku $p_{i+1}(x) = p'_i(x)$ i rozkład populacji w kolejnej generacji opisany jest równaniem (4.1). Można też zbadać rozkład po selekcji i mutacji $p_{i+1}(x)$ dla małych (dążących do zera) wartości wariancji mutacji. Mutacja znacząco przyspiesza proces zbieżności do rozkładu stacjonarnego, a zatem i do optimum globalnego (zobacz rys. 4.4). Jednocześnie, podtrzymując różnorodność populacji, mutacja o dużym zasięgu nie pozwala na dokładną lokalizację ekstremum. Gdy rozpatrzymy przypadek lokalny, bez mutacji ($\sigma^2 = 0$), wartość oczekiwana rozkładu (4.4) jest równa $\nu_{i+1}^2/\nu_i = 1/(1 + 2a\nu_i^2) < 1$, a współczynnik zawężania wynosi $\eta_{i+1} = \eta_i\nu_{i+1}^2/\nu_i$. Gdy $i \rightarrow \infty$, wówczas $\nu_i \rightarrow 0$ i współczynnik zawężania zbliża się do jedynki. Wtedy wartość oczekiwana bardzo wolno dąży do jej wartości stacjonarnej. Gdy mutacja jest obecna, wariancja rozkładu populacji, zgodnie z (4.6), nie może być mniejsza niż $\sigma^2/2$. Tak więc różnorodność w populacji jest zachowana, a zbieżność do rozkładu stacjonarnego jest szybsza niż w przypadku bez mutacji (rys. 4.13a). Zmiany wariancji rozkładu (4.16) w krajobrazie jedno- i dwumodalnej funkcji przystosowania, gdy nie występuje mutacja ($\sigma^2 = 0$), przedstawiono na rys. 4.13b. Widoczny efekt szybkiego ujednolicania populacji może spowodować, że rozkład populacji nie skoncentruje się na optimum globalnym, lecz na pewnym punkcie w jego otoczeniu (przykład tzw. przedwczesnej zbieżności rozkładu) (rys. 4.13a).

Przedstawiona analiza modelu nieskończonych populacji pozwala na sformułowanie pewnych spostrzeżeń dotyczących konstrukcji fenotypowych algorytmów ewolucyjnych. Aby zachować dobre właściwości eksploatacyjne i eksploracyjne poszukiwań ewolucyjnych, zasięg mutacji powinien być większy na początku poszukiwań i maleć w ich trakcie. Spostrzeżenie potwierdza wyniki uzyskane w pracy [163]. Wniosek ten jest znany i wykorzystywany w praktyce, np. adaptacja parametru σ mutacji w strategiach ewolucyjnych [8, 46, 170] czy zmiana temperatury w symulowanym wyżarzaniu [121].

Interesującą właściwością modelu nieskończonych populacji wydaje się występowanie dwumodalnych rozkładów w jednorodnych populacjach, dla których wszystkie cechy należą do jednej klasy. W przypadku wzgórz o równych wysokościach i małych σ^2 mutacji populacje zajmują oba optima, nie wyróżniając żadnego.



Rys. 4.13. a) Wartości oczekiwane η_i rozkładu populacji dla różnych parametrów rozkładu początkowego η_0 , $\nu_0^2 = 0,25$. Linie przerywane – bez mutacji ($\sigma^2 = 0$), ciągle – z mutacją ($\sigma^2 = 0,16$). Jednomodalna funkcja celu (D.2), $a = 5$.
b) Wariancja ν_i^2 w ewolucji bez mutacji ($\sigma^2 = 0$) dla różnych wartości ν_0^2 .

Jedno- i dwumodalne gaussowskie funkcje celu, $a = 5$, $h = 2$

Fig. 4.13. a) Expected value of distribution for various parameters of the initial distribution η_0 , $\nu_0^2 = 0,25$. Evolution without mutation ($\sigma^2 = 0$) marked with solid lines, and evolution with mutation ($\sigma^2 = 0.16$) with dotted lines.

Unimodal fitness (D.2), $a = 5$.

b) Variance ν_i^2 for evolution without mutation ($\sigma^2 = 0$) for different values of ν_0^2 .

Uni- and bimodal fitness, $a = 5$, $h = 2$

Rozkład równowagi może być traktowany jako suma dwóch rozkładów skoncentrowanych na poszczególnych szczytach. Wartość oczekiwana rozkładu łącznego jest jednak położona w obszarze siodła. Znając jedynie wartość oczekiwaną rozkładu, można traktować go jako jednomodalny i skupiony między optimami. W rzeczywistości jednak tylko bardzo mała część populacji ma cechy zbliżone do łącznej wartości oczekiwanej, podczas gdy większość ma cechy bliskie jednemu lub drugiemu ekstremum. Opisany przykład wskazuje jak istotnie różne wyniki można otrzymać, analizując całe rozkłady lub aproksymując je ze statystyk rozkładów (patrz też [159]).

4.3.5. Zbiór Cantora generowany przez wartości oczekiwane rozkładu

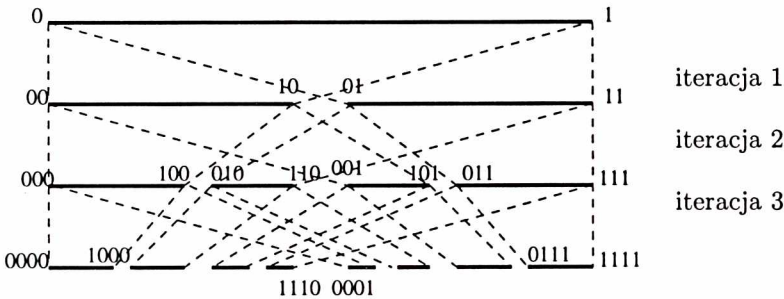
Ewolucja rozkładów nieskończonych populacji przypomina proces generowania fraktali przez układy odwzorowań iterowanych (ang. *iterated function system* IFS) [156]. Fakt ten wykorzystano, między innymi, w opisanej powyżej fraktalnej metodzie analizy rozkładów granicznych. Bliższe badania wartości oczekiwanych rozkładu pokazują także inne ciekawe fraktalne właściwości badanego modelu nieskończonych populacji [101]. W krajobrazie dwumodalnej funkcji przysto-

sowania, dla $d = 1$, równania opisujące wartości oczekiwane frakcji lokalnej i globalnej (4.12), (4.13) definiują afiniczne przekształcenia zwięzające postaci

$$F_{\text{loc}}(x) = cx,$$

$$F_{\text{glob}}(x) = cx + (1 - c), \quad (4.25)$$

gdzie $c = 1/z = 1/(1 + 2a\nu_s^2) < 1$, a x jest położeniem wartości oczekiwanej danego rozkładu. Jeżeli rozpatrzmy dwa normalne rozkłady początkowe, ulokowane jeden w optimum lokalnym, a drugi w globalnym, to przyjmujemy, że ich wartości oczekiwane wyznaczają odcinek $[0, d]$. Iteracja transformacji (4.25) przekształca odcinek $[0, d]$ w dwa odcinki (o łącznej długości równej $c < 1$). Krańce odcinków są wyznaczone przez wartości oczekiwane rozkładów wygenerowanych na podstawie rozkładów bieżącej iteracji. W granicy (dla nieskończonej liczby iteracji) otrzymujemy zbiór Cantora. Wynik przekształcenia zależy od parametrów a oraz σ^2 (przez wpływ na ν_s^2). Kilka pierwszych iteracji generujących zbiór Cantora przedstawiono na rys. 4.14. Klasyczny zbiór na odcinku $[0, 1]$ i ze współczynnikiem zawężania 3 otrzymamy dla $d = 1$ i $z = 3$, czyli gdy $a\nu_s^2 = 1$ i $a\sigma^2 = 2/3$.



Rys. 4.14. Zbiór Cantora tworzony przez ewolucję wartości oczekiwanych dwóch rozkładów początkowych ulokowanych w optimum funkcji przystosowania, $\sigma^2 = 0,04$
 Fig. 4.14. The Cantor set generated by the evolution of expected values of two initial distributions of trait in optima of the fitness function, $\sigma^2 = 0.04$

Skomplikowane twory fraktalne można więc odkryć nawet w tak prostym modelu. Jest to kolejny przykład potwierdzający słynne słowa M.F. Barnsleya: *fractals are everywhere* [11].

5. Dynamika bardzo małych populacji

W rozdziale 4 przedstawiono analizę modelu ewolucji fenotypowej z miękką selekcją dla nieskończonych populacji. Takie uproszczenie, aczkolwiek bardzo popularne w literaturze, może czasami prowadzić do mylnych wniosków, ponieważ w rzeczywistości symulowane populacje są zawsze skończone. Zwykle trajektorie badane w modelach nieskończonych nie pokazują zachowania pojedynczych uruchomień algorytmu, ale raczej przebiegi uśrednione po wszystkich możliwych realizacjach. Tak jest dla badanego modelu z miękką selekcją w krajobrazach gaussowskich funkcji celu, gdzie trajektorie nieskończonych populacji dobrze aproksymują uśrednione przebiegi dużych populacji. W małych populacjach obserwuje się wysoki stopień fluktuacji zachowań z generacji na generację. Predykcje ich ewolucji, uzyskane z pomocą modeli nieskończonych, mogą być niedokładne, a często takie uśrednione wyniki są diametralnie różne od rzeczywistych zachowań populacji [125, 159, 199, 203].

Analiza drugiej skrajności na osi liczebności populacji, mianowicie bardzo małych populacji, może przynieść wnioski bliższe rzeczywistości, tym bardziej, że w wielu zastosowaniach praktycznych stosuje się właśnie małe populacje. Badania symulacyjne wykazały, że niewielkie populacje zachowują się co prawda bardziej chaotycznie niż liczebniejsze, ale są bardziej mobilne, a więc bardziej efektywne w przekraczaniu siodła między optimami (także tych bardzo szerokich i wielowymiarowych) [29, 64, 186]. Dodatkowo, wykorzystanie małych populacji jest niekiedy konieczne w problemach, dla których obliczanie funkcji celu jest bardzo czasochłonne bądź kosztowne.

Ze względu na złożony charakter opisujących je procesów stochastycznych, niewiele prac poświęconych jest analizie teoretycznej małych populacji. Również i w tym przypadku dokonuje się dodatkowych założeń upraszczających, na przykład populacje opisywane są w terminach średnich lub innych makroskopowych charakterystyk. Dla problemów dyskretnych o małych rozmiarach bardzo małe populacje mogą być opisywane wprost za pomocą teorii łańcuchów Markowa (zobacz rozdz. 3.2). Także podejście wykorzystujące mechanikę statystyczną stosuje się do skończonych populacji [159] (zobacz rozdz. 3.4). Badania wartości oczekiwanych dla skończonych populacji wektorów binarnych, w tym dwuelementowych, opisano w [22]. Rozważano tylko operator krzyżowania, bez mutacji. Podobnie jak w modelu Vose'a [203], założono, że przestrzeń poszukiwań jest skończona, a populacje opisywano jako wektor częstości występowania w niej danego typu. Określono warunki, dla których właściwości populacji mogą być wystarczająco dokładnie określone przez analizę tylko wartości oczekiwanych populacji. Poka-

zano, że przy pewnych założeniach dotyczących długości wektorów, populacja jest zbieżna do stanu, w którym proporcje osobników są w równowadze (ang. *linkage equilibrium*). Wykorzystując między innymi populacje dwuelementowe, Culberson [30] badał teoretycznie działanie mutacji i krzyżowania w AG, analizując strukturę przestrzeni poszukiwań powodowaną działaniem tych operatorów.

W niniejszym rozdziale przedstawiono analizę zachowań dwuelementowych populacji dla modelu poszukiwań ewolucyjnych z miękką selekcją, badając trajektorie systemu dynamicznego generowanego przez wartości oczekiwane stanu populacji. Zaprezentowane wyniki oparto na pracach [28, 44, 70, 71, 97, 101, 105, 104, 106, 108, 109, 110, 111, 113].

5.1. Przestrzeń stanów populacji

Ewoluuujące populacje rozpatrywano dotychczas w *przestrzeni typów*, czyli osobników $T = R^n$, gdzie każdy punkt przestrzeni odpowiadał typowi elementu. Innym podejściem może być analiza w *przestrzeni stanów populacji* S , zawierającej wszystkie możliwe populacje. Cała populacja jest więc reprezentowana przez pojedynczy punkt, a ewolucja jest rozpatrywana jako ruch w tej przestrzeni. Analizę stanów populacji można spotkać w pracach poświęconych teorii metod ewolucyjnych wykorzystujących łańcuchy Markowa oraz teorię układów dynamicznych (zobacz rozdz. 3.2 i 3.3). Stosowane w tym rozdziale podejście różni się znacząco od teorii Vose'a [203], gdzie nieskończone populacje ewoluują w skończonej dyskretnej przestrzeni o K stanach. Wymiar przestrzeni nie zależy od wielkości populacji, gdyż stan opisany jest przez K -wymiarowy wektor częstości występowania typów w populacji. Populacje są multizbiorami (ang. *multi-sets*), nieuporządkowanymi zbiorami wieloelementowymi z powtarzającymi się elementami. W prezentowanym przypadku przestrzeń stanów populacji S jest nieograniczona i ciągła, a jej wymiar zależy wprost od wielkości populacji. Ponieważ dynamika populacji nie zależy od uszeregowania w niej osobników, populacja może być traktowana jako uporządkowany zbiór m -elementowy (ang. *m-tuple*), a nie jako zbiór nieuporządkowany.

Przestrzeń stanów S ma bardziej złożoną strukturę niż przestrzeń typów T . Wymiar przestrzeni S zależy od wielkości populacji i jest m razy większy od wymiaru przestrzeni T : $\dim(S) = n \cdot m$. Ze względu na niezależność ewolucji od kolejności ustawienia osobników w populacji (populacja dwuelementowa $P = \{\mathbf{x}_1, \mathbf{x}_2\}$ jest równoważna populacji $P = \{\mathbf{x}_2, \mathbf{x}_1\}$), w celu ułatwienia analizy zdefiniowano relację równoważności U uporządkowanych ciągów. Relacja U utożsamia wszystkie punkty odpowiadające permutacjom osobników w populacji. W konsekwencji przestrzeń stanów populacji staje się przestrzenią ilorazową, z relacją równoważności U : $S_U = S/U = R^{m \cdot n}/U$.

Funkcja gęstości prawdopodobieństwa rozkładu łącznego stanu dwielementowej populacji w przestrzeni stanów S jest iloczynem FGP m rozkładów (2.3). Przyjmijmy zasadę wyboru jednego z nierozróżnialnych (równoważnych względem relacji równoważności U) elementów zbioru. Niech $\mathbf{s}^i = (\mathbf{x}_1^i, \mathbf{x}_2^i, \dots, \mathbf{x}_m^i)$ reprezentuje stan populacji w generacji i -tej zgodnie z przyjętą relacją równoważności U . W przestrzeni ilorazowej S_U rozkład stanu populacji w $(i+1)$ generacji dany jest formułą

$$\tilde{f}_{S_U}^{i+1}(\mathbf{s}|\mathbf{s}^i) = m! \prod_{j=1}^m f_T^{i+1}(\mathbf{x}_j|\mathbf{s}^i) = m! \prod_{j=1}^m \sum_{k=1}^m \alpha(\mathbf{x}_k^i|\mathbf{s}^i) g(\mathbf{x}_j, \mathbf{x}_k^i), \quad (5.1)$$

gdzie $\alpha(\mathbf{x}_k^i|\mathbf{s}^i)$ opisuje prawdopodobieństwo wyboru elementu $\mathbf{x}_k^i$, a $g(\mathbf{x}, \mathbf{x}_k^i)$ jest rozkładem mutacji k -tego osobnika, $g(\mathbf{x}, \mathbf{x}_k^i) = \mathcal{N}(\mathbf{x}_k^i, \sigma^2)$. Funkcja przystosowania jest naturalnie zdefiniowana w przestrzeni typów T . Po wprowadzeniu przestrzeni stanów należy określić w niej odpowiednik funkcji przystosowania. Wydaje się, że przyjęcie jako wskaźnika przystosowania całej populacji jej średniego przystosowania (rys. 5.1) jest najbardziej naturalne¹. W biologii fenotyp średniego osobnika definiuje się także w krajobrazie adaptacyjnym średniego przystosowania [214]. W algorytmach ewolucyjnych krajobraz średniego przystosowania wykorzystuje się w modelach fizyki stochastycznej [159] oraz genetyki populacyjnej [144]. Powierzchnie jakości w S_U mają bardziej skomplikowaną strukturę niż w przestrzeni T , przede wszystkim mogą pojawić się nowe optima lokalne (rys. 5.1). Ewolucja w przestrzeni stanów generuje więc nieskończoną sekwencję stanów populacji $\mathbf{s}^0, \mathbf{s}^1, \mathbf{s}^2, \dots$ opisujących ruch populacji w krajobrazie średniego przystosowania.

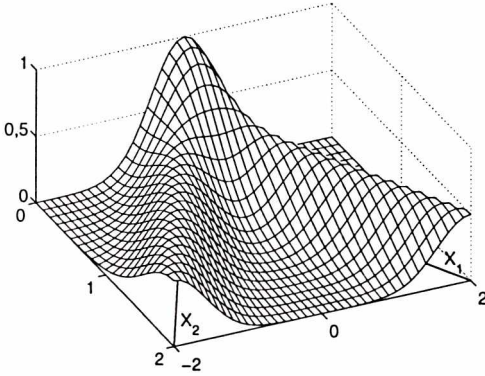
Pomimo skomplikowanej struktury przestrzeni stanów, analiza populacji w tej przestrzeni przynosi korzyści, zwłaszcza wtedy, gdy rozpatrujemy bardzo małe populacje. W przypadku populacji dwielementowych ($m = 2$) i jednowymiarowej przestrzeni poszukiwań ($n = 1$) przestrzeń S jest dwuwymiarowa, co umożliwia przedstawienie całej populacji jako punktu na płaszczyźnie. Dodatkowo, możliwe jest obliczenie rozkładów prawdopodobieństwa stanów populacji oraz ich wartości oczekiwanych i wariancji.

5.2. Dwieelementowa populacja w przestrzeni stanów

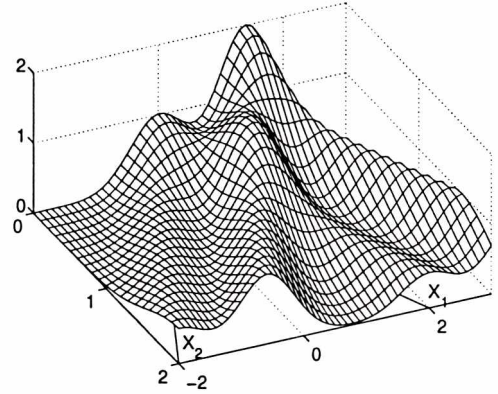
Rozpatrujemy dwielementową populację w jednowymiarowej przestrzeni poszukiwań $P = (x_1, x_2)$. Ponieważ przestrzeń stanów jest dwuwymiarowa, stan populacji jest punktem w R^2 . W tym przypadku FGP rozkładu położenia osobnika x w przestrzeni typów T w $(i+1)$ generacji zadana jest jako

¹ Alternatywą jest przyjęcie maksymalnego/minimalnego przystosowania populacji.

a) $q(x) = \exp(-5x^2)$



b) $q(x) = \exp(-5x^2) + 2 \exp[-5(x-1)^2]$



Rys. 5.1. Powierzchnia średniego przystosowania w przestrzeni stanów populacji S_U

Fig. 5.1. Surface of average fitness in the space of states S_U

$$f_T^{i+1}(x|\mathbf{s}^i) = \alpha(x_1^i|\mathbf{s}^i)g(x, x_1^i) + \alpha(x_2^i|\mathbf{s}^i)g(x, x_2^i), \quad (5.2)$$

gdzie $\alpha(x_k^i) = q(x_k^i)/(q(x_1^i) + q(x_2^i))$, $k = 1, 2$. W ilorazowej przestrzeni stanów S_U FGP rozkładu stanu dwuelementowej populacji (5.1) jest dana wzorem

$$\begin{aligned} \tilde{f}_{S_U}^{i+1}(x_1, x_2|\mathbf{s}^i) &= 2f_T^{i+1}(x_1|\mathbf{s}^i)f_T^{i+1}(x_2|\mathbf{s}^i) = 2(\alpha(x_1^i|\mathbf{s}^i)g(x_1, x_1^i) \\ &+ \alpha(x_2^i|\mathbf{s}^i)g(x_1, x_2^i))(\alpha(x_1^i|\mathbf{s}^i)g(x_2, x_1^i) + \alpha(x_2^i|\mathbf{s}^i)g(x_2, x_2^i)). \end{aligned} \quad (5.3)$$

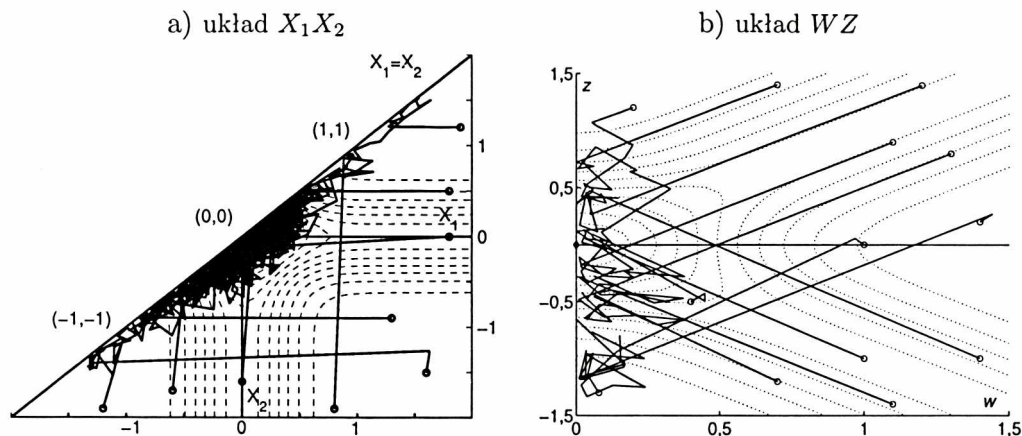
Zdefiniujmy relację równoważności U , porządkującą malejąco elementy populacji:

$$R^2 \rightarrow S_U \subset R^2 : (x_1^i, x_2^i) \rightarrow \begin{cases} (x_1^i, x_2^i) & \text{dla } x_1^i \geq x_2^i \\ (x_2^i, x_1^i) & \text{dla } x_1^i < x_2^i \end{cases}.$$

Relacja U zawęży przestrzeń sklejając, (utożsamiając) punkty symetryczne względem prostej $X_1 = X_2$ oraz identyfikując przestrzeń ilorazową S_U z prawą półpłaszczyzną ograniczoną przez prostą. Prostą $X_1 = X_2$ nazwano *osią identyczności*, ponieważ osobniki leżące na niej mają identyczne typy (populacja jest homogeniczna) (rys. 5.2a).

Obrót układu współrzędnych X_1X_2 o kąt $\gamma = \pi/4$ w kierunku przeciwnym do ruchu wskazówek zegara ułatwia analizę ewoluujących populacji w przestrzeni stanów S_U (rys. 5.2b). Obrót przekształca współrzędne x_1, x_2 w nowe współrzędne w i z :

$$\begin{bmatrix} w \\ z \end{bmatrix} = \begin{bmatrix} \cos \gamma & -\sin \gamma \\ \sin \gamma & \cos \gamma \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \frac{\sqrt{2}}{2} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}. \quad (5.4)$$



Rys. 5.2. Ewolucja dwuelementowej populacji w przestrzeni stanów S_U w dwóch układach współrzędnych. Pokazano 10 generacji dla różnych stanów początkowych (oznaczonych $\circ$) w krajobrazie funkcji $q(x) = \exp(-5x^2)$, $\sigma = 0,1$

Fig. 5.2. Evolution of two-element population in the space of states S_U in two coordinate systems. The 10 generations for different initial states (depicted as $\circ$) are shown in landscape of the fitness $q(x) = \exp(-5x^2)$, $\sigma = 0.1$

Tak więc stan populacji w i -tej generacji opisany współrzędnymi $\mathbf{s}^i = (x_1^i, x_2^i)$ zostaje przetransformowany w stan $\tilde{\mathbf{s}}^i = (w^i, z^i)^2$, dany zależnościami

$$w^i = (x_1^i - x_2^i)/\sqrt{2}, \quad z^i = (x_1^i + x_2^i)/\sqrt{2}.$$

W nowym układzie współrzędnych WZ ilorazowa przestrzeń stanów S_U jest ograniczona do prawej półpłaszczyzny ($w \geq 0$), a oś identyczności staje się osią Z układu współrzędnych. Współrzędne w i z mają dogodne interpretacje. Współrzędna w opisuje odległość stanu populacji od osi identyczności. Ponieważ $w = \sqrt{2}D[\mathbf{x}]$, gdzie $D[\mathbf{x}]$ jest odchyleniem standardowym populacji, może być więc interpretowane jako miara różnorodności populacji $P(\mathbf{x})$. Im mniejsze w , tym mniejsze różnice w typach osobników i populacja jest bardziej jednorodna. Współrzędna z określa położenie stanu populacji wzdłuż osi identyczności i jest równa $z = \sqrt{2}E[\mathbf{x}]$, gdzie $E[\mathbf{x}]$ jest średnią populacji P .

W zależności od funkcji przystosowania i stanu bieżącego, rozkład prawdopodobieństwa zmiany stanu populacji w kolejnej generacji (5.3) może być jedno- lub wielomodalny [28, 70, 106]. Rozkłady (5.3) oraz przykładowe przebiegi ewolucji w układzie współrzędnych WZ przedstawiono w lewej i środkowej kolumnie rys. 5.3 i 5.4 dla czterech różnych stanów początkowych w krajobrazach odpowiednio jedno- oraz dwumodalnej gaussowskiej funkcji celu. Dla jednomodalnej funkcji celu

² Dalej, dla uproszczenia zapisu pomijamy znak tyldy przy $\tilde{\mathbf{s}}$ i identyfikujemy kontekstowo przestrzeń.

$q(x) = \exp(-5x^2)$ (D.2), której maksimum przypada dla $x = 0$, rozkład prawdopodobieństwa zmiany stanu początkowego $\mathbf{s}^0 = (0,0; -0,8)$ (rys. 5.3a, lewa kolumna) jest jednomodalny z optimum w stanie początkowym. Położenie populacji w kolejnej generacji zmieni się więc nieznacznie i będzie wynikiem działania mutacji. Wielkość przesunięcia zależy od odchylenia standardowego mutacji. Rozkłady prawdopodobieństwa (5.3) dla kolejnych stanów początkowych są już wielomodalne. Dla stanów $\mathbf{s}^0 = (0,4; -0,4)$ i $\mathbf{s}^0 = (0,8; -0,1)$ (rys. 5.3b, 5.3c, lewa kolumna) optimum globalne rozkładu położone jest w pobliżu maksimum funkcji celu. Prawdopodobieństwo, że populacja w kolejnej generacji znajdzie się w pobliżu maksimum przystosowania jest więc dla tych stanów początkowych największe. Populacje mogą jednak, z niezerowym prawdopodobieństwem, pozostać w sąsiedztwie stanu początkowego lub znaleźć się w pobliżu osi identyczności, lecz w znacznej odległości od optimum (rys. 5.3c). Rozkład prawdopodobieństwa zmiany stanu populacji znajdującej się początkowo w $\mathbf{s}^0 = (0,9; 0,0)$ (rys. 5.3d, lewa kolumna) ma trzy optima o podobnej wysokości, lecz żadne z nich nie jest ulokowane w pobliżu maksimum funkcji celu. Z podobnym prawdopodobieństwem populacja w kolejnej generacji albo zmieni swoje położenie tylko nieznacznie pod wpływem mutacji, albo skoczy do jednego z dwóch stanów w pobliżu osi identyczności, lecz odległych od optimum funkcji celu.

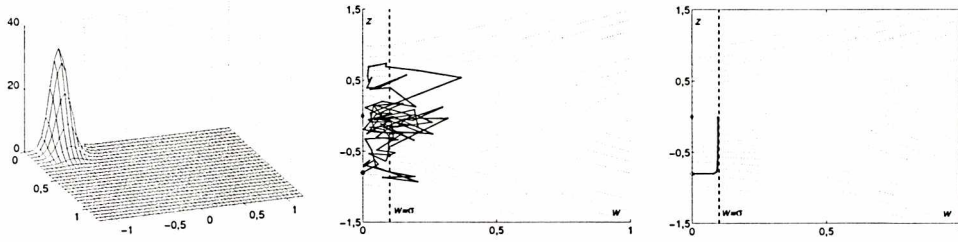
Dla dwumodalnej funkcji przystosowania $q(x) = \exp(-5x^2) + 2 \exp[-5(x-1)^2]$ (D.4) sytuacja jest analogiczna. Rozkład prawdopodobieństwa zmiany stanu populacji ze stanu $\mathbf{s}^0 = (1,3; 1,5)$ (rys. 5.4a, lewa kolumna), ma wyraźne maksimum w punkcie odpowiadającym optimum lokalnemu funkcji celu. Drugie optimum rozkładu, położone wokół stanu początkowego, jest tak niewielkie, że rozkład można traktować jako jednomodalny. Podobnie wygląda rozkład (5.3) ze stanu $\mathbf{s}^0 = (1,3; 0,0)$ (rys. 5.4c, lewa kolumna), ale skoncentrowany jest na optimum globalnym. W przypadkach b) i d) obserwujemy rozkład o trzech optimach, przy czym wierzchołki rozkładów przesunięte są albo w stronę optimum lokalnego, albo globalnego.

Przykładowe trajektorie 100 generacji dla dwuelementowej populacji w układzie współrzędnych WZ przedstawiono w środkowej kolumnie rys. 5.3 i 5.4. Dla wszystkich stanów początkowych obserwuje się „skok” populacji w pobliże osi Z , a następnie przeszukiwanie przestrzeni z mniejszym już „krokiem”. Dla przypadków wielomodalnych rozkładów populacji w krajobrazie dwumodalnej funkcji przystosowania (rys. 5.4b, d, kolumna środkowa) pokazano jakościowo odmienne realizacje trajektorii ewolucji obrazujące skoki populacji do różnych optimów.

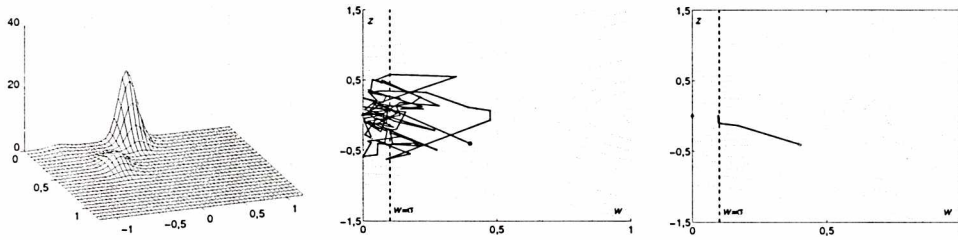
5.3. Analiza wartości oczekiwanych stanów populacji

Jak wspomniano wcześniej, proces ewolucyjny opisany rozkładami (5.1) jest procesem stochastycznym, którego analiza w ogólnym przypadku jest trudna. Przedstawienie ewoluujących populacji w przestrzeni stanów S_U umożliwia badanie śred-

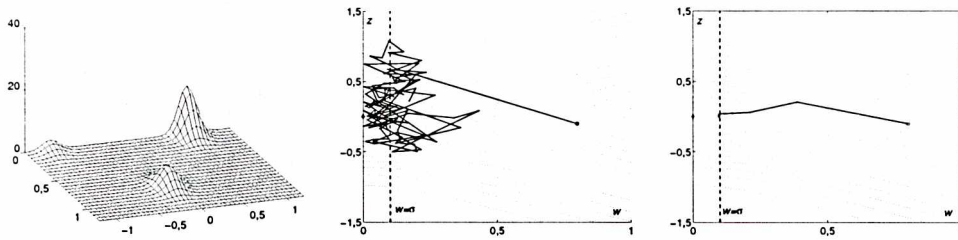
a) $s^0 = (0,0;-0,8)$



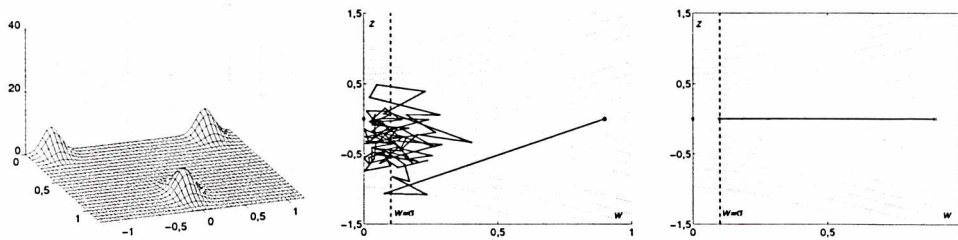
b) $s^0 = (0,4;-0,4)$



c) $s^0 = (0,8;-0,1)$

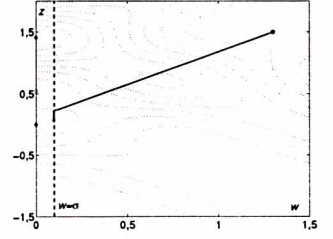
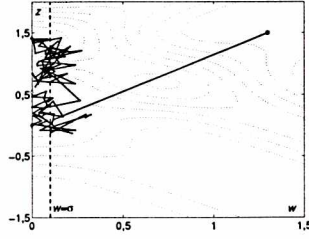
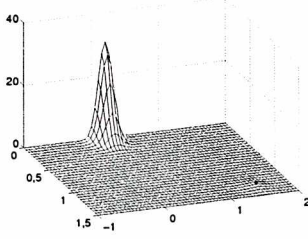
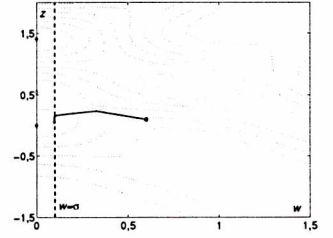
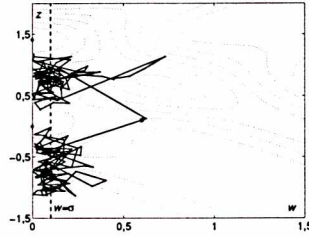
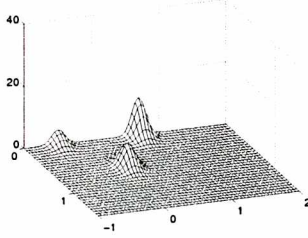
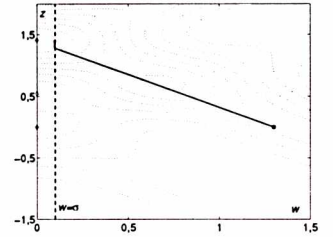
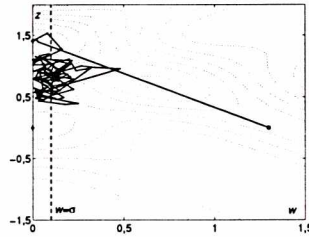
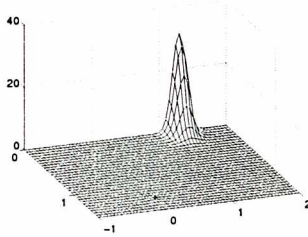
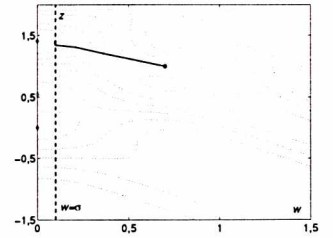
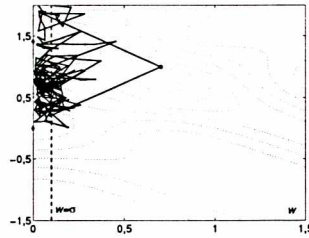
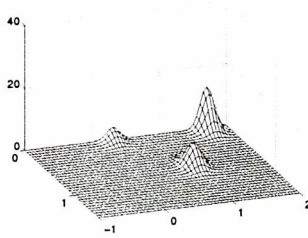


d) $s^0 = (0,9;0,0)$



Rys. 5.3. Dynamika stanów populacji w krajobrazie jednomodalnej funkcji celu (D.2) ($a = 5$, $\sigma = 0,1$) i wybranych stanów początkowych. W kolumnach przedstawiono: rozkłady prawdopodobieństwa zmiany stanu populacji w kolejnej generacji, przykładowe trajektorie ewolucji w 100 generacjach, trajektorie wartości oczekiwanych

Fig. 5.3. Dynamics of population states in the landscape of unimodal fitness function (D.2) ($a = 5$, $\sigma = 0.1$) and various initial states. In columns there are presented: the probability distributions of population state in the next generation, the examples of evolution trajectories in 100 generations, the trajectories of expected values

a) $s^0 = (1,3; 1,5)$ b) $s^0 = (0,6; 0,1)$ c) $s^0 = (1,3; 0,0)$ d) $s^0 = (0,7; 1,0)$ 

Rys. 5.4. Dynamika stanów populacji w krajobrazie dwumodalnej funkcji (D.4) ($a = 5$, $h = 2$, $d = 1$, $\sigma = 0,1$) i różnych stanów początkowych. W kolumnach przedstawiono: rozkłady prawdopodobieństwa zmiany stanu populacji w kolejnej generacji, przykładowe trajektorie ewolucji w 100 generacjach, trajektorie wartości oczekiwanych

Fig. 5.4. Dynamics of population states in the landscape of bimodal fitness (D.4) ($a = 5$, $h = 2$, $d = 1$, $\sigma = 0.1$) and various initial states. In columns there are presented: the probability distributions of population state in the next generation, the examples of evolution trajectories in 100 generations, the trajectories of expected values

niej dynamiki procesu. Mając daną wartość oczekiwaną rozkładu (5.1), która wskazuje oczekiwany stan populacji w kolejnej generacji, możemy rozpatrywać proces jako ewolucję reprezentanta wszystkich możliwych populacji w kolejnej generacji i wyznaczać trajektorie wartości oczekiwanych w czasie. Chociaż indywidualne realizacje procesu mogą się znacznie różnić od wartości uśrednionych, to uzyskana wiedza jest ważna, ponieważ umożliwia lepsze poznanie mechanizmów działających w procesach ewolucyjnych, zarówno rzeczywistych, jak i symulowanych.

Wartości oczekiwane stanu dwuelementowej populacji $\mathbf{s}^i = (w^i, z^i)$ w układzie współrzędnych WZ są obliczane na podstawie gęstości (5.3) wyrażonej w nowych współrzędnych. Wartości oczekiwane współrzędnych (w, z) wynoszą

$$E_{i+1} [w | \mathbf{s}^i] = \int_{\xi=0}^{\infty} \int_{z=-\infty}^{\infty} \xi \tilde{f}_{S_U}^{i+1}(\xi, z | \mathbf{s}^i) dz d\xi,$$

$$E_{i+1} [z | \mathbf{s}^i] = \int_{w=0}^{\infty} \int_{\xi=-\infty}^{\infty} \xi \tilde{f}_{S_U}^{i+1}(w, \xi | \mathbf{s}^i) d\xi dw.$$

Po podstawieniu mutacji gaussowskiej, zamianie zmiennych i żmudnych obliczeniach [44] otrzymano następujące wyrażenia opisujące wartości oczekiwane stanu:

$$E_{i+1} [w | \mathbf{s}^i] = \sqrt{\frac{2}{\pi}} \sigma + (1 - (\Psi^i)^2) \Theta(w^i / \sigma) \sigma, \quad (5.5)$$

$$E_{i+1} [z | \mathbf{s}^i] = z^i + \Psi^i w^i, \quad (5.6)$$

gdzie

$$\Psi(w, z) = \frac{q((w+z)/\sqrt{2}) - q((z-w)/\sqrt{2})}{q((w+z)/\sqrt{2}) + q((z-w)/\sqrt{2})}, \quad \Psi^i = \Psi(w^i, z^i),$$

$$\phi_0(\xi) = \frac{1}{\sqrt{2\pi}} \left[\exp\left(-\frac{\xi^2}{2}\right) - 1 \right], \quad \Phi_0(\xi) = \frac{1}{\sqrt{2\pi}} \int_0^\xi \exp\left(-\frac{t^2}{2}\right) dt,$$

$$\Theta(\xi) = \phi_0(\xi) + \xi \Phi_0(\xi).$$

Wprowadźmy oznaczenia dla wartości przystosowania osobników i -tej generacji:

$$q_1^i = q(x_1^i) = q((w^i + z^i)/\sqrt{2}), \quad q_2^i = q(x_2^i) = q((z^i - w^i)/\sqrt{2}),$$

wtedy $\Psi^i = \Psi(w^i, z^i) = (q_1^i - q_2^i)/(q_1^i + q_2^i)$.

Trajektorie wartości oczekiwanych dla jedno- i dwumodalnej gaussowskich funkcji przystosowania przedstawiono w prawych kolumnach rysunków 5.3 i 5.4. Na szczególną uwagę zasługuje fakt, że już w pierwszej iteracji następuje „skok” populacji w kierunku osi Z , widoczny także w trajektoriach ewolucji (kolumna środkowa), co oznacza, że początkowo duże rozproszenie populacji, mierzone współrzędną w , gwałtownie maleje. W kolejnych generacjach populacja bardzo wolno

zbliża się do punktu stałego (zaznaczonego gwiazdką): współrzędna w zmienia się nieznacznie lub wcale, zmiany współrzędnej z są rzędu σ . Przedstawiona dalej analiza teoretyczna pozwoli na wyjaśnienie tych zjawisk.

Za punkt wyjścia przyjmujemy równania (5.5) i (5.6). Prawa strona (5.5) jest sumą dwóch składników: pierwszy jest zależny jedynie od wartości odchylenia standardowego mutacji σ , drugi – od funkcji przystosowania (wpływającej na współczynnik Ψ) oraz bieżącego stanu populacji, a dokładniej od jego współrzędnej w . Określimy dolne i górne ograniczenie wartości oczekiwanej $E_{i+1}[w|s^i]$.

Twierdzenie 5.3.1. *Dla dowolnego stanu $s^i \in S_U$ wartość oczekiwana $E_{i+1}[w|s^i]$ jest ograniczona z dołu*

$$E_{i+1}[w|s^i] \geq \sqrt{\frac{2}{\pi}}\sigma. \quad (5.7)$$

Dowód: Funkcja przystosowania $q(x)$ jest nieujemna, tak więc $\Psi \in [-1, 1]$. Ponieważ $\forall x \geq 0 \Theta(x) \geq 0$, co oznacza, że drugi składnik sumy (5.5) jest nieujemny, więc twierdzenie 5.3.1 zachodzi. $\square$

Twierdzenie 5.3.2. *Dla dowolnego stanu $s^i \in S_U$ wartość oczekiwana $E_{i+1}[w|s^i]$ jest ograniczona od góry*

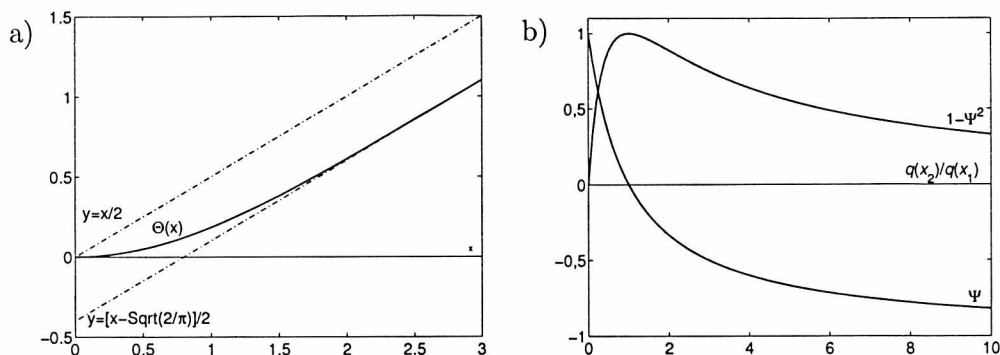
$$E_{i+1}[w|s^i] \leq \sqrt{\frac{2}{\pi}}\sigma + \frac{w^i}{2}. \quad (5.8)$$

Dowód: Dla $x > 0$ funkcja $\Theta(x)$ jest nieujemna, monotoniczna, wypukła. Gdy wartość argumentu x rośnie, $\phi_0 \rightarrow -1/\sqrt{2\pi}$, $\Phi_0 \rightarrow 1/2$, więc funkcja $\Theta(x) \rightarrow (x - \sqrt{2/\pi})/2$. Ponieważ $\phi_0 \leq 1/\sqrt{2\pi}$, a $\Phi_0 \leq 1/2$, więc funkcja $\Theta(x)$ jest ograniczona z góry przez prostą $y = x/2$. Po podstawieniu $x = w^i/\sigma$ zauważmy, że suma (5.5) może być ograniczona w sposób następujący: $E_{i+1}[w|s^i] \leq \sqrt{(2/\pi)}\sigma + (1 - (\Psi^i)^2)w^i/2$. Ponieważ $w^i \geq 0$ oraz $\Psi \in [-1, 1]$, górne ograniczenie (5.8) otrzymamy dla $\Psi^i = 0$. $\square$

Funkcja $\Theta(x)$ jest położona w obszarze ograniczonym od góry przez prostą $y = x/2$ i z dołu przez $y = (x - \sqrt{2/\pi})/2$ (rys. 5.5a). Z rysunku 5.5a oraz obliczeń asymptoty wynika, że dla wartości w takich, że $w/\sigma > 2$ funkcja $\Theta(x)$, która zależy od gęstości i dystrybucyj zmiennej losowej o rozkładzie normalnym $\mathcal{N}(0, 1)$, może być aproksymowana przez $\Theta(x) \simeq (x - \sqrt{2/\pi})/2$. Korzystając z tego przybliżenia, wartość oczekiwaną (5.5) dla $w > 2\sigma$ można aproksymować przez

$$E_{i+1}[w|s^i] \simeq \frac{1}{2} \left(\sqrt{\frac{2}{\pi}}(1 + (\Psi^i)^2)\sigma + (1 - (\Psi^i)^2)w^i \right). \quad (5.9)$$

Obliczenia $E_{i+1}[w|s^i]$ stają się w tym przypadku łatwiejsze, gdyż nie trzeba obliczać wartości funkcji $\Theta(x)$, a konieczna jest jedynie znajomość przystosowania osobników dla obliczenia funkcji Ψ . Zauważmy, że ograniczenia (5.7) i (5.8)



Rys. 5.5. Funkcje pomocnicze: a) $\Theta(x)$ leży między prostymi $y = x/2$ i $y = (x - \sqrt{2/\pi})/2$; b) Zależność Ψ oraz $1 - \Psi^2$ od stosunku jakości osobników $q(x_2)/q(x_1)$

Fig. 5.5. Auxiliary functions: a) $\Theta(x)$ lies between two lines $y = x/2$ and $y = (x - \sqrt{2/\pi})/2$; b) Ψ and $1 - \Psi^2$ depend on the fitness ratio of both individuals $q(x_2)/q(x_1)$

wartości oczekiwanej $E_{i+1}[w|\mathbf{s}^i]$ nie zależą od jakości osobników. Ujednolicanie populacji, mierzone przez współrzędną w , zachodzi więc w każdym krajobrazie funkcji przystosowania, nawet wtedy, gdy presja selekcyjna jest nieobecna (funkcja przystosowania w postaci płaskowyzu). Koncentracja populacji może następować natychmiast ($E_{i+1}[w|\mathbf{s}^i]$ bliskie wartości dolnego ograniczenia) lub w kilku krokach ($E_{i+1}[w|\mathbf{s}^i]$ w pobliżu górnego ograniczenia).

Jeżeli weźmiemy pod uwagę funkcję przystosowania, to wartość oczekiwana (5.5) osiągnie maksymalną wartość dla $\Psi \simeq 0$, ponieważ dla funkcji nieujemnych $\Psi \in [-1, 1]$ (rys. 5.5b). Współczynnik Ψ przyjmuje wartości bliskie zeru, gdy osobniki mają praktycznie jednakowe typy (stan populacji leży na osi Z lub w jej bliskim sąsiedztwie). $E_{i+1}[w|\mathbf{s}^i]$ osiąga wartość minimalną $\sqrt{2/\pi}\sigma$, kiedy różnice w przystosowaniu osobników są znaczne, czyli populacja jest znacznie rozproszona i $\Psi = \pm 1$, a więc $(1 - \Psi^2) \rightarrow 0$.

Ponieważ oczekiwane rozproszenie populacji zależy od jej bieżącej dyspersji (współrzędna w), wniosek 5.3.1 podaje wartości oczekiwane $E_{i+1}[w|\mathbf{s}^i]$ dla populacji o różnym rozproszeniu początkowym: I – homogeniczna populacja, składająca się z identycznych osobników, II – prawie homogeniczna populacja, składająca się z osobników z praktycznie jednakowymi typami, III – populacja różnorodna, składająca się z osobników o różnych typach.

Wniosek 5.3.1

I. Wartość oczekiwana współrzędnej w (rozproszenie) homogenicznej populacji ($w^i = 0$) zwiększa się

$$E_{i+1}[w|\mathbf{s}^i] = \sqrt{(2/\pi)\sigma}. \quad (5.10)$$

II. Wartość oczekiwana współrzędnej w prawie homogenicznej populacji zwiększa się

$$E_{i+1}[w|s^i] > w^i. \quad (5.11)$$

Nierówność (5.11) zachodzi dla $w^i < \beta\sigma$, gdzie współczynnik β jest obliczany z równania $E_{i+1}[w|s^i] = w^i$, przy założeniach $w^i = \beta\sigma$ i $(1 - (\Psi^i)^2) \simeq 1$, ($\beta \simeq 0,973$).

III. Wartość oczekiwana współrzędnej w różnorodnej populacji zmniejsza się

$$E_{i+1}[w|s^i] < w^i. \quad (5.12)$$

Dowód:

I. Homogeniczna populacja może przyjmować każdy stan na osi Z , gdzie $x_1^i = x_2^i$, $w^i = 0$, $q_1^i = q_2^i$ i $\Psi^i = 0$. Po podstawieniu danych do równania (5.5) otrzymamy (5.10).

II. Dla $w^i < \sqrt{2/\pi}\sigma$ prawa strona równania (5.11) wynika z twierdzenia 5.3.1. Jeżeli $w^i = \sqrt{2/\pi}\sigma$, to wartość oczekiwana $E_{i+1}[w|s^i] \simeq \sqrt{2/\pi}\sigma + 0,12(1 - (\Psi^i)^2)\sigma$ i przyjmując, że dla prawie homogenicznych populacji $(1 - (\Psi^i)^2) \simeq 1$ (rys. 5.5b), aproksymacja $E_{i+1}[w|s^i] \simeq 0,92\sigma > w^i$ zachodzi.

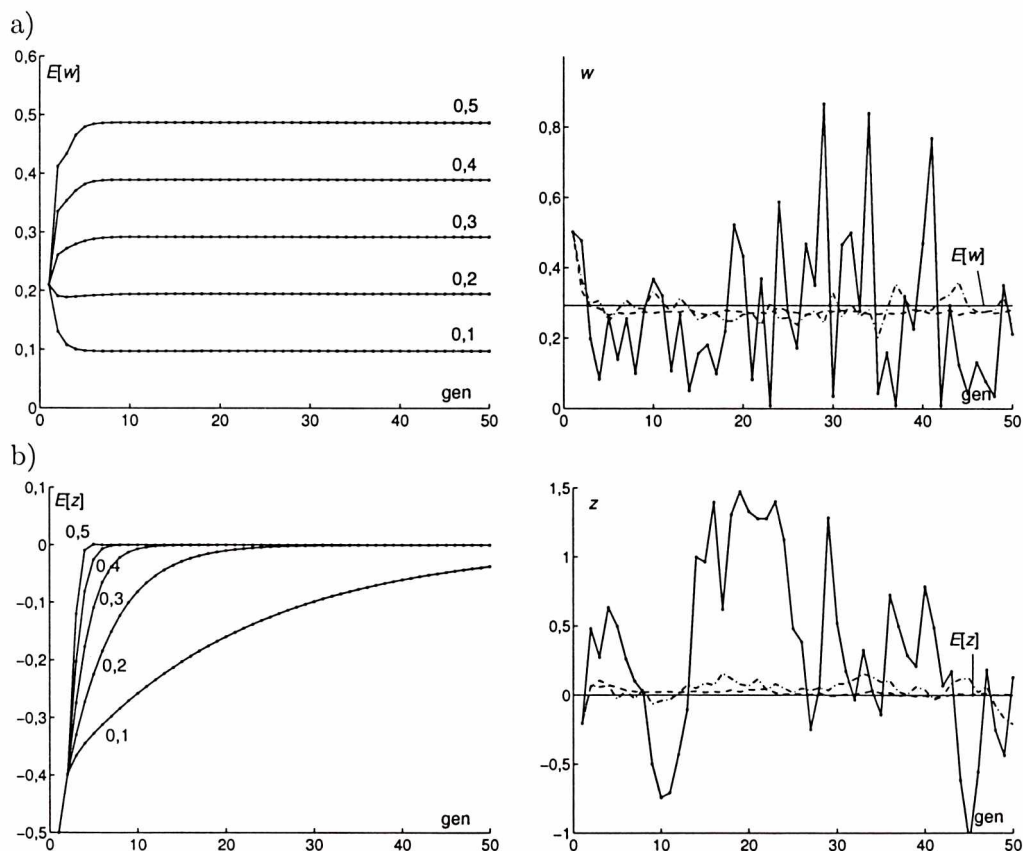
III. Dla różnorodnej populacji $\Psi^i \rightarrow 0$. Dla $w^i > 2\sigma$ stosujemy aproksymację (5.9). Wtedy $E_{i+1}[w|s^i] \simeq \sqrt{2/\pi}\sigma + w^i/2 < \sqrt{2/\pi}w^i/2 + w^i/2 < w^i$. $\square$

Z twierdzenia 5.3.2 i wniosku 5.3.1 wynika więc, że oczekiwana różnorodność początkowo rozproszonej populacji (duża wartość współrzędnej w) szybko maleje. Jeżeli wartość $w^i < 2\sqrt{2/\pi}\sigma$, to w kolejnej generacji oczekiwane rozproszenie populacji będzie mniejsze, $E_{i+1}[w|s^i] < w^i$. Zwiększenie wartości oczekiwanej współrzędnej w może nastąpić dla początkowych populacji homogenicznych lub prawie homogenicznych. Trajektorie wartości oczekiwanych jednorodnych populacji są odpychane od osi Z w kierunku linii $w = \sqrt{2/\pi}\sigma$. Mutacja zwiększy rozproszenie populacji i w kolejnej generacji oczekiwane typy osobników będą odległe od siebie o około $2\sigma/\sqrt{\pi} \simeq 1,13\sigma$. Trajektorie wartości oczekiwanych prawie homogenicznych populacji są nieznacznie przesunięte w prawo od linii $w = \sqrt{2/\pi}\sigma$, czyli oczekiwane rozproszenie także wzrasta.

Wartość oczekiwana współrzędnej z (5.6) zależy od znaku współczynnika Ψ . Jeżeli przystosowanie osobnika z typem x_1 jest wyższe/nizsze niż przystosowanie osobnika z typem x_2 , to wartość oczekiwana współrzędnej z wzrasta/maleje. Dla populacji homogenicznych wartość oczekiwana współrzędnej z pozostaje stała ($w^i = 0$, $\Psi^i = 0$). Gdy populacja znajduje się w pobliżu prostej $w = \sqrt{2/\pi}\sigma$ i wartość w praktycznie się nie zmienia, wówczas na zmianę trajektorii wartości oczekiwanych wpływa współrzędna z . Dla małych wartości współrzędnej w funkcja Ψ ma niewielką, bliską zera wartość bezwzględną, więc krok z jakim trajektoria przesuwa się w kierunku osi Z w jednej iteracji jest mniejszy niż $\sqrt{2/\pi}\sigma$.

W lewej kolumnie rys. 5.6 pokazano wartości oczekiwane $E[w]$ i $E[z]$ w kolejnych 50 generacjach dla kilku wartości odchylenia standardowego mutacji σ ,

w krajobrazie jednomodalnej funkcji celu (D.2). Im większe σ , tym szybciej wartości oczekiwane obu współrzędnych są zbliżone do wartości stacjonarnej, która dla $E[w]$ jest zbliżona do wartości σ , a dla $E[z]$ wynosi zero. W prawej kolumnie rys. 5.6 przedstawiono przykładowe wartości w, z otrzymane z przebiegu rzeczywistego ewolucji na tle średnich z 50. i 1000. uruchomień procesu ewolucyjnego oraz oczekiwanych wartości stacjonarnych obu współrzędnych. Przebiegi te mogą uświadomić różnice pomiędzy aproksymacyjnym procesem deterministycznym opi-

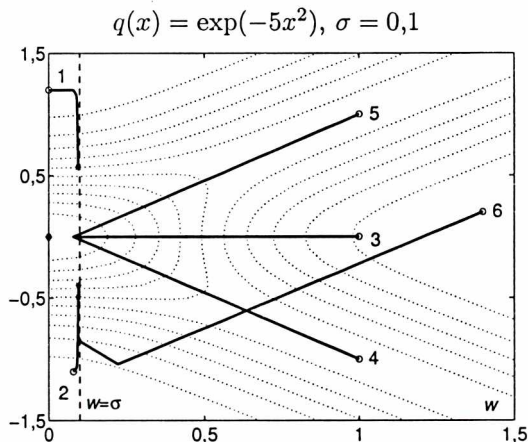


Rys. 5.6. Kolumna lewa: Wartości oczekiwane: a) $E[w]$, b) $E[z]$, dla różnych $\sigma = (0,1-0,5)$, w 50 generacjach. Kolumna prawa: Przykładowe wartości współrzędnej:

a) w , b) z , przebiegu rzeczywistego (linia ciągła), średnia z 50 (linia kropkowana)

i 1000 (linia przerywana) uruchomień. Funkcja celu: $q(x) = \exp(-5x^2)$, $\sigma = 0,3$

Fig. 5.6. Left column: Expected values: a) $E[w]$; b) $E[z]$ for various parameters $\sigma = (0.1-0.5)$ in 50 generations. Right column: Example values of coordinates: a) w , b) z , of actual evolutionary run (solid line), average from 50 runs (dotted line) and from 1000 runs (dashed line). Fitness function: $q(x) = \exp(-5x^2)$, $\sigma = 0.3$



Rys. 5.7. Trajektorie wartości oczekiwanych stanów populacji. Przedstawiono 20 generacji dla 6 stanów początkowych opisanych w tekście (oznaczonych przez o)
 Fig. 5.7. Trajectories of expected values of population states. 20 generations for 6 initial states described in text (and marked with o) are presented

sanym przez wartości oczekiwane a rzeczywistą realizacją procesu stochastycznego. Realizacje uśrednione przypominają wykresy wartości oczekiwanych tym bardziej, im więcej danych wykorzystano do ich obliczeń.

Przytoczona analiza wartości oczekiwanych wskazuje na istotną rolę prostej $w = \sqrt{2/\pi}\sigma$, która przyciąga trajektorie populacji znajdujących się w różnych stanach początkowych. Korzystając z terminologii układów dynamicznych, prostą tą można nazwać *atraktorem skoku* lub w terminologii biologicznej *kanalem ewolucyjnym*. Trajektorie wartości oczekiwanej dla populacji początkowej o dużej różnorodności (duża wartość w i duża różnica w jakości osobników q_1 i q_2) znajdzie się w kanale ewolucyjnym praktycznie już w pierwszej generacji. Taki skok oznacza znaczne zmniejszenie współrzędnej w , czyli zmniejszenie różnorodności populacji. Trajektorie wartości oczekiwanej dla populacji homogenicznych zostają natomiast odepchnięte od osi identyczności i przyciągnięte do atraktora skoku. Tak więc, niezależnie od początkowego stanu i po kilku zaledwie generacjach, trajektorie wartości oczekiwanej znajdują się w pobliżu atraktora. Nie tylko trajektorie wartości oczekiwanych, ale także populacje w przebiegach rzeczywistych będą przyciągane przez atraktor (zobacz rys. 5.3 i 5.4). Populacja w stanie znajdującym się w kanale ewolucyjnym jest skupiona, a typy osobników różnią się mniej więcej o odchylenie standardowe mutacji σ . Przyciągnięte przez atraktor populacje poruszają się następnie wolno wzdłuż kanału w kierunku optimum funkcji celu.

Przeanalizujemy obecnie wartości oczekiwane stanów populacji w kolejnej generacji dla kilku charakterystycznych stanów początkowych (rys. 5.7) oraz symetrycznej jednomodalnej funkcji celu $q(x) = \exp(-5x^2)$ (D.2) z optimum w zerze.

Stan 1. Trajektoria początkowo homogenicznej populacji zostaje odepchnięta od osi Z do atraktora skoku. Wartość oczekiwana współrzędnej w przyjmuje wielkość zgodną z wnioskiem 5.3.1.I, a wartość oczekiwana współrzędnej z pozostaje bez zmian.

Stan 2. Trajektoria populacji prawie homogenicznej, leżącej na linii $w = \sqrt{2/\pi}\sigma$, jest nieznacznie odepchnięta od tej linii, a oczekiwana różnorodność populacji wzrasta (wniosek 5.3.1.II): $E_{i+1}[w|\mathbf{s}^i] \simeq \sqrt{2/\pi}\sigma + 0,12(1 - \Psi^2)\sigma$. W zależności od znaku współczynnika Ψ , wartość oczekiwana współrzędnej z maleje lub rośnie przesuując populację w kierunku optimum $E_{i+1}[z|\mathbf{s}^i] = z^i + \Psi\sqrt{2/\pi}\sigma$.

Stan 3. Trajektorie populacji, których stan początkowy położony jest na osi W (czyli takie, których osobniki mają typy o przeciwnych znakach $x_1^i = -x_2^i$), pozostają na niej nadal. Wartość $E_{i+1}[z|\mathbf{s}^i] = 0$, ponieważ symetria funkcji $q(x) = q(-x)$ implikuje $\Psi = 0$. Zmniejsza się jedynie wartość oczekiwana współrzędnej w , która przyjmuje praktycznie połowę swojej początkowej wartości: $E_{i+1}[w|\mathbf{s}^i] = 1/2(\sqrt{2/\pi}\sigma + w^i)$. Trajektoria skacze do stanu na osi W położonego w połowie odległości od osi Z .

Stany 4 i 5. Trajektorie populacji o stanach leżących na osiach $W = Z$ lub $W = -Z$, gdzie jeden z osobników ma optymalny typ równy 0. Wartość oczekiwana współrzędnej w maleje, a wartość oczekiwana współrzędnej z dąży do zera (maleje lub rośnie). Trajektoria skacze do punktu przecięcia atraktora skoku z osią W .

Stan 6. Trajektoria populacji początkowej o dużej różnorodności skacze w pobliże atraktora skoku. Po dwóch generacjach wartość oczekiwana $E_{i+1}[w|\mathbf{s}^i] = \sqrt{2/\pi}\sigma$, a wartość oczekiwana współrzędnej z maleje $E_{i+1}[z|\mathbf{s}^i] = z^i - w^i$.

Analiza ewolucji populacji w przestrzeni stanów S_U umożliwia także obliczenie wariancji współrzędnych w i z :

$$D_{i+1}^2[w|\mathbf{s}^i] = \sigma^2 + (1 - (\Psi^i)^2) \frac{w^{i2}}{2} - (E_{i+1}[w|\mathbf{s}^i])^2, \quad (5.13)$$

$$D_{i+1}^2[z|\mathbf{s}^i] = \sigma^2 + (1 - (\Psi^i)^2) \frac{w^{i2}}{2}. \quad (5.14)$$

Wariancja współrzędnej w jest mniejsza od wariancji współrzędnej z o składnik $(E_{i+1}[w|\mathbf{s}^i])^2$ (patrz rys. 5.6). Gdy różnice w wartościach przystosowania osobników są znaczne, czyli $1 - (\Psi^i)^2 = 0$, to $D_{i+1}^2[w|\mathbf{s}^i] = (1 - 2/\pi)\sigma^2 \simeq 0,36\sigma^2$ i $D_{i+1}^2[z|\mathbf{s}^i] = \sigma^2$. Jeżeli rozproszenie populacji zmniejsza się, wartości wariancji (5.13) i (5.14) wzrastają. Korzystając z aproksymacji (5.7) można określić górne ograniczenie na wartości wariancji (5.13):

$$D_{i+1}^2[w|\mathbf{s}^i] \leq (1 - 1/(2\pi))\sigma^2 + w^{i2}/2 \simeq 0,84\sigma^2 + \frac{w^{i2}}{2}.$$

Górne ograniczenie wartości wariancji współrzędnej z (5.14) zależy od wartości σ oraz w i wynosi

$$D_{i+1}^2[z|\mathbf{s}^i] < \sigma^2 + \frac{w^{i2}}{2}.$$

Podsumowując, z przeprowadzonej analizy wynika, że ewolucja populacji dwuelementowej jest procesem o dwóch prędkościach. Krótkie okresy adaptacji prowadzą populację do obszarów o wyższym średnim przystosowaniu i redukują jej zróżnicowanie. Skupione populacje następnie wolno dryfują w kierunku pobliskich optimów funkcji celu.

5.4. Wartości oczekiwane stanu populacji jako dyskretny układ dynamiczny

Wartości oczekiwane stanów populacji opisane równaniami (5.5), (5.6) generują dyskretny układ dynamiczny określony w przestrzeni S_U :

$$\begin{cases} w^{i+1} = E_{i+1}[w|\mathbf{s}^i], \\ z^{i+1} = E_{i+1}[z|\mathbf{s}^i]. \end{cases} \quad (5.15)$$

Równania te definiują odwzorowanie

$$(w, z) \rightarrow F(w, z) = \begin{cases} F_1(w, z), \\ F_2(w, z). \end{cases} \quad (5.16)$$

Otrzymano w pełni deterministyczny model fenotypowego procesu ewolucyjnego dla populacji dwuelementowych. Układ (5.15) można wykorzystać do analizy aproksymacji średniego zachowania populacji w czasie. W stanie równowagi selekcyjno-mutacyjnej populacja jest położona w punkcie stałym, którego współrzędne oblicza się z układu równań:

$$w = F_1(w, z), \quad (5.17)$$

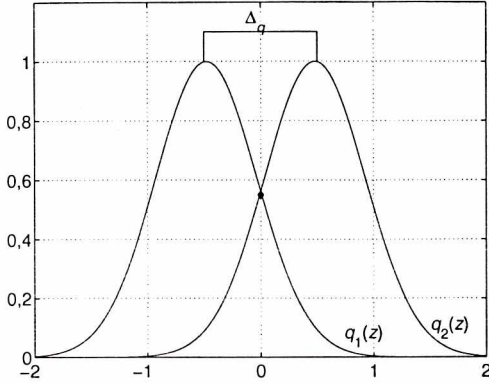
$$z = F_2(w, z). \quad (5.18)$$

Jeżeli uwzględnimy (5.5), równanie (5.17) jest równaniem przestępnym, którego przybliżone rozwiązanie wyznacza się numerycznie. Równanie (5.18) jest spełnione dla $w^i = 0$ lub $\Psi = 0$. Ponieważ $w^i = 0$ nie jest nigdy pierwiastkiem równania (5.17), zatem stan równowagi $s^s = (w^s, z^s)$ jest charakteryzowany warunkami

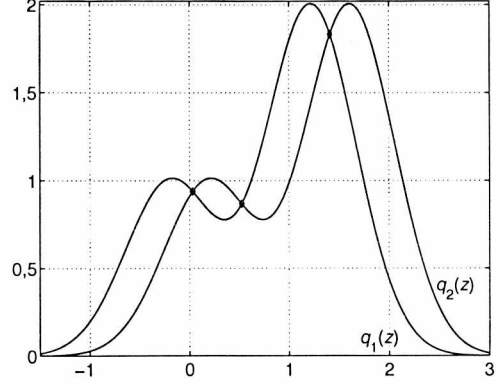
$$w^s \simeq 0,97\sigma, \quad (5.19)$$

$$\Psi(w^s, z^s) = 0. \quad (5.20)$$

a) $q(x) = \exp(-5x^2)$



b) $q(x) = \exp(-5x^2) + 2 \exp[-5(x-1)^2]$



Rys. 5.8. Współrzędne z punktów stałych otrzymane jako punkty przecięcia dwóch przesuniętych gaussowskich funkcji celu $q_1(z)$ i $q_2(z)$, $\sigma = 0,25$

Fig. 5.8. The z -coordinates of fixed points obtained as intersection points of two shifted Gaussian functions $q_1(z)$ and $q_2(z)$, $\sigma = 0.25$

Współrzędna w^s punktu stałego (5.19) nie zależy od funkcji przystosowania, natomiast współrzędna z^s (5.20) jest od niej zależna i jest rozwiązaniem równania

$$q((z^s + w^s)/\sqrt{2}) = q((z^s - w^s)/\sqrt{2}). \quad (5.21)$$

Rozwiązanie równania (5.21) otrzymuje się jako argument punktu przecięcia dwóch funkcji $\tilde{q}_1(z) = q((z + w^s)/\sqrt{2})$ i $\tilde{q}_2(z) = q((z - w^s)/\sqrt{2})$, przesuniętych względem siebie o $\Delta_q = \sqrt{2}w^s \simeq 1,37\sigma$ (rys. 5.8). Dla uproszczenia zapisu przyjmujemy $\tilde{q}_j(z) = q_j(z)$, $j = 1, 2$, identyfikując instancję funkcji q_j po liczbie argumentów. Liczba punktów przecięcia, czyli punktów stałych, zależy od liczby optimów funkcji przystosowania oraz przesunięcia Δ_q , które jest funkcją odchylenia standardowego mutacji σ . Punkty stałe położone są w pobliżu optimów i siodła funkcji celu. Gdy wartość σ wzrasta, liczba punktów stałych maleje, ponieważ przesunięcie Δ_q między funkcjami $q_1(z)$ i $q_2(z)$ rośnie.

W celu znalezienia warunków stabilności punktów stałych układu dynamicznego (5.16) obliczono jacobian zwartościowany dla współrzędnych punktu równowagi:

$$\begin{bmatrix} \frac{\partial F_1(w, z)}{\partial w} & \frac{\partial F_1(w, z)}{\partial z} \\ \frac{\partial F_2(w, z)}{\partial w} & \frac{\partial F_2(w, z)}{\partial z} \end{bmatrix}_{w=w^s, z=z^s} = \begin{bmatrix} \Phi_0(w) & 0 \\ w \frac{\partial \Psi(w, z)}{\partial w} & w \frac{\partial \Psi(w, z)}{\partial z} + 1 \end{bmatrix}_{w=w^s, z=z^s} \quad (5.22)$$

Macierz (5.22) jest trójkątna, więc jej wartości własne są równe elementom na przekątnej: $\lambda_1 = \Phi_0(w^s)$ oraz $\lambda_2 = 1 + w^s (\partial \Psi(w, z)/\partial z)_{w=w^s, z=z^s}$. Punkt stały

jest stabilny, jeśli wartość bezwzględna wartości własnych jest nie większa niż 1. Ponieważ $w^s > 0$, wartość bezwzględna pierwszej wartości własnej λ_1 jest mniejsza od jedności. Druga wartość własna zależy od funkcji przystosowania. Aby punkt stały był stabilny, musi być spełniona następująca nierówność:

$$-2 \leq w^s \frac{\partial \Psi(w^s, z)}{\partial z} \Big|_{z=z^s} \leq 0. \quad (5.23)$$

Ponieważ w punkcie stałym spełnione jest równanie (5.21), nierówność (5.23) można przedstawić w rozwiniętej postaci:

$$-2 \leq w^s \frac{\partial q_1(z)/\partial z - \partial q_2(z)/\partial z}{q_1(z) + q_2(z)} \Big|_{z=z^s} < 0. \quad (5.24)$$

Dla nieujemnych funkcji celu, prawa nierówność z (5.24) jest spełniona, gdy

$$\frac{\partial q_1(z)}{\partial z} \Big|_{z=z^s} < \frac{\partial q_2(z)}{\partial z} \Big|_{z=z^s}. \quad (5.25)$$

Dla punktów stałych leżących w pobliżu optimum funkcji celu nierówność (5.25) jest spełniona i punkty są stabilne dopóty, dopóki lewa nierówność (5.24) jest spełniona. Z nierówności (5.25) wynika następujące twierdzenie.

Twierdzenie 5.4.1. *Punkty stałe leżące w pobliżu siodła są niestabilne.*

Dowód: W punktach przecięcia się przesuniętych funkcji przystosowania $q_1(z)$ oraz $q_2(z)$ nachylenia ich zboczy są przeciwne (styczne mają współczynniki kierunkowe o różnych znakach). Nierówność (5.25) jest spełniona, jeśli nachylenie funkcji $q_1(z)$ jest malejące ($\partial q_1/\partial z < 0$), a nachylenie funkcji $q_2(z)$ jest rosnące ($\partial q_2/\partial z < 0$). W punkcie siodłowym nachylenie zbocza funkcji $q_1(z)$ wzrasta ($\partial q_1/\partial z > 0$), a nachylenie zbocza funkcji $q_2(z)$ maleje ($\partial q_2/\partial z < 0$). Tak więc, punkt stały zlokalizowany przy siodle jest niestabilny. $\square$

5.5. Analiza punktów stałych i ich stabilności

Dalsze rozważania ogólne układu dynamicznego (5.16) są utrudnione i mogą być przeprowadzone analitycznie dla szczególnych klas funkcji przystosowania. W kolejnych podrozdziałach przedstawimy dokładną analizę punktów stałych układu opisującego dynamikę wartości oczekiwanych zmiennych w i z , ich obszarów przyciągania oraz stabilności dla wybranych funkcji celu.

5.5.1. Punkty stałe

Położenie punktów stałych (a dokładniej współrzędnej z^s) oraz ich stabilność zależą od funkcji celu oraz od wartości odchylenia standardowego mutacji σ . W tym

podrozdziale omówiono zachowanie asymptotyczne układu dynamicznego (5.16) dla kilku charakterystycznych funkcji celu. Rozpatrzono dwie klasy funkcji: jednomodalne i dwumodalne. Oprócz wykorzystywanych w rozdziale 4 jedno- i dwumodalnych funkcji gaussowskich (D.2), (D.4) oraz dwumodalnej funkcji trójkątnej (D.6) rozważano również jednomodalną funkcję trójkątną (D.5). Dodatkowo obie klasy funkcji badano w wersji symetrycznej i asymetrycznej (funkcje (D.1) oraz (D.3)).

Jednomodalne funkcje celu

Jednomodalna funkcja trójkątna (D.5) oraz gaussowska (D.2) mają optimum lokalne w zerze. W wersji symetrycznej oba ramiona funkcji trójkątnej są równe ($A=B$) i nachylenia obu zboczy funkcji gaussowskiej takie same. Funkcje jednomodalne mają nie więcej niż jeden punkt stały, ponieważ funkcje $q_1(z)$ i $q_2(z)$ mogą przeciąć się tylko w jednym punkcie (rys. 5.8a) lub wcale (w przypadku gdy funkcje są zdefiniowane na ograniczonym przedziale jak (D.5)).

Przypadek symetryczny. Dla funkcji symetrycznych, gdy $q(x) = q(-x)$, rozwiązaniem równania (5.21) jest $z^s = 0$. Tak więc punkt $\omega^s = (0,97\sigma; 0)$ jest jedynym punktem stałym dla jednomodalnych symetrycznych funkcji przystosowania. Populacja znajdująca się w stanie odpowiadającym punktowi stałemu składa się z dwóch elementów o jednakowej jakości, lecz położonych po przeciwnych stronach zboczy funkcji celu w odległości około $0,69\sigma$ od optimum ($x_1^s \simeq 0,69\sigma$ oraz $x_2^s \simeq -0,69\sigma$). Proces ewolucji wartości oczekiwanych populacji dwuelementowej jest więc zbieżny nie do optimum funkcji przystosowania, lecz do stanu w położonego w pewnej odległości od optimum. Na wielkość oddalenia od optimum ma wpływ mutacja. Im mniejsza wartość odchylenia standardowego mutacji, tym lokalizacja optimum jest dokładniejsza. W skrajnym przypadku, gdy $\sigma = 0$, $w^s = 0$ i punkt równowagi będzie leżał w optimum.

Przypadek asymetryczny. Asymetria funkcji celu wpływa na wartość współrzędnej z punktu stałego. Przesuwa się ona z zera w stronę segmentu funkcji o większej „masie”. Współrzędne z punktu stałego wyznaczamy, korzystając z równania (5.21). Dla asymetrycznej funkcji trójkątnej (D.5) (rys. 5.9a)

$$z^s = \frac{B - A}{B + A} w^s. \quad (5.26)$$

Współrzędne z^s są określone tylko dla $w^s \leq (B + A)/\sqrt{2}$, gdyż dla większych wartości w^s wykresy $q_1(z)$ i $q_2(z)$ się nie przecinają. Dla asymetrycznych funkcji gaussowskich (D.3) współrzędną z^s opisuje równanie (rys. 5.9b)

$$z^s = w^s \frac{a_1 + a_2 - 2\sqrt{a_1 a_2}}{a_1 - a_2}. \quad (5.27)$$

Im większa wartość odchylenia standardowego mutacji, tym bardziej punkt stały oddala się od optimum funkcji celu.

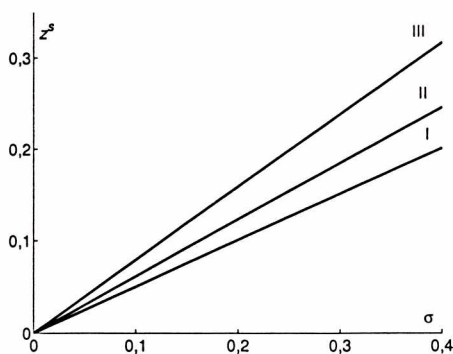
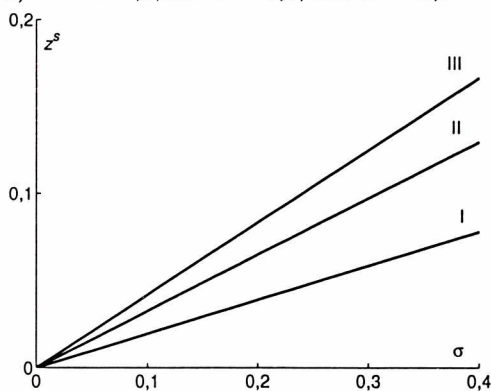
Dwumodalne funkcje celu

W przypadku funkcji dwumodalnych, układ dynamiczny (5.16) ma jeden lub trzy punkty stałe. Zwykle punkty te znajdują się w pobliżu obu optimum i siodła.

Przypadek symetryczny. Jako przykład symetrycznej funkcji dwumodalnej rozpatrzmy sumę dwóch jednomodalnych funkcji gaussowskich o jednakowej wysokości, której oś symetrii leży w zerze: $q(x) = \exp[-5(x+0,5)^2] + \exp[-5(x-0,5)^2]$. Układ dynamiczny (5.16) ma dwa symetryczne punkty stałe w pobliżu optimum i jeden siodłowy (rys. 5.11a). Funkcje $q_1(z)$ i $q_2(z)$ przecinają się w więcej niż jednym punkcie, jeżeli przesunięcie Δ_q pomiędzy nimi jest mniejsze niż odległość pomiędzy optimum wynosząca d czyli $\sqrt{2}w^s < d$. Wraz ze wzrostem odchylenia standardowego mutacji punkty stałe związane z optimummi (w^s, z^s) i $(w^s, -z^s)$, zbliżają się do siebie, aż staną się jednym punktem. Gdy odchylenie standardowe mutacji jest większe od wartości $\sigma \simeq 0,73d$, punkt siodłowy $w^s = (0,97\sigma; 0)$ jest jedynym punktem stałym. Dla tak dużych wartości σ , większych niż odległość pomiędzy wzgórzami, pojedyncza mutacja może łatwo spowodować przeskok populacji do obszaru przyciągania drugiego optimum i populacja nie może się ustabilizować na żadnym z nich. Im mniejsza jest wartość odchylenia standardowego mutacji, tym punkty stałe dokładniej lokalizują optima.

a) I. $B = 1,5$; II. $B = 2,0$; III. $B = 2,5$

b) I. $a_2 = 0,1$; II. $a_2 = 0,05$; III. $a_2 = 0,01$



Rys. 5.9. Współrzędne z^s punktów stałych jako funkcja σ . Jednomodalne funkcje celu: a) trójkątne (D.5), $A = 1$; b) gaussowskie (D.3), $a_1 = 1,0$

Fig. 5.9. The z^s -coordinates of fixed points as the function of σ . Unimodal fitness functions: a) tent (D.5), $A = 1$; b) Gaussian (D.3), $a_1 = 1.0$

Przypadek asymetryczny. Podobnie jak dla funkcji symetrycznych punkty stałe położone są w pobliżu optimów i siodła. W przypadku dwumodalnej funkcji trójkątnej (D.6), funkcje $q_1(z)$ i $q_2(z)$ mogą przecinać się w jednym, dwóch³ lub trzech punktach. Współrzędne z punktów stałych otrzymane z różnych kombinacji przecinających się ramion funkcji (D.6) wraz z warunkami ich wyboru zebrano w tabeli 5.1, a przykłady przecięć funkcji $q_1(z)$ i $q_2(z)$ przedstawiono na rys. 5.10. Wraz ze wzrostem odchylenia standardowego mutacji punkty stałe leżące w pobliżu optimum lokalnego i siodła zbliżają się do siebie i dla $w^s = (A + B)/\sqrt{2}$ znikają. Pozostały, globalny punkt stały przesuwają się w kierunku siodła, a dla $w^s > (A + C)/\sqrt{2}$ także zanika ($q_1(z)$ i $q_2(z)$ się nie przecinają) (rys. 5.11c).

Dla dwumodalnej funkcji gaussowskiej ze wzgórzami o różnej wysokości: $q(x) = \exp(-5x^2) + h \exp[-5(x-1)^2]$ (D.4), gdzie $h > 1$, punkty stałe można wyznaczyć numerycznie, korzystając z równania (5.21). Punkty stałe leżące w pobliżu optimum lokalnego i siodła zbliżają się do siebie wraz ze wzrostem σ , a następnie zanikają, pozostawiając tylko punkt stały związany z optimum globalnym (rys. 5.11b).

Wielomodalne funkcje celu

Dla funkcji wielomodalnych punktów stałych może być więcej, jeśli odległość między optimami funkcji celu jest większa od przesunięcia Δ_q . Dla funkcji przystosowania o trzech optimach punktów stałych jest 1, 3 lub 5, dla funkcji o czterech optimach 1, 3, 5 lub 7 itd. Ogólnie, dla funkcji o k optimach układ dynamiczny ma do $2k + 1$ punktów stałych zlokalizowanych w pobliżu optimów oraz siodeł funkcji przystosowania. W przypadku funkcji symetrycznych jeden z punktów jest zawsze położony na osi symetrii, podczas gdy inne tworzą symetryczne pary. Wraz ze wzrostem wartości σ punkty stałe w pobliżu niektórych optimów mogą zanikać.

5.5.2. Obszary przyciągania punktów stałych

Jeżeli układ dynamiczny ma więcej niż jeden punkt stały, to każdy z nich ma swój *obszar przyciągania* (zwany także zbiorem atrakcji), czyli zbiór stanów początkowych, dla których trajektorie dążą do odpowiednich punktów stałych [156, 192]. Formalnie, *obszarem przyciągania* $B(\mathbf{x}^s)$ danego punktu stałego $\mathbf{x}^s$ nazywamy zbiór stanów początkowych modelu procesu ewolucyjnego opisanego funkcją F , z których trajektorie są zbieżne do $\mathbf{x}^s$:

$$B(\mathbf{x}^s) = \{\mathbf{x} | F^i(\mathbf{x}) \xrightarrow{i \rightarrow \infty} \mathbf{x}^s, F(\mathbf{x}^s) = \mathbf{x}^s\}.$$

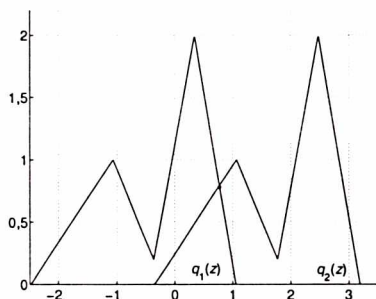
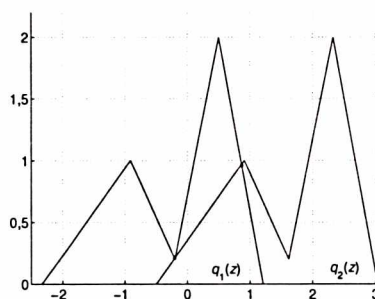
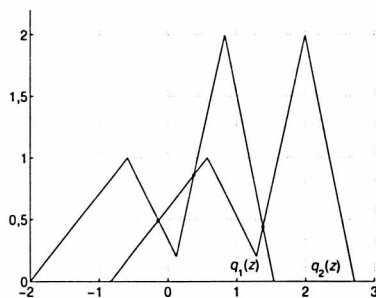
Rozpatrzmy obszary przyciągania dla wielomodalnych funkcji przystosowania, dla których układ dynamiczny (5.16) ma więcej niż jeden punkt stały. Na rysunku 5.12 (wiersz górny) przedstawiono obszary przyciągania sumy funkcji

³ Przypadek szczególny.

Tabela 5.1. Równania współrzędnej z^s punktów stałych dla kombinacji przecinających się ramion funkcji $q_1(z)$ i $q_2(z)$. Kolejne ramiona opisano: I = $[-A, 0)$, II = $[0, B)$, III = $[B, 1)$, IV = $[1, D]$ (rys. D.1d), a indeksy dolne determinują funkcję q_1 lub q_2

Table 5.1. Equations of the z^s -coordinates of fixed points for the combinations of intersecting branches of the functions $q_1(z)$ and $q_2(z)$. Branches are described as follows: I = $[-A, 0)$, II = $[0, B)$, III = $[B, 1)$, IV = $[1, D]$ (Fig. D.1d); subscripts determine the function q_1 or q_2

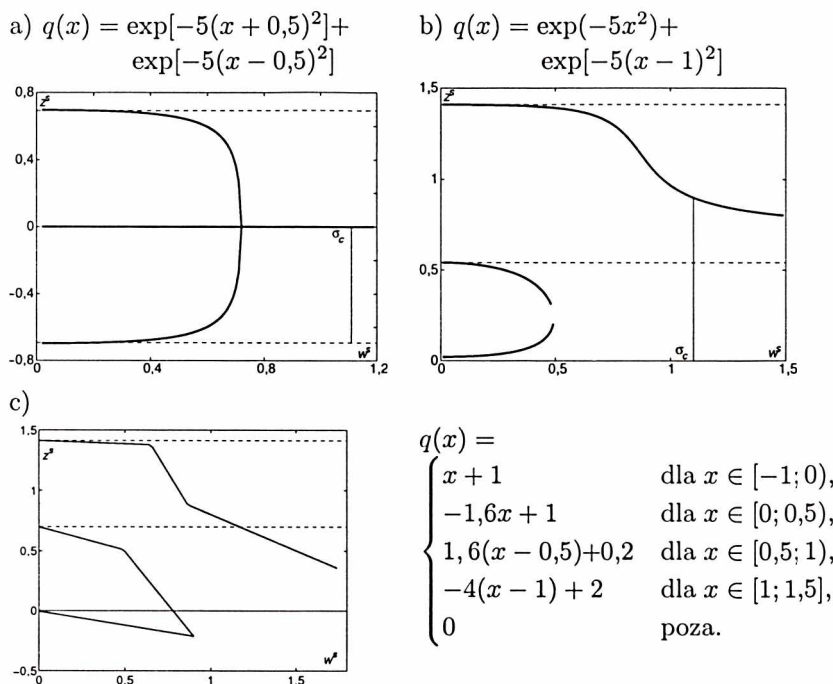
Ramiona	Równanie współrzędnej z^s
II ₁ , I ₂	$z^s = w^s(B - A + Ah_1)/(B + A - Ah_1)$
III ₁ , I ₂	$z^s = \frac{w^s(B - A(h - h_1) - 1) - \sqrt{2A(B - Bh + h_1 - 1)}}{B + A(h - h_1) - 1}$
III ₁ , II ₂	$z^s = w^s[B(2h_1 - h - 1) - h_1 + 1]/(B - Bh + h_1 - 1) + \sqrt{2B}$
IV ₁ , I ₂	$z^s = [w^s(D - Ah - 1) - \sqrt{2A(D - Dh - 1)}]/(D + Ah - 1)$
IV ₁ , II ₂	$z^s = \frac{w^s[-Bh + D(h_1 - 1) - h_1 + 1] - \sqrt{2B(D - Dh - 1)}}{Bh + D(h_1 - 1) - h_1 + 1}$
IV ₁ , III ₂	$z^s = w^s[D(h_1 - h) - Bh - h_1 + 2h]/(B - Bh + h_1 - 1) + \sqrt{2}$

a) $\sigma = 1,1$, $\Delta_q = 1,5$ b) $\sigma = 0,94$, $\Delta_q = 1,3$ c) $\sigma = 0,6$, $\Delta_q = 0,8$ funkcja celu: $q(x) =$

$$\begin{cases} x + 1 & \text{dla } x \in [-1; 0), \\ -1,6x + 1 & \text{dla } x \in [0; 0,5), \\ 1,6(x - 0,5) + 0,2 & \text{dla } x \in [0,5; 1), \\ -4(x - 1) + 2 & \text{dla } x \in [1; 1,5], \\ 0 & \text{poza.} \end{cases}$$

Rys. 5.10. Przecięcia funkcji trójkątnych $q_1(z)$ i $q_2(z)$ w jednym, dwóch i trzech punktach dla różnych wartości σ

Fig. 5.10. Intersections of bimodal tent functions $q_1(z)$ and $q_2(z)$ in one, two and three points for different values of the parameter σ



Rys. 5.11. Położenie punktów stałych układu (5.16) dla różnych funkcji celu
 Fig. 5.11. Locations of fixed points of dynamical system (5.16) for different fitness functions

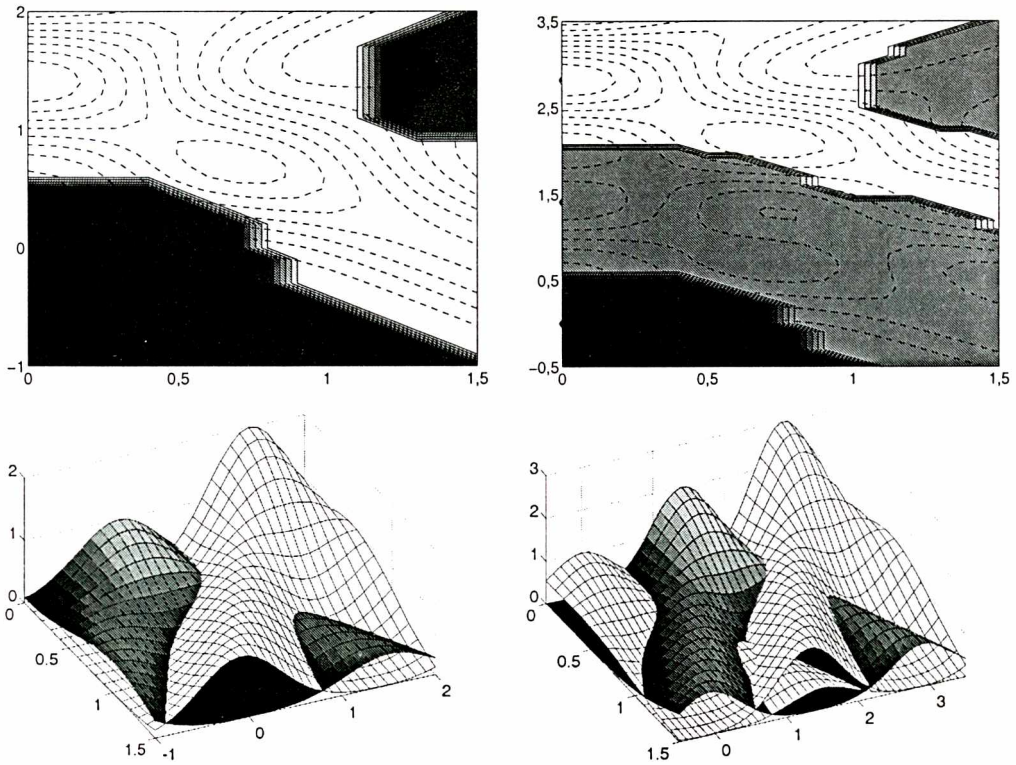
gaussowskich o dwóch i trzech optimach (D.1). Na wszystkich rysunkach obszarów przyciągania kolorem białym oznaczono obszar przyciągania optimum globalnego, kolorem czarnym optimum lokalnego, a szarym optimum pośredniego dla funkcji o trzech optimach. W obu przypadkach zauważamy, że istnieją pewne stany przyciągane przez optima lokalne, pomimo iż są one położone bliżej optimum globalnego (np. w prawym górnym rogu rys. 5.12a). Wyjaśnienia tego faktu poszukamy na rys. 5.12 (wiersz dolny), na którym przedstawiono osobno każdą składową z sumy funkcji gaussowskich w krajobrazie średniego przystosowania. Zdefiniujemy *obszar wpływu* danego optimum jako zbiór stanów, w których wkład danej funkcji składowej do średniego przystosowania jest wyższy niż wkład pozostałych składowych. Dla funkcji dwumodalnej (rys. 5.12a) obszar wpływu optimum lokalnego jest zbiorem stanów $s = (x_1, x_2)$, dla których

$$h_1 \exp(-a(x_1 - d_1)^2) + h_1 \exp(-a(x_2 - d_1)^2) > h_2 \exp(-a(x_1 - d_2)^2) + h_2 \exp(-a(x_2 - d_2)^2),$$

gdzie $x_1 = (z + w)/\sqrt{2}$, $x_2 = (z - w)/\sqrt{2}$. Ponieważ obszar wpływu optimum lokalnego obejmuje także stany w prawym górnym rogu rys. 5.12a,

a) $q(x) = \exp(-5x^2) + 2 \exp[-5(x-1)^2]$

b) $q(x) = \exp(-5x^2) + 2 \exp[-5(x-1)^2] + 3 \exp[-5(x-2)^2]$



Rys. 5.12. Obszary przyciągania (górny wiersz) i obszary wpływu optimów składowych (wiersz dolny) dla wielomodalnych funkcji celu, $\sigma = 0,1$

Fig. 5.12. Basins of attraction (the upper pictures) and impact areas of optima (the lower pictures) for multimodal fitness functions, $\sigma = 0.1$

więc trajektorie zainicjowane w tych stanach będą przyciągane przez optimum lokalne, co można także zaobserwować na rys. 5.4a. W przestrzeni ilorazowej zakłócone więc są relacje sąsiedztwa. Przykładowe trajektorie ewolucji i stanów oczekiwanych przedstawione na rys. 5.4 zainicjowano w stanach należących do obszarów przyciągania różnych optimów. Populacje ze stanów początkowych pokazanych na rys. 5.4a, b są przyciągane przez optimum lokalne, a te ze stanów na rys. 5.4c, d przez optimum globalne. Pomimo iż stany początkowe należą do obszarów przyciągania danych optimów, ewolucja rozpoczynająca się w nich może doprowadzić populację zarówno w pobliże jednego, jak i drugiego wzgórza (rys. 5.4b, d). Dzieje się tak, gdy rozkład (5.3) dla danego stanu początkowego jest wielomodalny

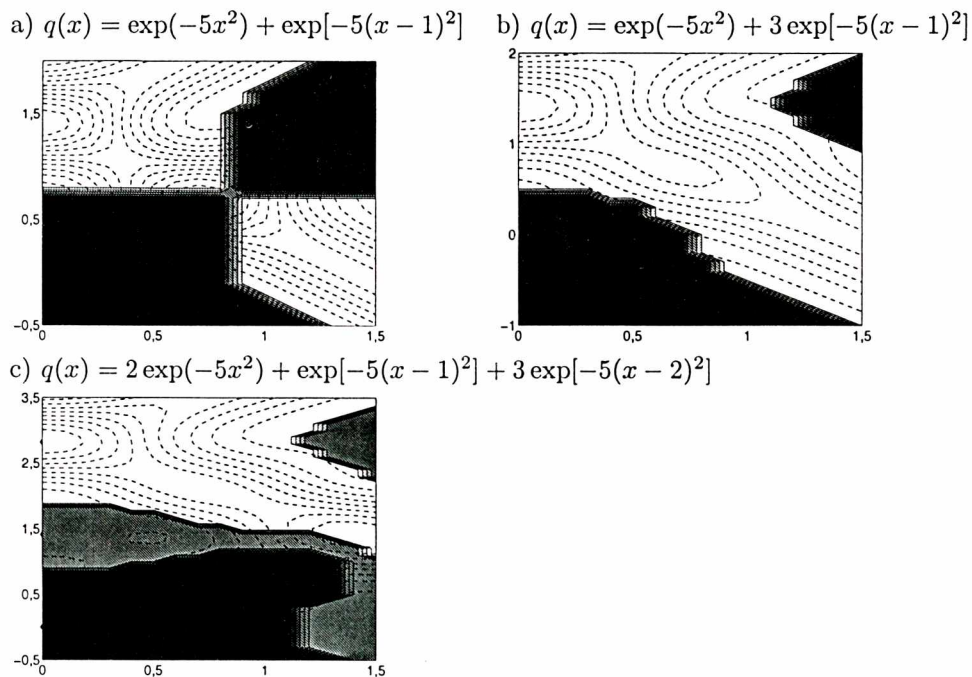
z optimami o podobnej wysokości. Natomiast trajektorie stanów oczekiwanych są w tym przypadku zdeterminowane przez obszary przyciągania. Opisane zjawisko wskazuje na różnicę między zachowaniem stochastycznym ewoluującej populacji a zachowaniem deterministycznym jej wartości oczekiwanych.

Kształt i wielkość zbiorów przyciągania zależy od postaci funkcji celu. Zbadamy wpływ parametrów wielomodalnych funkcji gaussowskich: nachylenia zbocza i różnic wysokości pomiędzy wzgórzami, na kształty obszarów przyciągania. W przypadku symetrycznych funkcji dwumodalnych o wzgórzach równej wysokości, można by się spodziewać, że zbiory przyciągania będą równej wielkości (miary) i położone w pobliżu danego optimum. W rzeczywistości obszary te mają bardziej skomplikowaną strukturę i są położone wzdłuż grzbietów zawierających stany, w których jeden osobnik jest optymalny (rys. 5.13a). Potwierdza to wcześniejsze obserwacje, że populacje przyciągane są głównie przez grzbiety funkcji przystosowania, a ewolucja przebiega wzdłuż grzbietów [64]. Różnice wysokości wzgórz wpływają na wielkości zbiorów przyciągania, powodując wzrost wielkości obszaru należącego do najwyższego wzgórza (rys. 5.13b, c). Gdy różnica wysokości wzgórz maleje, obszar przyciągania niższego wzgórza rozszerza się. Nachylenie zboczy funkcji celu ma wpływ na wielkość obszarów przyciągania, a nie na ich kształt (rys. 5.14). Zarówno gdy zbocza obu wzgórz mają bardziej strome nachylenie (parametr a rośnie) (rys. 5.14a) lub gdy wzgórze globalne jest bardzo wąskie (rys. 5.13b), powiększa się obszar przyciągania optimum lokalnego, pozostawiając w zbiorze przyciągania optimum globalnego stany przyległe do niego oraz wąski obszar prowadzący wzdłuż grzbietu obszaru wpływu tego optimum. Obszary przyciągania punktów stałych zależą także od wielkości odchylenia standardowego mutacji. Jeżeli wartość σ wzrasta, to kształt zbiorów przyciągania może się zmienić, ponieważ niektóre punkty stałe, związane z optimami lokalnymi, zanikają.

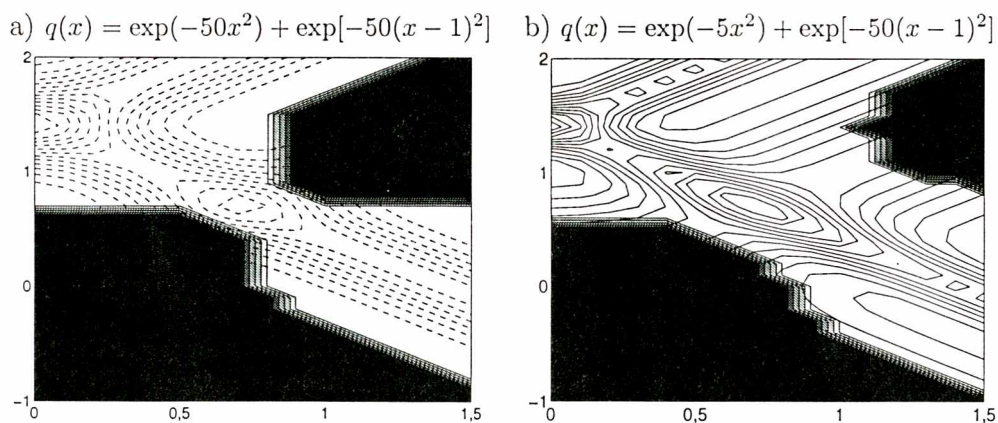
5.5.3. Stabilność punktów stałych

Stabilność punktu stałego jest określona przez drugą wartość własną λ_2 macierzy Jacobiego i, zgodnie z nierównością (5.23), zależy od postaci funkcji celu oraz odchylenia standardowego mutacji. Dla analizowanych funkcji przystosowania punkty stałe są stabilne dla małych wartości σ i stają się niestabilne dla większych σ . Zdefiniowano zatem *wartość krytyczną odchylenia standardowego mutacji* σ_c , dla której punkt stały traci stabilność. Zjawisko przejścia fazowego dla pewnych wartości odchylenia standardowego mutacji tzw. *error threshold*⁴ jest znane zarówno wśród badaczy układów dynamicznych [155, 212] jak i biologów ewolucyjnych [47, 49]. Dla układu dynamicznego (5.15) wartość σ_c jest wyznaczana z wa-

⁴ *Error threshold* to krytyczna wartość mutacji, powyżej której struktura populacji załamuje się, a populacja rozprasza w przestrzeni poszukiwań [47].



Rys. 5.13. Wpływ wysokości optimów funkcji celu na obszary przyciągania, $\sigma = 0,1$
 Fig. 5.13. Basins of attraction as a function of the optimum height, $\sigma = 0.1$



Rys. 5.14. Wpływ nachylenia zboczy funkcji celu na obszary przyciągania, $\sigma = 0,1$
 Fig. 5.14. Basins of attraction as a function of the hillside slope, $\sigma = 0.1$

runku $\lambda_2 = -1$ lub $\lambda_2 = 1$. W dalszych obliczeniach zamiast wprost z wartości σ_c wygodniej jest korzystać z wartości krytycznej współrzędnej w , ponieważ $w_c \simeq 0,97\sigma_c$, więc $\sigma_c \simeq 1,03w_c$. Dalej przeanalizujemy stabilność punktów stałych układu dynamicznego (5.16) dla przyjętych funkcji celu.

Jednomodalne funkcje celu

Stabilność jedyne punktu stałego dla funkcji jednomodalnych zależy od odchylenia standardowego mutacji oraz parametrów nachylenia zbocza a (dla funkcji gaussowskiej (D.2)) lub A (dla funkcji trójkątnej (D.5)). Przykładowe trajektorie wartości oczekiwanych z dwóch stanów początkowych dla różnych wartości σ przedstawiono na rys. 5.15.

Przypadek symetryczny. Dla funkcji symetrycznej trójkątnej (D.5) λ_2 wynosi

$$\lambda_2 = 1 - w^s / (A - w^s). \quad (5.28)$$

Druga wartość własna przyjmuje więc wartości mniejsze od jedności (rys. 5.16a), a układ będzie niestabilny, jeśli λ_2 zmaleje poniżej wartości -1 , czyli gdy spełniona będzie nierówność $2 > w^s / (A - w^s)$. Wartość krytyczna współrzędnej w wynosi więc $w_c = 2/3A$.

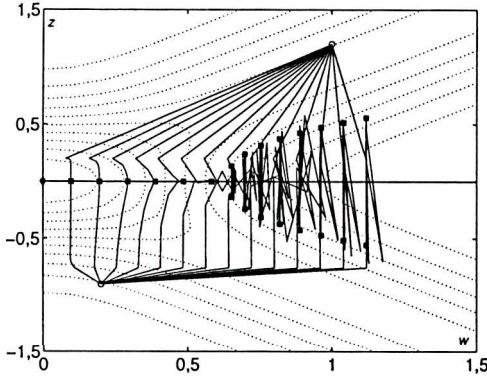
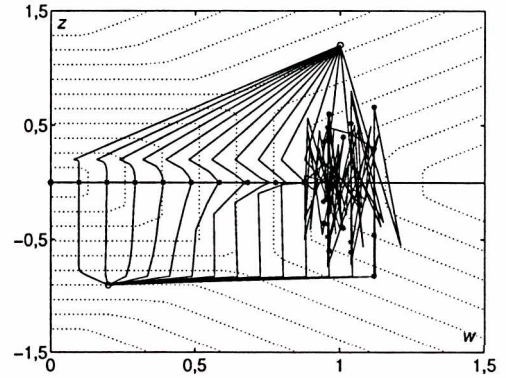
Dla symetrycznej funkcji gaussowskiej wartość λ_2 jest obliczana ze wzoru

$$\lambda_2 = 1 - a(w^s)^2. \quad (5.29)$$

Nierówność (5.23) przyjmuje postać $-2 < -a(w^s)^2 < 0$. Prawa nierówność jest zawsze spełniona, a z rozwiązania lewej otrzymamy wartość krytyczną, która wynosi $w_c = \sqrt{2/a}$ (rys. 5.16b).

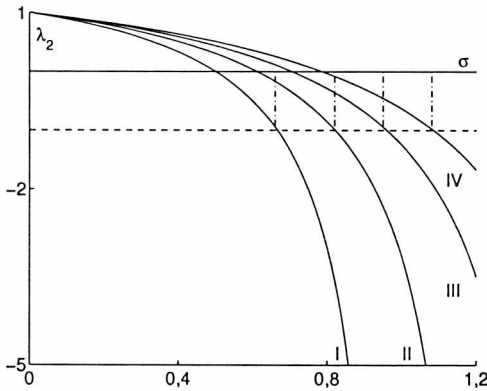
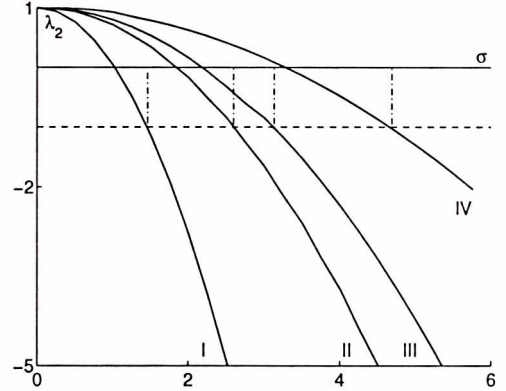
Dla krytycznych wartości odchylenia standardowego mutacji σ_c układ dynamiczny (5.16) traci stabilność i występuje bifurkacja. Układ staje się niestabilny dla $\sigma > \sigma_c$ i pojawia się stabilna orbita periodyczna o okresie 2 (rys. 5.17a). Mutacje o bardzo dużym zasięgu mogą więc spowodować, że potomkowie optymalnego rodzica znajdą się w obszarach o bardzo małym przystosowaniu. Różnice w jakości takich potomków są bardzo małe, presja selekcyjna także, największy wpływ na populację ma zatem mutacja (z dużym σ), co objawia się zachowaniami niestabilnymi.

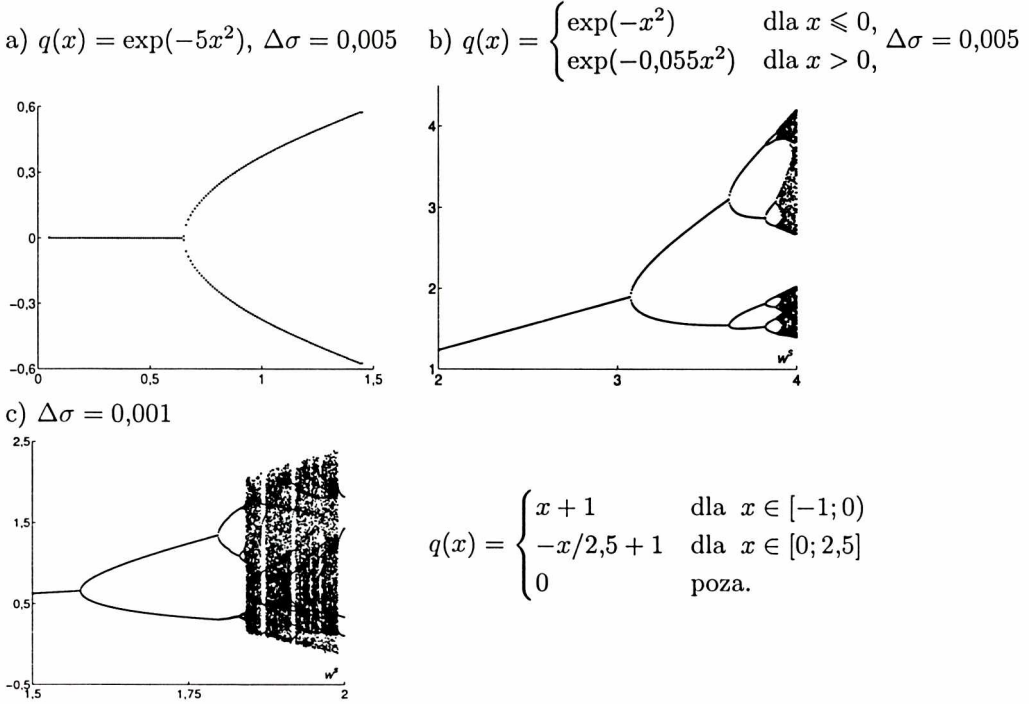
Na rysunku 5.18 przedstawiono graficzną analizę iteracji odwzorowania $F_2(w, z)$ dla $w = w^s$, różnych stanów początkowych i wartości $\sigma > \sigma_c$. We wszystkich przypadkach trajektorie w stanie ustalonym oscylują pomiędzy dwoma stanami orbity periodycznej. Na rysunku 5.15a można zauważyć, że bifurkacje pojawiają się, gdy σ osiągnie wartość, dla której w krajobrazie średniego przystosowania pojawiają się dwa grzbiety. Grzbiety odpowiadają stanom, w których jeden z osobników ma typ zbliżony do optymalnego. Punkty orbity znajdują się na wewnętrznych zboczach tych grzbietów.

a) $q(x) = \exp(-5x^2)$ b) $q(x) = x + 1$ dla $x \in [-1; 0]$,
 $-x + 1$ dla $x \in [0; 1]$ 

Rys. 5.15. Trajektorie wartości oczekiwanych dla jednomodalnych funkcji celu.

Przedstawiono stabilne punkty stałe oraz orbity periodyczne (oznaczone *)

dla $\sigma \in [0,1; 1,4]$, $\Delta\sigma = 0,1$. Stany początkowe $s^0 = (0,2; -0,9)$ i $s^0 = (1,0; 1,2)$ oznaczono $\circ$ Fig. 5.15. Trajectories of expected values of population states for unimodal fitness functions. Stable fixed points and periodical orbits (marked with *) are presented for $\sigma \in [0.1, 1.4]$, $\Delta\sigma = 0.1$. Initial states: $s^0 = (0.2, -0.9)$, $s^0 = (1.0, 1.2)$ are marked with $\circ$ a) I. $B = 1,0$, II. $B = 1,5$, III. $B = 2,0$,
IV. $B = 2,5$ b) I. $a_2 = 1,0$, II. $a_2 = 0,1$, III. $a_2 = 0,05$,
IV. $a_2 = 0,01$ Rys. 5.16. Druga wartość własna jako funkcja σ ; a) funkcje trójkątne (D.5), $A = 1,0$,b) funkcje gaussowskie (D.3), $a_1 = 1,0$ Fig. 5.16. The second eigenvalue as a function of σ ; a) tent functions (D.5), $A = 1.0$,b) Gaussian functions (D.3), $a_1 = 1.0$

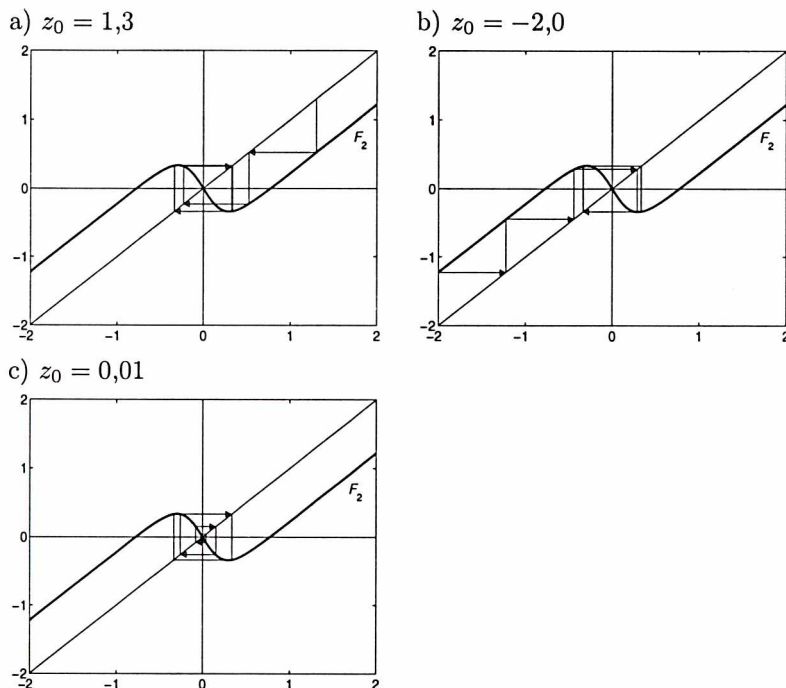


Rys. 5.17. Diagramy bifurkacyjne dla jednomodalnych funkcji celu
Fig. 5.17. Bifurcation diagrams for unimodal fitness functions

Jeżeli spojrzymy na krajobraz funkcji celu, to się okaże, że stany składające się na orbitę opisują dwie populacje ulokowane symetrycznie na obu zboczach funkcji celu (rys. 5.19). W obu populacjach jeden z osobników ma wysoką jakość, podczas gdy drugi, o typie położonym po przeciwnej stronie wzgórza, ma znacznie mniejsze przystosowanie. W kolejnej generacji następuje zamiana miejsc, mianowicie $x_1^{i+1} = -x_2^i$ oraz $x_2^{i+1} = -x_1^i$. Na przykład gdy $\sigma = 0,8$, orbita składa się z dwóch populacji (pokazanych na rys. 5.19) $(x_1^i; x_2^i) = (0,66; -0,33)$ oraz $(x_1^{i+1}; x_2^{i+1}) = (0,33; -0,66)$. Odległość euklidesowa pomiędzy rodzicem a potomkiem wynosi około $\sigma/2$.

Przypadek asymetryczny. Dla asymetrycznych funkcji jednomodalnych druga wartość własna przyjmuje także wartości mniejsze od jedności (rys. 5.16), a wartości krytyczne w_c są większe, niż w przypadku symetrycznym. Dla funkcji trójkątnej asymetrycznej druga wartość własna wynosi

$$\lambda_2 = 1 - w^s \frac{(A+B)^2}{2AB(A+B-2w^s)}, \quad (5.30)$$



Rys. 5.18. Graficzna analiza iteracji $F_2(w, z)$, dla $w = w^s$ i różnych z_0 , $\sigma = 0,8$

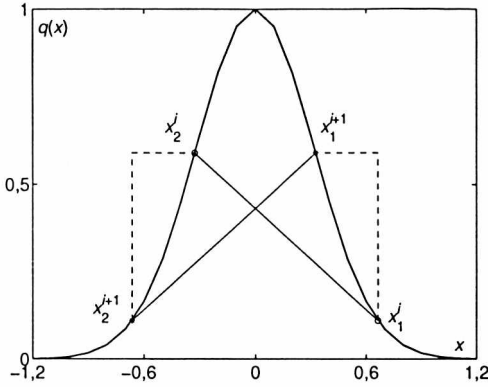
Fig. 5.18. Graphical analysis of the iteration $F_2(w, z)$ for $w = w^s$ and different z_0 , $\sigma = 0.8$

a wartość krytyczna jest równa $w_c = (4AB(A+B))/(8AB + (A+B)^2)$. W przypadku gaussowskiej funkcji asymetrycznej wartość λ_2 jest dana wzorem

$$\lambda_2 = 1 - w^s(z^s(a_{q1} - a_{q2}) - w^s(a_{q2} + a_{q1}))/2, \quad (5.31)$$

gdzie parametry a_{q1} i a_{q2} oznaczają nachylenia stoków funkcji $q_1(z)$ i $q_2(z)$ w punkcie przecięcia ich wykresów. Po podstawieniu wartości współrzędnej z^s danej wzorem (5.27) otrzymujemy $\lambda_2 = 1 - (w^s)^2 \sqrt{a_{q1}a_{q2}}$, a wartość krytyczna jest równa $w_c = \sqrt{2/\sqrt{a_{q1}a_{q2}}}$.

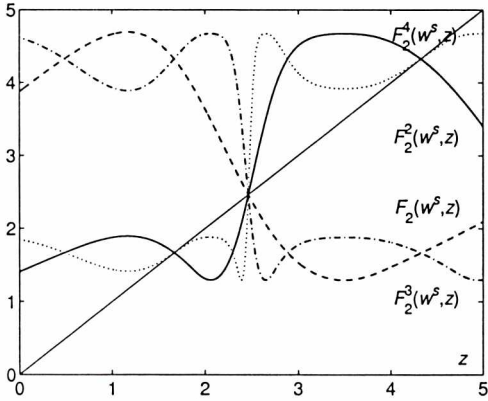
Analiza zachowania układu dynamicznego (5.16) dla dużych σ w przypadku funkcji asymetrycznych ujawniła sekwencję bifurkacji podwajania okresu prowadzącą do chaosu (rys. 5.17b, c). W przypadku funkcji trójkątnej tylko kilka bifurkacji jest widocznych, a zachowania chaotyczne pojawiają się stosunkowo szybko. W obszarach chaotycznych widoczne są także okienka periodyczne o okresie 5 i 10. Iteracja układu dynamicznego dla asymetrycznej funkcji gaussowskiej prowadzi do serii kolejnych bifurkacji i orbit o okresach 2, 4, 8, ... Okazuje się, że zachowania chaotyczne zależą od stopnia asymetrii funkcji celu i pojawiają się, gdy jest ona



○ – i -ta generacja
 * – $(i + 1)$ generacja
 osobniki jednej generacji połączono linią ciągłą, a rodziców i ich potomków linią przerywaną

Rys. 5.19. Dwie populacje tworzące orbitę periodyczną w krajobrazie jednomodalnej funkcji celu $q(x) = \exp(-5x^2)$, $\sigma = 0,8$

Fig. 5.19. Two populations forming the periodic orbit in the landscape of the unimodal fitness function $q(x) = \exp(-5x^2)$, $\sigma = 0.8$



$F_2(w^s, z)$ - linia przerywana,
 $F_2^2(w^s, z)$ - linia ciągła,
 $F_2^3(w^s, z)$ - linia przerywano-kropkowa,
 $F_2^4(w^s, z)$ - linia kropkowa

funkcja celu:

$$q(x) = \begin{cases} \exp(-x^2) & \text{dla } x \leq 0 \\ \exp(-0,05x^2) & \text{dla } x > 0 \end{cases}$$

Rys. 5.20. Złożenia jednej, dwóch, trzech i czterech funkcji F_2

Fig. 5.20. Composition of one, two, three and four functions F_2

znaczna, dla funkcji gaussowskiej, gdy $a_1/a_2 > 10$, a dla funkcji trójkątnej, gdy $B \geq 1,5$.

Wokół stanu równowagi współczynnik $\Psi \simeq 0$, a funkcja F_2 (liniowa ze względu na Ψ) wpływa na dynamikę układu (5.16) bardziej niż F_1 (kwadratowa ze względu na Ψ). Na rysunku 5.20 przedstawiono wykres złożenia jednej, dwóch i czterech funkcji F_2 dla gaussowskiej asymetrycznej jednomodalnej funkcji jakości (D.3). Punkty przecięcia wykresów z prostą $f(z) = z$ wskazują na istnienie orbit o okresach, odpowiednio, 1, 2 i 4. Oczekuje się, że podobne zachowania będzie wykazywać analizowany model ewolucji dla innych funkcji jednomodalnych, gdyż wiele z nich

może być aproksymowane przez segmenty liniowe (wtedy lokalnie zachowują się jak funkcje trójkątne lub funkcje gaussowskie).

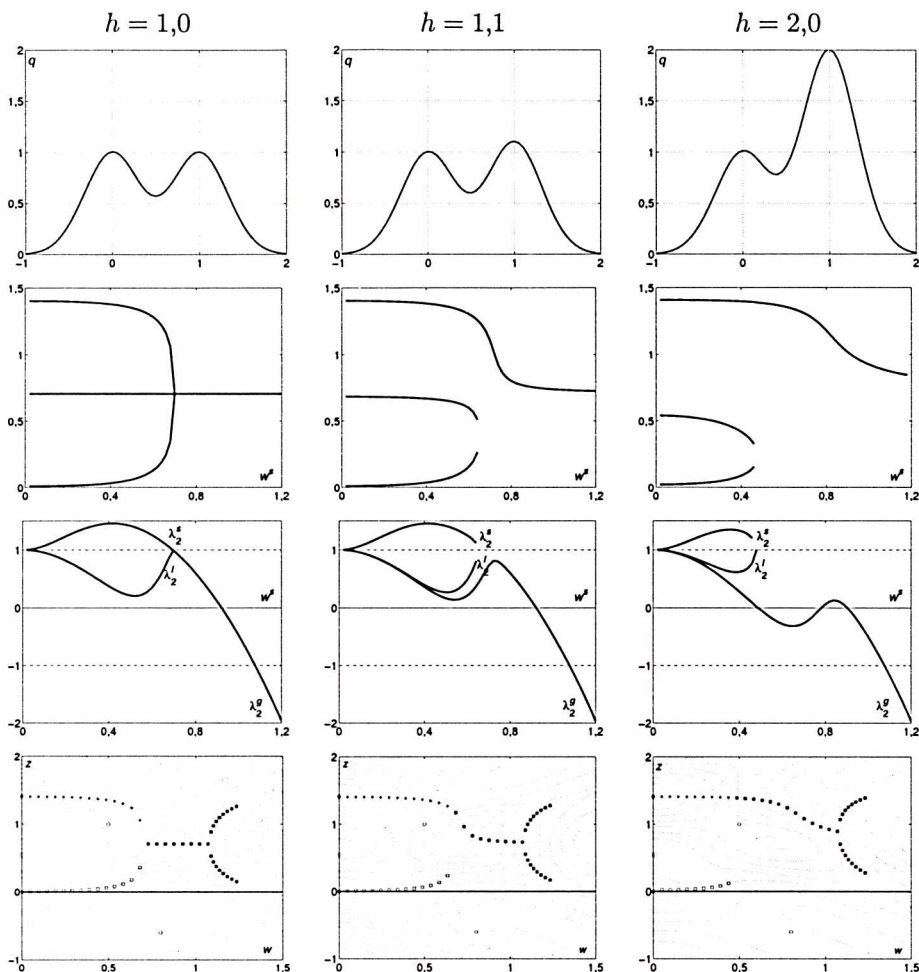
Dwumodalne funkcje celu

Analiza stabilności punktów stałych układu dynamicznego (5.16) w krajobrazie dwumodalnych funkcji przystosowania możliwa jest jedynie numerycznie, dlatego w niniejszych rozważaniach ograniczono się do sumy dwóch funkcji gaussowskich (D.4). Takie same zachowania obserwowano dla dwumodalnej asymetrycznej funkcji trójkątnej (D.6) [103]. Stabilność punktów stałych zależy od ich położenia, postaci funkcji celu oraz odchylenia standardowego mutacji. Twierdzenie 5.4.1 mówi, że siodłowe punkty stałe są niestabilne. Punkty stałe położone w pobliżu optimów są stabilne dla małych wartości odchylenia standardowego mutacji. Dla większych σ obserwuje się zachowania niestabilne.

Przypadek symetryczny. Dla funkcji symetrycznych ze wzgórzami o równych wysokościach oba punkty stałe położone w pobliżu optimów są stabilne, ich drugie wartości własne są równe, dodatnie i mniejsze od jedności (rys. 5.21a). Po połączeniu się punktów stałych pozostały punkt (siodłowy) jest stabilny dla pewnego zakresu wartości σ , a po przekroczeniu wartości krytycznej staje się niestabilny i pojawia się orbita o okresie 2.

Przypadek asymetryczny. Asymetria w przypadku funkcji dwumodalnych może oznaczać różnice nie tylko w wysokościach optimów, ale także w szerokościach wzgórz opisywanych parametrem $\zeta^2 = 1/(2a)$ (rys. 5.21–5.23). Jeżeli szerokości obu wzgórz są równe, układ dynamiczny zachowuje się podobnie jak dla przypadku symetrycznego. Jedynie wraz ze wzrostem wysokości optimum globalnego punkty stałe siodłowy i optimum lokalnego zanikają szybciej. Pozostały punkt stały dla dużych wartości σ staje się niestabilny i pojawia się orbita o okresie 2. Wartość krytyczna σ , dla której pojawia się bifurkacja, jest taka sama ($\sigma_c \simeq 1,15$), niezależnie od różnicy wysokości wzgórz.

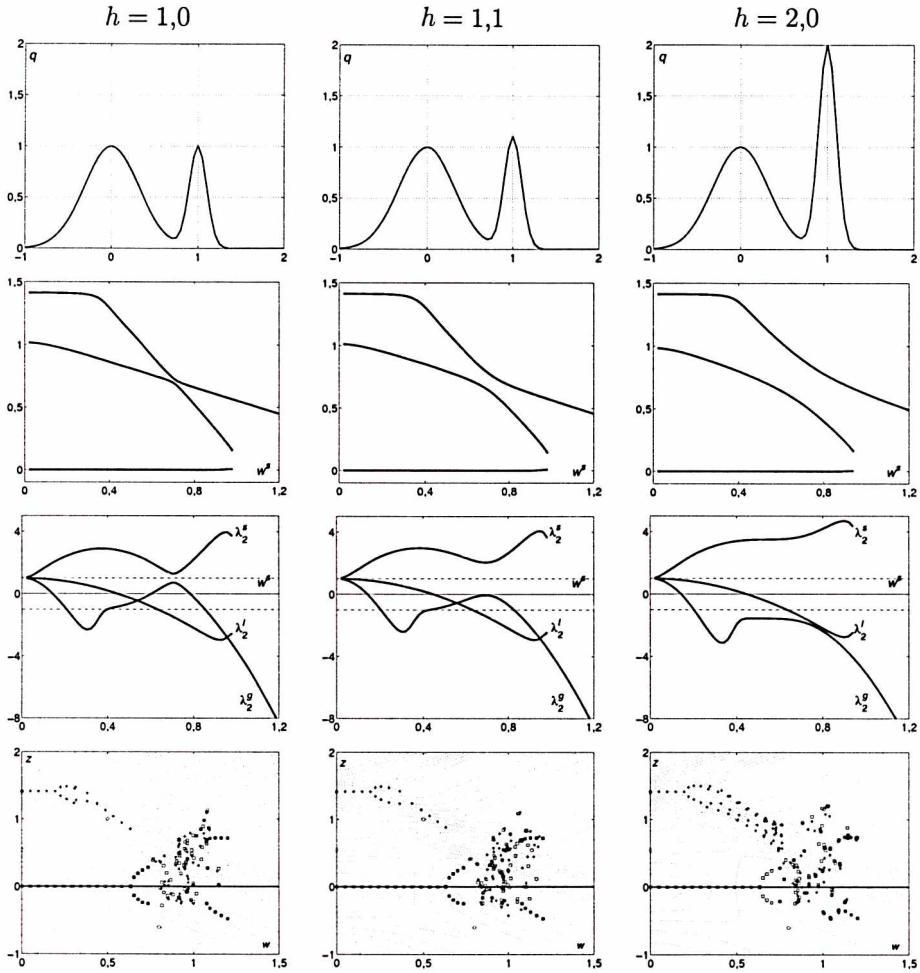
Bardziej interesujący wydaje się przypadek, gdy oba wzgórza mają różną szerokość, szczególnie wtedy, gdy wzgórze globalne jest węższe niż lokalne. Zmieniają się wtedy lokalizacja i stabilność punktów stałych. Pojawiają się bifurkacje podwajania okresu, bifurkacje różnych okresów oraz chaos. Gdy $a_2 = 10a_1$ ($\zeta_2^2 = 0,1\zeta_1^2$) (rys. 5.22), punkty stałe związane z optimami zmieniają stabilność znacznie częściej. Bifurkacja punktu przy optimum globalnym pojawia się znacznie wcześniej ($\sigma \simeq 0,25$) niż w przypadku symetrycznym. Orbita periodyczna występuje dla niewielkiego przedziału zmienności $\sigma \in (0,25; 0,45)$. Następnie pojawia się orbita o podwojonym okresie, która ponownie staje się stabilna i dla $\sigma \simeq 0,7$ punkt ten traci stabilność definitywnie. Gdy wzgórze globalne jest dwa razy wyższe od lokalnego ($h = 2$), punkt stały związany z optimum globalnym raz tracąc stabilność, już jej nie odzyskuje. Obserwuje się bifurkacje podwajania okresu, czyli orbity



Rys. 5.21. Symulacje w krajobrazie funkcji celu $q(x) = \exp(-5x^2) + h \exp[-5(x-1)^2]$.

W kolejnych wierszach przedstawiono profil funkcji, położenie punktów stałych, wykresy drugich wartości własnych oraz punkty stałe trajektorii wartości oczekiwanych. Punkty równowagi trajektorii pokazano dla dwóch stanów początkowych $s^0 = (0,8; -0,6)$ oraz $s^0 = (0,4; 1,0)$, dla zakresu odchylenia standardowego mutacji $\sigma \in [0,05; 1,55]$, $\Delta\sigma = 0,05$. Stany początkowe oznaczono przez $\circ$, a punkty stałe odpowiednio dla stanów początkowych przez $\square$ oraz $*$

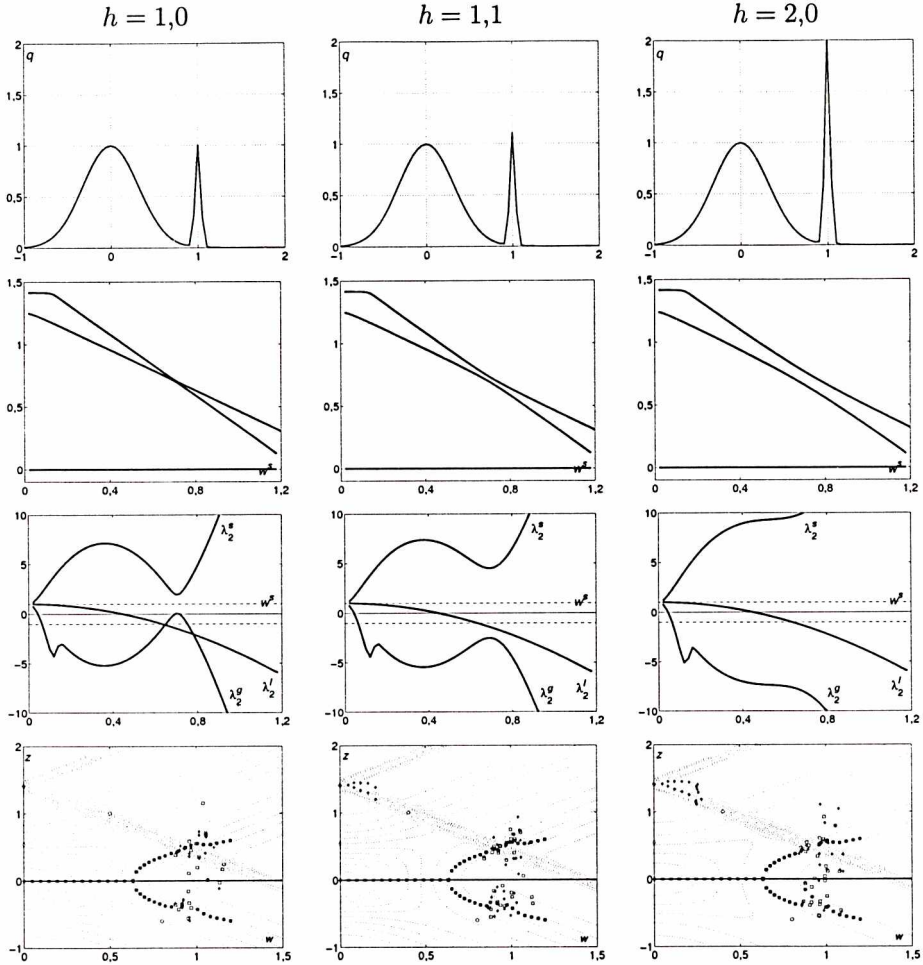
Fig. 5.21. Simulations in the landscape of the fitness function $q(x) = \exp(-5x^2) + h \exp[-5(x-1)^2]$. The function profile, locations of fixed points, magnitude of the second eigenvalue, and the fixed points of the trajectories expected values are presented in consecutive rows. Equilibrium points of trajectories are shown for two initial states: $s^0 = (0,8; -0,6)$ and $s^0 = (0,4; 1,0)$ for $\sigma \in [0,05; 1,55]$, $\Delta\sigma = 0,05$. Initial states are marked with circles, while equilibrium points for two initial states are represented by $\square$ and $*$, respectively



Rys. 5.22. Symulacje w krajobrazie funkcji $q(x) = \exp(-5x^2) + h \exp[-50(x-1)^2]$. W kolejnych wierszach przedstawiono profil funkcji celu, położenie punktów stałych, wykresy drugiej wartości własnych oraz punkty stałe trajektorii wartości oczekiwanych. Punkty równowagi trajektorii pokazano dla dwóch stanów początkowych $s^0 = (0,8; -0,6)$ oraz $s^0 = (0,4; 1,0)$, dla zakresu $\sigma \in [0,05; 1,55]$, $\Delta\sigma = 0,05$.

Stany początkowe oznaczono przez $\circ$, a punkty stałe odpowiednio dla stanów początkowych przez $\square$ oraz $*$

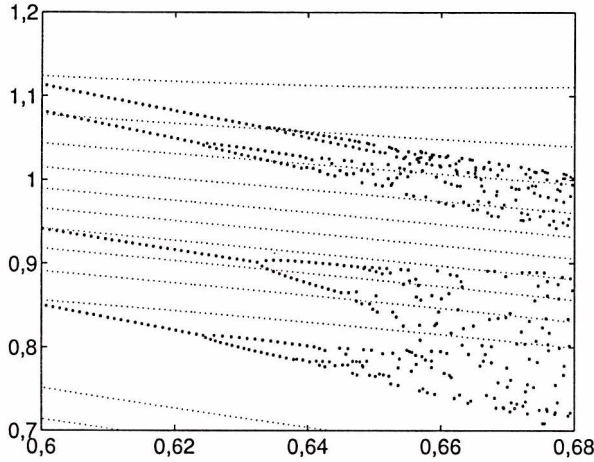
Fig. 5.22. Simulations in the landscape of the fitness function $q(x) = \exp(-5x^2) + h \exp[-50(x-1)^2]$. The function profile, locations of fixed points, magnitude of the second eigenvalue, and equilibrium points of the trajectories of expected values are presented in consecutive rows. Equilibrium points of trajectories are shown for two initial states: $s^0 = (0,8; -0,6)$ and $s^0 = (0,4; 1,0)$ for $\sigma \in [0,05; 1,55]$, $\Delta\sigma = 0,05$. Initial states are marked with circles, while equilibrium points for two initial states are represented by $\square$ and $*$, respectively



Rys. 5.23. Symulacje w krajobrazie funkcji $q(x) = \exp(-5x^2) + h \exp[-500(x-1)^2]$. W kolejnych wierszach przedstawiono profil funkcji celu, położenie punktów stałych, wykresy drugiej wartości własnych oraz punkty stałe trajektorii wartości oczekiwanych. Punkty równowagi trajektorii pokazano dla dwóch stanów początkowych $s^0 = (0,8; -0,6)$ oraz $s^0 = (0,4; 1,0)$, dla zakresu $\sigma \in [0,05; 1,55]$, $\Delta\sigma = 0,05$.

Stany początkowe oznaczono przez $\circ$, a punkty stałe odpowiednio dla stanów początkowych przez $\square$ oraz $*$

Fig. 5.23. Simulations in the landscape of the fitness function $q(x) = \exp(-5x^2) + h \exp[-500(x-1)^2]$. The function profile, locations of fixed points, magnitude of the second eigenvalue, and the fixed points of the trajectories of expected values are presented in consecutive rows. Equilibrium points of trajectories are shown for two initial states: $s^0 = (0.8; -0.6)$ and $s^0 = (0.4; 1.0)$ for $\sigma \in [0.05; 1.55]$, $\Delta\sigma = 0.05$. Initial states are marked with circles, while equilibrium points for two initial states are represented by $\square$ and $*$, respectively



Rys. 5.24. Punkty stałe trajektorii wartości oczekiwanych dla stanu $s^0 = (0.4; 1.0)$ oraz zakresu $\sigma \in [0.6; 0.68]$, $\Delta\sigma = 0.001$ (powiększony fragment rys. 5.22). Przedstawiono ostatnie 20 z 10 000 generacji

Fig. 5.24. Fixed points of the trajectories of expected values for initial state $s^0 = (0.4; 1.0)$ and the range of $\sigma \in [0.6; 0.68]$, $\Delta\sigma = 0.001$ (enlarged fragment of Fig. 5.22). The last 20 of 10 000 generations shown

o okresach 4, 8, 16 prowadzące do chaosu. Serię takich bifurkacji przedstawiono w szczególach na rys. 5.24. W danych numerycznych zaobserwowano także orbity o okresie 3 i 6. Gdy wzgórze globalne jest jeszcze węższe ($\zeta_2^2 = 0.01\zeta_1^2$, rys. 5.23), punkt stały związany z optimum globalnym traci stabilność bardzo szybko, już dla $\sigma \simeq 0.1$. Dla wzgórz o niewielkiej różnicy wysokości punkt ten staje się ponownie stabilny dla krótkiego przedziału wartości odchylenia standardowego mutacji $\sigma \in (0.65; 0.8)$. Punkt stały związany z optimum lokalnym traci stabilność dla $\sigma \simeq 0.65$. W przedziałach niestabilności obserwuje się ponownie orbity o okresie 2 i 4, a ciąg bifurkacji podwajania okresu prowadzi do zachowania chaotycznego. Zaobserwowana sekwencja orbit periodycznych była taka sama jak przewiduje twierdzenie Sarkowskiego [39].

Im słabsze jest oddziaływanie optimum globalnego na populację, tym bardziej skomplikowane staje się zachowanie układu (5.16). Punkt stały związany z optimum globalnym staje się niestabilny już dla relatywnie małego σ . Prawdopodobieństwo znalezienia takiego ekstremum jest małe i rośnie wraz ze zmniejszaniem się odchylenia standardowego mutacji. Zachowania niestabilne, orbity periodyczne, oscylacje, chaos zaobserwowano także przy analizie modeli układów dynamicznych algorytmów genetycznych z kodowaniem binarnym bez krzyżowania [210, 211, 213].

5.6. Badanie zachowań chaotycznych

Układ dynamiczny (5.16) wykazuje zachowania niestabilne i chaotyczne dla asymetrycznych funkcji celu i dużych wartości odchylenia standardowego mutacji. Teoria układów dynamicznych oferuje kilka metod badania chaosu [192]. Oprócz analizy przebiegu trajektorii czasowych, z której korzystano w poprzednim podrozdziale, stosuje się analizę wykresów fazowych, widma mocy sygnału, funkcje autokorelacji, wykładniki Lapunowa oraz przekroje Poincarego. Zachowania asymptotyczne badanego układu dynamicznego przeanalizujemy na podstawie pierwszych czterech kryteriów [112, 114].

Wykresy fazowe przedstawiają zależność prędkości wzdłuż trajektorii od położenia układu. Wykresy wykonuje się przez bezpośrednią rejestrację zmiennych (x, y) , zmiennej wraz z prędkością $(x, \dot{x})$ lub zmiennej opóźnionej o czas τ $(x(t), x(t - \tau))$. Podobnie jak dla przebiegów czasowych nie jest to bardzo dokładna metoda i pozwala jedynie na wstępną ocenę zachowań układu. Istnieją układy niechaotyczne, których wykresy fazowe przypominają układy chaotyczne. *Widmo mocy* otrzymuje się jako kwadrat modułu transformaty Fouriera $X(\omega)$ trajektorii czasowej $x(t)$: $P(\omega) = |X(\omega)|^2$. Dla układów chaotycznych w widmie występuje wiele częstotliwości, których moc koncentruje się w dolnym zakresie częstotliwości. Obliczanie widma mocy jest trudne i zwykle wymaga zastosowania procedur numerycznych do obliczenia transformaty Fouriera (DFT lub FFT), jednak jest to znacznie bardziej wiarygodne kryterium niż przebiegi czasowe czy wykresy fazowe. *Funkcja autokorelacji* jest równoważna transformacie Fouriera widma mocy. Dla przebiegu czasowego $x(t)$ o wartości oczekiwanej $E[x(t)]$ funkcję autokorelacji definiuje się jako: $C(\tau) = E[x(t)x(t+\tau)] - E[x(t)]^2$. Funkcja jest miarą podobieństwa w przebiegach i zależy jedynie od parametru opóźnienia τ między trajektorią i jej przesuniętą kopią. Procesy okresowe lub stałe są mocno skorelowane, tak więc ich funkcja autokorelacji jest stała lub oscyluje. Dla układów chaotycznych funkcja autokorelacji szybko maleje bądź zanika do zera.

Wykładnik Lapunowa jest miarą wrażliwości układu na warunki początkowe, inaczej miarą oddalania się trajektorii rozpoczynających się w dwóch blisko położonych punktach. W detekcji chaosu znaczenie ma maksymalny, dominujący wykładnik, który określa „najgorsze” zachowania układu. Dominujący wykładnik Lapunowa obliczany jest z formuły

$$\lambda = \frac{1}{N} \ln \mu_{\max}(D\mathbf{F}^N(\mathbf{s}^0)),$$

gdzie $\mu_{\max}(A)$ jest wartością własną macierzy A o maksymalnej wartości bezwzględnej, a $D\mathbf{F}^N$ jest skumulowaną, w N krokach, macierzą Jacobiego odwzorowania $\mathbf{F}$

$$D\mathbf{F}^N(\mathbf{s}^0) = D\mathbf{F}|_{\mathbf{F}^{N-1}(\mathbf{s}^0)} D\mathbf{F}|_{\mathbf{F}^{N-2}(\mathbf{s}^0)} \dots D\mathbf{F}(\mathbf{s}^0) = \prod_{i=1}^N D\mathbf{F}(\mathbf{F}^{i-1}(\mathbf{s}^0)),$$

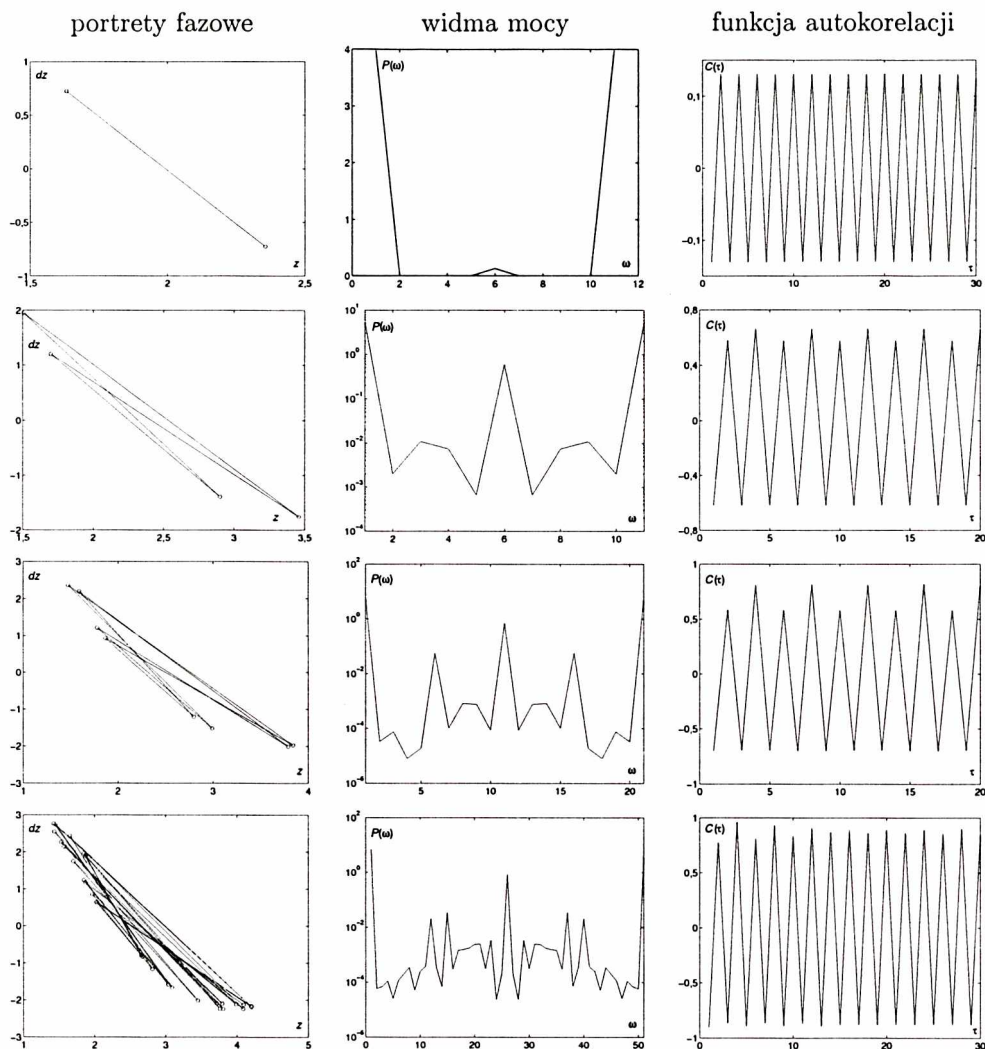
gdzie $\mathbf{s}^0$ jest stanem początkowym, $\mathbf{F}^0(\mathbf{s}) \equiv \mathbf{s}$. Dla rozważanego przypadku, gdy $\mathbf{s} = [w, z]^T$ i $\mathbf{F} = [F_1, F_2]^T$, mamy

$$D\mathbf{F} = \frac{\partial \mathbf{F}}{\partial \mathbf{s}} = \begin{pmatrix} \frac{\partial F_1}{\partial w} & \frac{\partial F_1}{\partial z} \\ \frac{\partial F_2}{\partial w} & \frac{\partial F_2}{\partial z} \end{pmatrix}.$$

Dla układów, które są zbieżne do punktów stałych lub okresowych, zachodzi $\lambda < 0$. Takie układy są asymptotycznie stabilne, a im większa wartość ujemna wykładnika, tym większy jest margines stabilności. Dla $\lambda = 0$ punkt stały jest neutralny (np. punkty bifurkacji), a układ jest w pewnego rodzaju stanie ustalonym określanym jako stabilność Lapunowa. Wartości $\lambda > 0$ są oznaką chaosu. Chociaż zerowy dominujący wykładnik nie zawsze oznacza zachowanie chaotyczne, to w większości przypadków jest to rzetelne kryterium.

Zachowania chaotyczne układu (5.16) badano dla funkcji jedno- i dwumodalnych, gaussowskich (D.2), (D.4) oraz trójkątnych (D.5). Na rysunkach 5.25, 5.26 oraz 5.27 przedstawiono wykresy fazowe, widmo mocy i funkcje autokorelacji dla funkcji gaussowskich oraz różnych odchyłeń standardowych mutacji. Ponieważ na zachowanie układu w punkcie stałym ma wpływ przede wszystkim współrzędna z , na rysunkach przedstawiono wartości kryteriów tylko dla tej współrzędnej. Dla funkcji dwumodalnej przeanalizowano dodatkowo dwa stany początkowe należące do różnych obszarów przyciągania optimum lokalnego: $\mathbf{s}^0 = (0,8; -0,6)$ i optimum globalnego: $\mathbf{s}^0 = (0,5; 1,0)$.

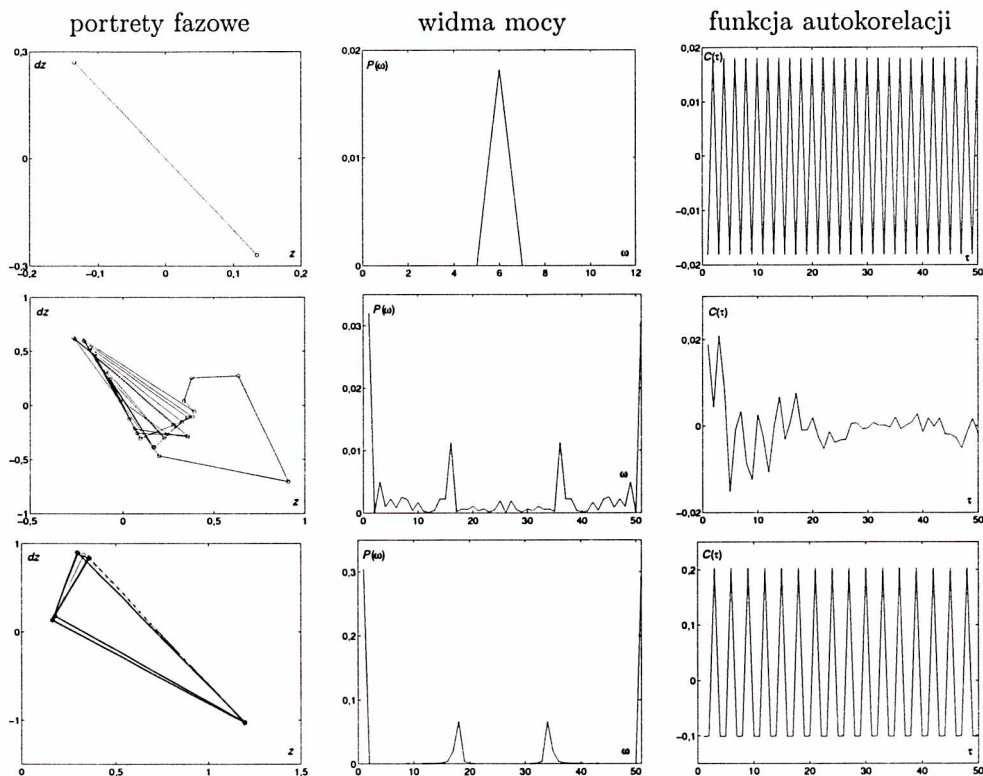
Dla funkcji jednomodalnej asymetrycznej (rys. 5.25) wszystkie trzy metody potwierdziły występowanie orbit o okresie 2, 4 oraz 8 (odpowiednio dla $\sigma = 3,2; 3,7; 3,85$). Nie wykazano, oczekiwanego na podstawie rys. 5.17b, zachowania chaotycznego dla $\sigma = 4,0$. Chociaż portret fazowy wygląda chaotycznie, zarówno analiza widma mocy, jak i funkcji autokorelacji pokazuje, że jest to orbita, prawdopodobnie o dużym okresie. Na rysunkach 5.26 i 5.27 znajdziemy potwierdzenie, że układ (5.16) przejawia zachowania chaotyczne. Dla $\sigma = 0,9$ oraz stanu początkowego w obszarze przyciągania optimum lokalnego, a także $\sigma = 0,8$ oraz stanu początkowego w obszarze przyciągania optimum globalnego wykresy fazowe wyglądają chaotycznie: widmo mocy rozpościera się po wszystkich częstotliwościach, a funkcja autokorelacji maleje do zera, co jest potwierdzeniem występowania chaosu. Ostatni wykres fazowy na rys. 5.26 pokazuje orbitę o okresie 3, co zgodnie z twierdzeniem Sarkowskiego oznacza występowanie orbit o wszystkich innych okresach.



Rys. 5.25. Analiza zachowań asymptotycznych układu dynamicznego (5.16) dla jednomodalnej asymetrycznej funkcji gaussowskiej (D.3) ($a_1 = 1$, $a_2 = 0,05$), $\sigma = 3,2; 3,7; 3,85; 4,0$ (w kolejnych wierszach)

Fig. 5.25. Analysis of asymptotic behavior of dynamical system (5.16) for unimodal asymmetrical Gaussian fitness function (D.3) ($a_1 = 1$, $a_2 = 0.05$), $\sigma = 3.2; 3.7; 3.85; 4.0$ (in consecutive rows)

Wykładniki Lapunowa obliczono dla funkcji gaussowskich (D.3), (D.4) oraz trójkątnej funkcji jednomodalnej (D.5). Na rysunku 5.28 przedstawiono wykresy dominującego wykładnika Lapunowa dla funkcji jednomodalnych. Dla symetrycz-

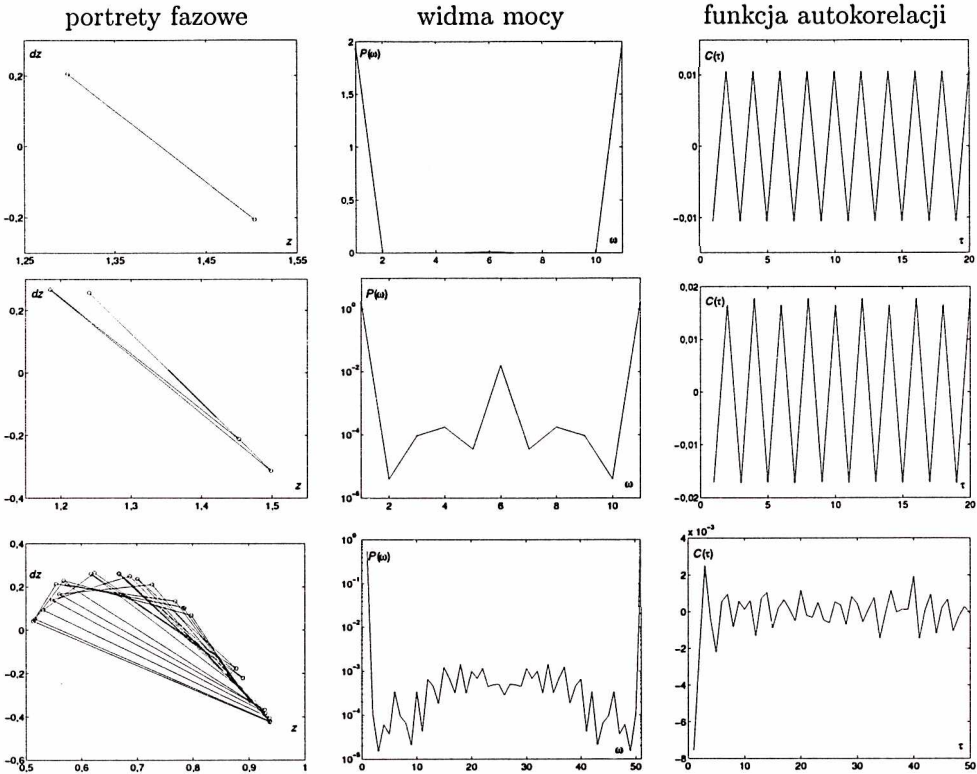


Rys. 5.26. Analiza zachowań asymptotycznych układu dynamicznego (5.16) dla dwumodalnej funkcji gaussowskiej (D.4) ($h = 2, a_1 = 5, a_2 = 50$), $s_0 = (0.8; -0.6)$, $\sigma = 0.7; 0.9; 1.2$ (w kolejnych wierszach)

Fig. 5.26. Analysis of asymptotic behavior of dynamical system (5.16) for bimodal Gaussian fitness function (D.4) ($h = 2, a_1 = 5, a_2 = 50$), $s_0 = (0.8; -0.6)$, $\sigma = 0.7; 0.9; 1.2$ (in consecutive rows)

nej jednomodalnej funkcji gaussowskiej oraz funkcji asymetrycznych (gaussowskiej i trójkątnej) i małych σ wykładniki są ujemne, przyjmując wartość zero w punktach bifurkacji. Zachowania chaotyczne układu zostały potwierdzone ($\lambda > 0$) dla dużych wartości σ w przypadku funkcji asymetrycznych (porównaj rys. 5.17). Wykładniki Lapunowa, obliczone dla funkcji dwumodalnej i stanów początkowych w obszarach przyciągania obu optimum, potwierdziły występowanie chaosu dla dużych odchyleń standardowych mutacji (rys. 5.29b). Dla funkcji jednomodalnej trójkątnej oraz dwumodalnej gaussowskiej obszar chaosu poprzedzielany jest niewielkimi (w skali σ) pasami zachowań stabilnych (okna okresowe).

Badano także wpływ różnych parametrów na dokładność obliczania λ [112]. Okazało się, że liczba iteracji N , w których liczono wykładniki wpływa na



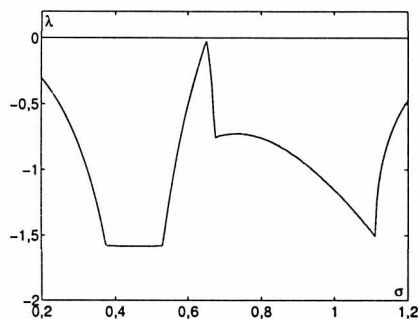
Rys. 5.27. Analiza zachowań asymptotycznych układu dynamicznego (5.16) dla dwumodalnej funkcji gaussowskiej (D.4) ($h = 2, a_1 = 5, a_2 = 50$), $s_0 = (0.5; 1.0)$, $\sigma = 0.3; 0.4; 0.8$ (w kolejnych wierszach)

Fig. 5.27. Analysis of asymptotic behavior of dynamical system (5.16) for bimodal Gaussian fitness function (D.4) ($h = 2, a_1 = 5, a_2 = 50$), $s_0 = (0.5; 1.0)$, $\sigma = 0.7; 0.9; 1.2$ (in consecutive rows)

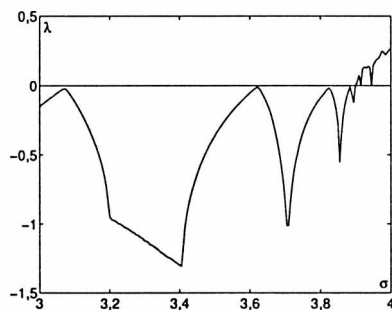
obliczoną wartość λ . Nie może być ona zbyt mała ze względu na błędy obliczeń numerycznych. Zwykle kilkaset iteracji (ok. 200) jest wystarczające. Jednak do detekcji orbit o dużym okresie potrzebna jest większa wartość N . Wybór stanu początkowego praktycznie nie ma wpływu na wartość wykładnika.

Podsumowując, przedstawiona w podrozdziale analiza potwierdziła występowanie w układzie dynamicznym (5.16) orbit periodycznych o różnych okresach oraz chaosu. Wykresy fazowe, widmo mocy oraz funkcje autokorelacji są stosunkowo łatwe do implementacji i szybkie w działaniu, ale czasami ich wyniki są sporne oraz zależą od ustawień parametrów (np. liczby iteracji czy opóźnienia τ). Wykładniki Lapunowa są bardziej wiarygodnym, aczkolwiek bardziej obliczeniowo złożonym, kryterium badania zachowania układu dynamicznego.

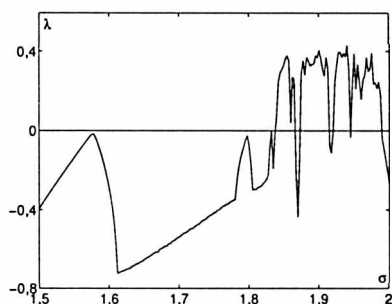
a) $q(x) = \exp(-5x^2)$



b) $q(x) = \exp(-x^2)$ dla $x \leq 0$,
 $\exp(-0,05x^2)$ dla $x > 0$



c)



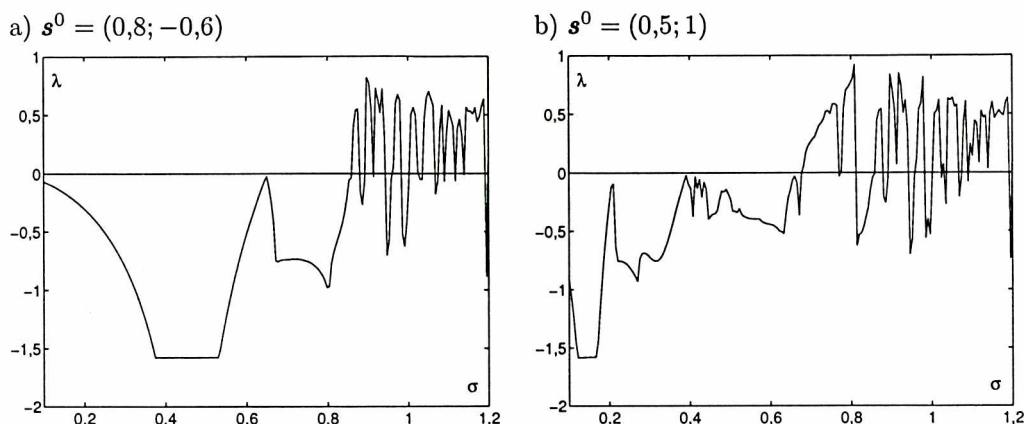
$$q(x) = \begin{cases} x + 1 & \text{dla } x \in [-1; 0) \\ -x/2,5 + 1 & \text{dla } x \in [0; 2,5] \\ 0 & \text{poza} \end{cases}$$

Rys. 5.28. Dominujący wykładnik Lapunowa jako funkcja σ
 Fig. 5.28. The dominant Lyapunov exponent as a function of σ

5.7. Oczekiwany czas zbieżności

Liczba iteracji potrzebna do osiągnięcia optimum lokalnego jest ważnym wskaźnikiem efektywności metod optymalizacyjnych. Dla metod ewolucyjnych oczekiwany czas zbieżności obliczano przede wszystkim stosując łańcuchy Markowa [194, 208]. W przypadku procesów dynamicznych określa się konieczny czas zbieżności do punktu stałego.

Oczekiwany czas zbieżności dla układu dynamicznego (5.16) analizowano w krajobrazach jednomodalnej oraz dwumodalnej gaussowskiej funkcji celu. Na rysunku 5.30 przedstawiono uzyskane za pomocą symulacji numerycznej wyniki reprezentujące liczbę generacji potrzebną do stabilizacji procesu, czyli do momentu, gdy wartość bezwzględna różnicy pomiędzy wartościami oczekiwanymi współrzędnych w i z w dwóch kolejnych generacjach jest mniejsza niż założona granica dokładności $\varepsilon = 10^{-8}$. Wartość ε ustalono empirycznie, tak aby zachować dużą dokładność, bez wpływu błędów obliczeniowych w rozsądnym czasie obliczeń. Zwiększanie dokładności nie wpływa na kształt wykresu, powoduje natomiast zwiększenie liczby



Rys. 5.29. Dominujący wykładnik Lapunowa jako funkcja σ dla funkcji celu

$q(x) = \exp(-5x^2) + 2 \exp[-50(x-1)^2]$ oraz dwóch stanów początkowych

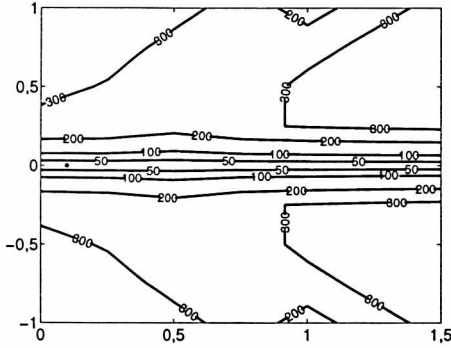
Fig. 5.29. The dominant Lyapunov exponent as a function of σ for fitness function $q(x) = \exp(-5x^2) + 2 \exp[-50(x-1)^2]$ and two initial states

generacji potrzebnych do znalezienia punktu stałego. Stany początkowe wybrane zostały na siatce pokrywającej prezentowany obszar.

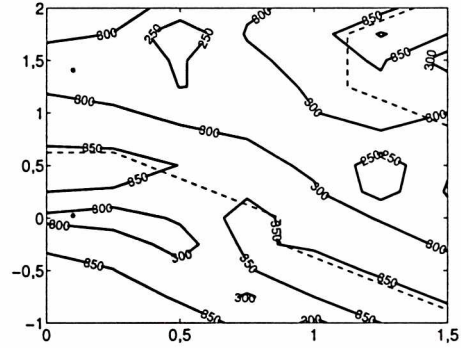
Dla funkcji jednomodalnych najmniejszą liczbę generacji do osiągnięcia stanu ustalonego potrzebują populacje, których stany początkowe są położone na i w najbliższym otoczeniu osi W . Są to więc populacje znajdujące się już w pobliżu optimum. Nieco dłużej ewoluują populacje, w których jeden z elementów ma bardzo dużą jakość (stany położone na dwóch grzbietach biegnących wzdłuż prostych $w = z$ i $w = -z$ w krajobrazie średniego przystosowania). Najwięcej czasu, aby się ustabilizować, potrzebują populacje zainicjalizowane w stanach o małej wartości średniego przystosowania. Dla funkcji dwumodalnej najszybciej stabilizują się populacje o stanach początkowych należących do zbioru przyciągania optimum globalnego.

Wcześniejsze rozważania pokazały, że analizowany proces ewolucyjny bardzo szybko lokuje populacje w stanach w pobliżu kanału ewolucyjnego. Sprawdźmy więc, ile generacji potrzebnych jest, aby populacja znalazła się w okolicy prostej $w = \sqrt{2/\pi}\sigma$. Dla uproszczenia obliczeń zbadamy jak szybko współrzędna w stanu populacji przyjmuje wartość poniżej σ . Na rysunku 5.31 przedstawiono liczbę generacji potrzebną do osiągnięcia prostej $w = \sigma$ dla jedno- i dwumodalnej funkcji gausowskiej. Populacje znajdują się w pobliżu kanału ewolucyjnego znacznie szybciej niż zajmuje im znalezienie punktu stałego, średnio potrzebują tylko kilku generacji. Efekt ten nie zależy od wartości odchylenia standardowego mutacji i odpowiada opisanym wcześniej skokom trajektorii wartości oczekiwanych (zob. rys. 5.3, 5.4). Większość czasu (mierzonego w generacjach) trajektorie modelu spędzają w pobliżu prostej $w = \sigma$, gdzie różnica w typach osobników wynosi około odchylenie

a) $q(x) = \exp(-5x^2)$

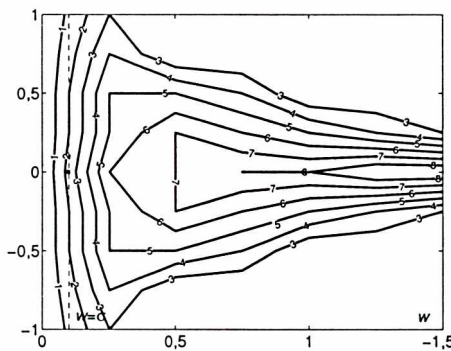


b) $q(x) = \exp(-5x^2) + 2 \exp[-5(x-1)^2]$

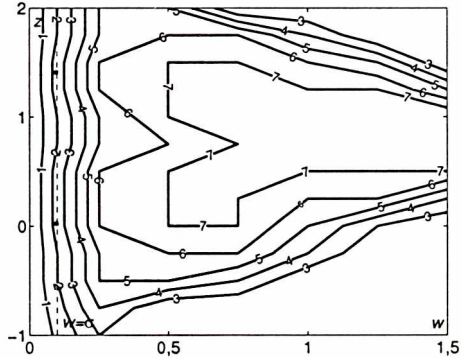


Rys. 5.30. Liczba generacji do stabilizacji procesu dla $\sigma = 0,1$. Stany początkowe z siatki: $w \in [0; 1,5]$, a) $z \in [-1,0; 1,0]$, b) $z \in [-1,5; 1,5]$, o kroku 0,25. Punkty stałe oznaczono przez *. Linie przerywane oddzielają obszary przyciągania obu maksimów Fig. 5.30. The number of generations necessary to stabilize the process for $\sigma = 0.1$. Initial states are located on the grid: $w \in [0; 1,5]$, a) $z \in [-1,0; 1,0]$, b) $z \in [-1,5; 1,5]$, with 0.25 step. Fixed points are marked with *. Dotted lines separate the basins of attraction of both maxima

a) $q(x) = \exp(-5x^2)$



b) $q(x) = \exp(-5x^2) + 2 \exp[-5(x-1)^2]$

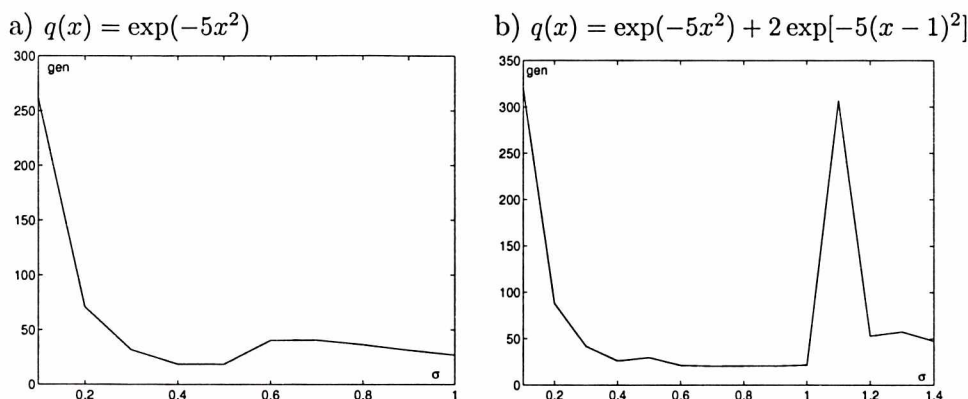


Rys. 5.31. Liczba generacji potrzebnych do osiągnięcia prostej $w = \sigma$ dla $\sigma = 0,1$.

Stany początkowe z siatki jak na rys. 5.30. Punkty stałe zaznaczono przez *
Fig. 5.31. The number of generations necessary to reach the line $w = \sigma$ for $\sigma = 0.1$. Initial states are located on the grid (see Fig. 5.30). Fixed points are marked with *

standardowe mutacji i presja selekcyjna jest słaba. Trajektorie wolno przesuwają się wzdłuż osi Z , w kierunku optimum z krokiem nie większym niż σ .

Analiza punktów stałych pokazała, że dokładna lokalizacja optimum wymaga zmniejszenia odchylenia standardowego mutacji, a tym samym zwiększa się liczba iteracji koniecznych do znalezienia optimum. Średni czas potrzebny do osiągnięcia



Rys. 5.32. Średni czas osiągnięcia punktu stałego lub orbity w funkcji σ .

Liczbę generacji uśredniono z 90. trajektorii inicjowanych w stanach położonych na siatce jak na rys. 5.30, 5.31

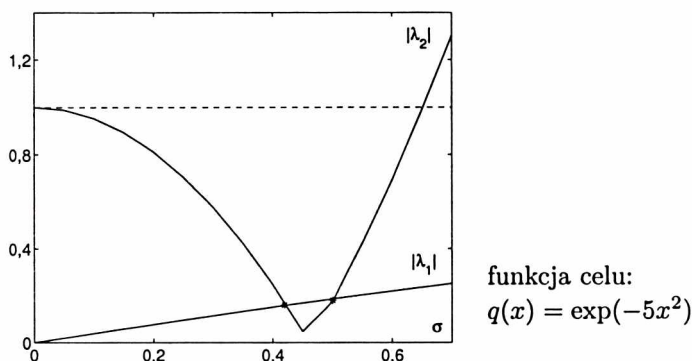
Fig. 5.32. Mean time necessary to reach the fixed point or orbit as the function of σ .

The number of generations are the averages of 90 trajectories started from states located on the grid (see Figs. 5.30, 5.31)

punktu stałego dla różnych wartości σ przedstawiono na rys. 5.32. Dla dużych wartości σ populacja potrzebuje mniej niż sto generacji, podczas gdy dla $\sigma < 0,2$ kilkaset. Tak więc praktyczny wybór odpowiedniej (optymalnej) wartości odchylenia standardowego mutacji powinien zależeć od założonego celu eksperymentu lub jego fazy. Do „zgrubnego” i szybkiego zlokalizowania optimum wystarczy przyjęcie dużych wartości σ . Większa dokładność wymaga mniejszych wartości σ i dłuższego czasu obliczeń⁵. Nagły wzrost liczby obliczeń na rys. 5.32b dla funkcji dwumodalnych występuje w pobliżu punktu bifurkacji i wynika ze zmiany typu stabilności – zbieżności do orbity, a nie do punktu. Dla funkcji jednomodalnych wzrost potrzebnych obliczeń dla bifurkacyjnych wartości σ też występuje, jest jednak znacznie mniejszy.

Oczekiwany czas zbieżności można analizować także lokalnie w bliskim otoczeniu punktów stałych, badając wartości bezwzględne wartości własnych λ_1 i λ_2 macierzy przybliżenia liniowego (5.22). Przyjmując, że punkt stały jest stabilny, czas zbieżności asymptotycznej jest determinowany przez wartość własną przyjmującą większą wartość bezwzględną. Przeanalizujemy populację w stanie położonym niedaleko od punktu stałego dla jednomodalnej funkcji gaussowskiej (D.2) (rys. 5.33). Dla bardzo małych wartości σ zachodzi $|\lambda_1| \simeq 0$, podczas gdy $|\lambda_2| \rightarrow 1$ i zbieżność do punktu stałego jest bardzo wolna. Dla wzrastających wartości σ wzrasta także wartość $|\lambda_1|$, a wartość $|\lambda_2|$ maleje. Interesujące są punkty, a więc i wielkość σ , dla których wartość bezwzględna obu wartości własnych jest równa. Czas zbieżno-

⁵ Do doboru parametru σ w zadaniach praktycznych wrócimy w kolejnym rozdziale.



Rys. 5.33. Wartości bezwzględne wartości własnych układu dynamicznego (5.16)

Fig. 5.33. Absolute values of the eigenvalues of the dynamical system (5.16)

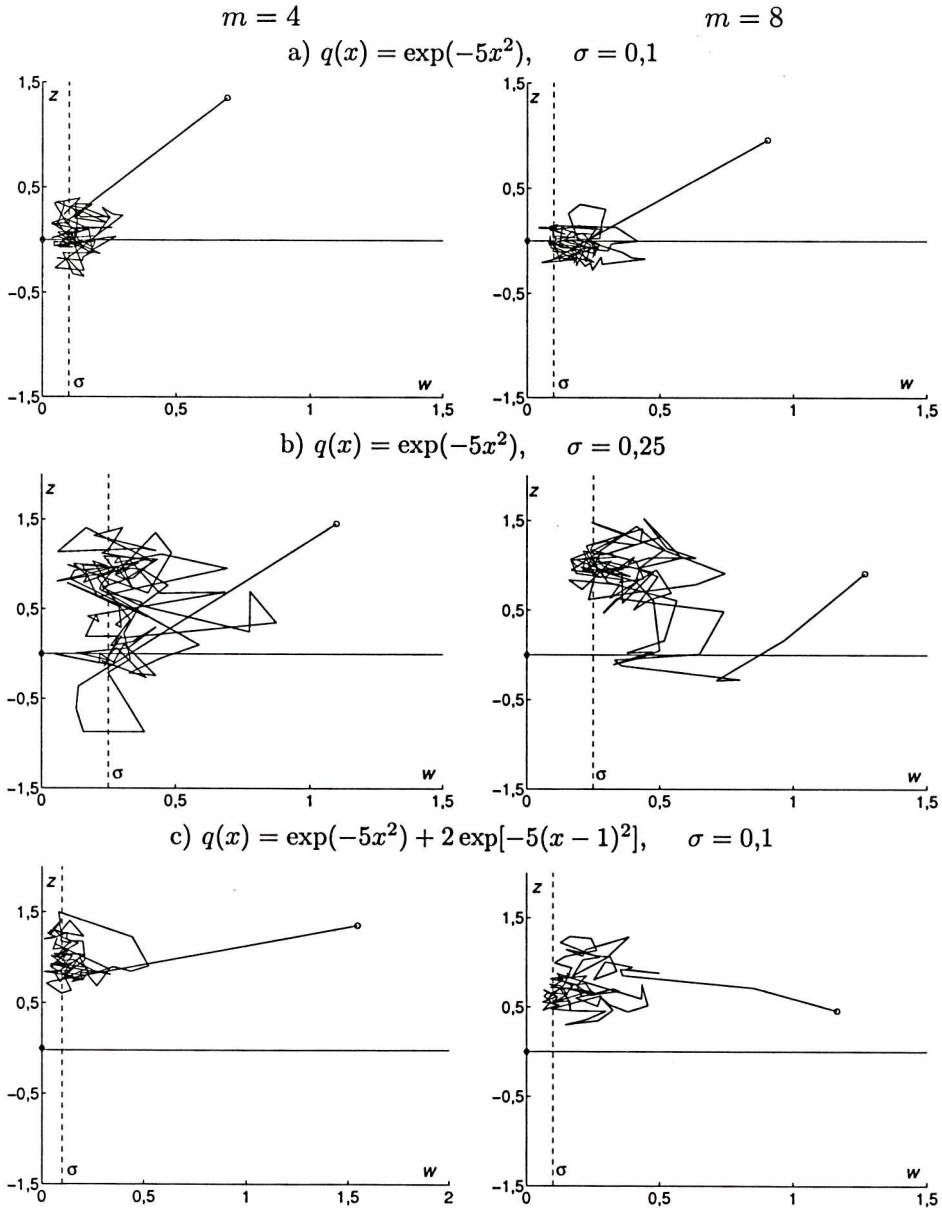
ści nie jest wtedy „wydłużany” przez żadną wartość własną. Porównując wyniki z rys. 5.33 i 5.32a można zauważyć, że lokalna analiza jest zbieżna z wynikami statystycznymi, szczególnie dla zakresu $\sigma \in [0, 2; 0, 6]$. Można zatem oczekiwać, że (przynajmniej dla badanej funkcji przystosowania) ruch wzdłuż prostej $w = \sigma$ ma taką samą dynamikę jak zachowania w pobliżu punktu stałego.

Przeprowadzona w tym podrozdziale analiza oczekiwanego czasu zbieżności do punktów stałych potwierdziła raz jeszcze dwoistą dynamikę badanego procesu. Szybki ruch w kierunku kanału ewolucyjnego poprzedza powolną zbieżność do punktu stałego. Małe różnice w typach prawie homogenicznej populacji powodują, że presja selekcyjna jest słaba i proces ewolucji wyhamowuje. Metody ewolucyjne szybko lokalizują okolice optimów, lecz proces znalezienia ich dokładnych wartości jest powolny. Dobrą praktyką jest więc łączenie metod ewolucyjnych z klasycznymi metodami lokalnymi, które szybko i precyzyjnie określą położenie zlokalizowanych przez ewolucję optimów.

5.8. Uwagi dotyczące kilkuelementowych populacji

Zwiększenie liczebności populacji bądź wymiarowości przestrzeni poszukiwań powoduje, że analiza ewolucji staje się bardzo trudna (jeżeli wręcz możliwa), ponieważ wymiarowość przestrzeni stanów populacji szybko wzrasta. Dodatkowo dochodzą jeszcze ograniczenia przestrzeni związane z relacją równoważności. Zachowania większych populacji można więc analizować za pomocą głównie symulacji komputerowych.

Analogicznie do przypadku populacji dwuelementowych, aby przedstawić populację z m osobnikami opisanymi przez jedną tylko cechę ($n = 1$), wprowadzimy



Rys. 5.34. Ewolucja małych populacji we współrzędnych (w, z)
 Fig. 5.34. Evolution of small populations in the coordinates (w, z)

nowe współrzędne:

$$w^i = \sqrt{\frac{1}{m} \sum_{j=1}^m (x_j^i - z^i)^2}, \quad z^i = \frac{1}{m} \sum_{j=1}^m x_j^i.$$

Współrzędna w jest miarą rozproszenia populacji (odległość stanu populacji od osi jednorodności, na której znajdują się stany odpowiadające jednorodnym populacjom) zaś współrzędna z opisuje średnią populacji (lokalizację stanu wzdłuż osi jednorodności). W związku z przyjętymi uproszczeniami funkcja przystosowania nie może być skojarzona ze współrzędnymi w i z , dlatego mapa średniego przystosowania nie pojawia się już na rysunkach.

Przedstawione symulacje ewolucji kilkuelementowych populacji w krajobrazie jedno- i dwumodalnej funkcji przystosowania (rys. 5.34) powtarzają charakterystyczne zachowania populacji dwuelementowych. Obserwuje się skok populacji w kierunku osi jednorodności, co oznacza ujednolicanie populacji oraz wolne fluktuacje wzdłuż osi identyczności w kierunku optimum. Populacje początkowo znacznie rozproszone szybko się skupiają, niezależnie od funkcji przystosowania. Populacja przeszukuje przestrzeń jako klastery o średnicy rzędu σ . Wydaje się, że fluktuacje populacji w okolicy optimum są mniejsze dla większych populacji. Im większa populacja, tym stany zajmowane przez populację są bliżej osi w i ułożone dalej od osi jednorodności. Większe populacje są więc bardziej różnorodne, lecz mają mniejszą zdolność penetracji przestrzeni poszukiwań.

Podobnie zwiększanie wymiarowości przestrzeni poszukiwań nie wpływa na zmianę dynamiki ewolucji [69]. Nadal obserwuje się szybką unifikację populacji oraz powolny ruch w kierunku optimum. Opisane badania procesu ewolucyjnego dla większych populacji korespondują z wynikami otrzymanymi dla modeli biologii molekularnej [215].

6. Przykłady wykorzystania analizy dynamiki populacji

Przedstawione w poprzednich rozdziałach wyniki analizy dynamiki modeli populacji ewoluujących zgodnie z metodą poszukiwań ewolucyjnych z miękką selekcją można wykorzystać także do badań innych aspektów metod ewolucyjnych, między innymi do analizy różnorodności populacji oraz sposobów określania nieznanymi parametrów procesu ewolucji i funkcji celu [107, 111, 115, 118].

6.1. Różnorodność populacji

W biologii *różnorodność genetyczna* jest sumą aktualnych lub potencjalnych informacji genetycznych oraz możliwości zawartych w genach osobników, populacji lub gatunków¹. W populacjach naturalnych podtrzymywanie zróżnicowania jest ważne, aby populacje mogły szybko adaptować się do zmieniającego się otoczenia, eksplorować nowe nisze w środowisku oraz posiadać system obrony przeciwko wrogom.

Dla użytkowników metod ewolucyjnych różnorodność populacji jest wskaźnikiem mówiącym jak bardzo osobniki danej populacji różnią się od siebie. Obserwowanie zróżnicowania populacji w trakcie działania algorytmu jest ważne z kilku powodów. Uważa się, że utrzymanie różnorodności populacji na odpowiednim poziomie może zapobiec *przedwczesnej zbieżności* (ang. *premature convergence*)² metody do optimum lokalnego [37, 78]. Populacje o dużej różnorodności mogą eksplorować rozległe obszary przestrzeni w poszukiwaniu nowych optimumów. Duże zróżnicowanie populacji ma znaczenie dla szybkiej adaptacji w dynamicznie zmieniającym się środowisku. Spadek zróżnicowania poniżej pewnego niskiego poziomu może być oznaką znalezienia optimum i posłużyć jako warunek zatrzymania lub ponownego uruchomienia algorytmu [139].

Badania symulacyjne ewolucji fenotypowej [28, 63, 64] wykazały, że początkowo rozproszone populacje (o rozproszeniu rzędu kilku σ) szybko koncentrują się i przeszukują przestrzeń jako zwarty klaster osobników. Skupianie się populacji jest spowodowane głównie działaniem selekcji, która wybiera do reprodukcji tylko niektóre osobniki. Zjawisko to, znane także w populacjach naturalnych, jest

¹ Definicja ta jest podawana w dokumentach organizacji FAO, źródło: www.fao.org.

² Zjawisko przedwczesnej zbieżności występuje, gdy populacja osiągnie taki stan, że operatory ewolucyjne nie są w stanie generować potomków lepszych niż ich rodzice.

zwane przez ewolucjonistów *dryfem genetycznym* [119]. Analiza szybkiej utraty początkowo dużej różnorodności populacji pokazuje, że może do niej dojść nawet wtedy, gdy na populację nie działa żadna presja selekcyjna, a mechanizmem odpowiedzialnym jest kumulacja błędów próbkowania (ang. *sampling errors*) [25, 83]. Zwiększenie reprezentacji danego typu w populacji w wyniku losowania rodziców może doprowadzić do dominacji jednego elementu³.

Fakt szybkiej koncentracji początkowo rozproszonej populacji potwierdza analiza wartości oczekiwanej współrzędnej w dwuelementowej populacji, zadanej przez równanie (5.5), opisującej rozproszenie populacji. Początkowo duże wartości w szybko maleją w każdym krajobrazie adaptacyjnym, w którym populacja ewoluuje, a więc nawet wtedy, gdy nie działa presja selekcyjna (zobacz też rys. 5.6). Współrzędna w^s punktu stałego (opisana równaniem (5.19)) jest stała dla danego odchylenia standardowego mutacji i nie zależy od funkcji przystosowania. Podobna prawidłowość jest obserwowana w symulacjach ewolucji dla populacji o większej liczebności [25, 64]. Rozproszenie kilkuelementowej populacji szybko maleje, a następnie waha się w granicach odchylenia standardowego mutacji (por. rys. 5.34). Dla populacji nieskończonych wariancja rozkładu populacji (4.6) osiąga stan stabilny także już w kilku iteracjach. W tym przypadku wartość wariancji rozkładu stacjonarnego zależy nie tylko od σ mutacji, ale także od krajobrazu adaptacyjnego (nachylenia zbocza gaussowskiej funkcji jakości). Na węższych szczytach różnorodność populacji jest mniejsza.

Uzasadniony wydaje się wniosek, że skupianie się ewoluujących populacji jest istotną cechą procesu ewolucji. Na wielkość rozproszenia populacji wpływają dodatkowo stosowane w metodzie operatory. Selekcja i część operatorów rekombinacji zmniejszają różnorodność populacji, natomiast mutacja jest mechanizmem, który ją podtrzymuje lub zwiększa. W kolejnych podrozdziałach rozpatrzmy wpływ różnych sposobów reprodukcji na zróżnicowanie populacji.

6.1.1. Miary zróżnicowania populacji

Dla analizy rozproszenia populacji należy zdefiniować jego miarę. W literaturze spotyka się wiele wskaźników różnorodności populacji w zależności od sposobu kodowania [10, 129, 134]. Dla osobników opisanych wektorem cech rzeczywistoliczbowych najbardziej naturalną miarą rozproszenia wydaje się odchylenie od średniej w populacji (osobnika o cechach równych średnim wartościom danej współrzędnej dla całej populacji), co dla jednowymiarowych przestrzeni poszukiwań odpowiada odchyleniu standardowemu. Dla rzeczywistych przestrzeni

³ Łatwo to pokazać na przykładzie losowania kolorowych kul z urny [25].

n -wymiarowych, w pracy [69], zaproponowano wykorzystanie odległości euklidesowej pomiędzy osobnikami:

$$div_1(\mathbf{s}) = \left(\sum_{j=1}^n D_j^2 \right)^{1/2}, \quad (6.1)$$

gdzie $\mathbf{s}$ opisuje stan populacji $\mathbf{s} = (x_1, x_2, \dots, x_m)$, a D_j^2 jest wariancją

$$D_j^2 = \frac{1}{m} \sum_{k=1}^m (x_{k,j} - \bar{x}_j)^2 \quad \text{ i } \quad \bar{x}_j = \frac{1}{m} \sum_{k=1}^m x_{k,j},$$

$j = 1, 2, \dots, n$. Miara (6.1) jest podobna do miar różnorodności stosowanych w biologii [135].

Wskaźniki rozproszenia stosowane w algorytmach genetycznych zwykle zliczają różnice między genotypami osobników. W celu łatwiejszego porównania operatorów dla kodowania binarnego i rzeczywistoliczbowego zaadaptowano miarę zdefiniowaną jako średnią odległość między wszystkimi parami osobników w populacji [10, 165]:

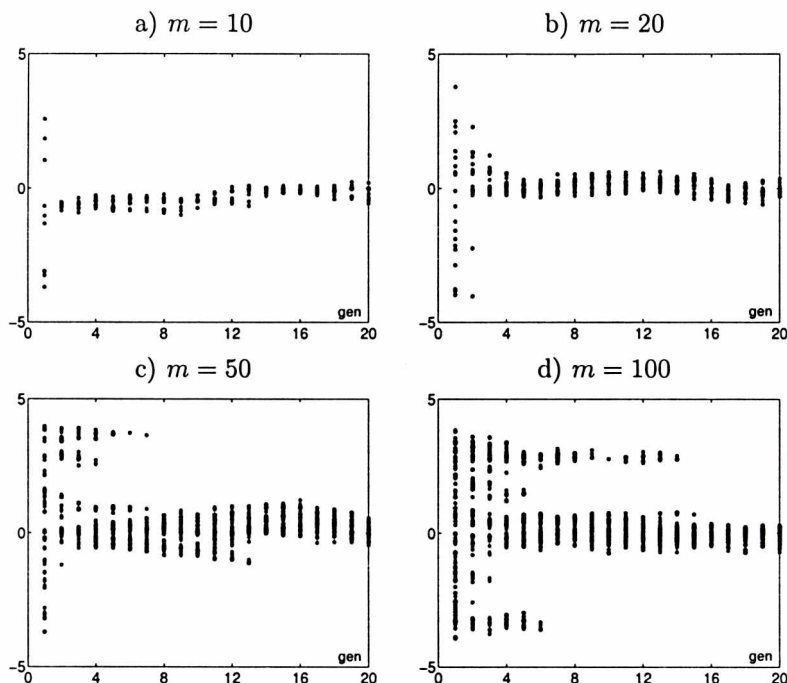
$$div_2(\mathbf{s}) = \frac{2}{m(m-1)} \sum_{i=1}^{m-1} \sum_{j=i+1}^m |\mathbf{x}_i, \mathbf{x}_j|, \quad (6.2)$$

gdzie m jest liczebnością populacji, a $|\cdot, \cdot|$ jest odległością Hamminga dla wektorów binarnych $\mathbf{x}_i$ i $\mathbf{x}_j$, natomiast odległością euklidesową $|\mathbf{x}_i, \mathbf{x}_j| = \sqrt{\sum_{k=1}^n (x_{i,k} - x_{j,k})^2}$ dla wektorów rzeczywistoliczbowych. Aby uzyskać wartości rozproszenia mniej zależne od wielkości populacji, w symulacjach wskaźnik (6.2) normowano, dzieląc przez liczbę różnych par osobników w populacji.

6.1.2. Operatory selekcji

Selekcja jest głównym operatorem zmniejszającym różnorodność populacji. Analizowano, czy na wielkość zmian różnorodności ma wpływ intensywność nacisku selektywnego („miękość” selekcji). Zbadano sześć typów selekcji opisanych w rozdziale 2.2: proporcjonalną, twardą, turniejową binarną, elitarną (towarzyszącą selekcji proporcjonalnej), progową i losową.

Prezentowane symulacje przeprowadzono w krajobrazie jednowymiarowej funkcji celu $q(x) = \exp(-5x^2)$. Populację początkową losowano z rozkładem jednostajnym na odcinku $[-4; 4]$, aby zapewnić jej jak największe rozproszenie. Ewolucja przebiegała zgodnie z algorytmem przedstawionym na rys. 2.1. Po selekcji osobników według wybranej procedury dokonywano mutacji z rozkładem normalnym o odchyleniu standardowym σ . Wartości średnie obliczano ze 100 niezależnych uruchomień algorytmu.

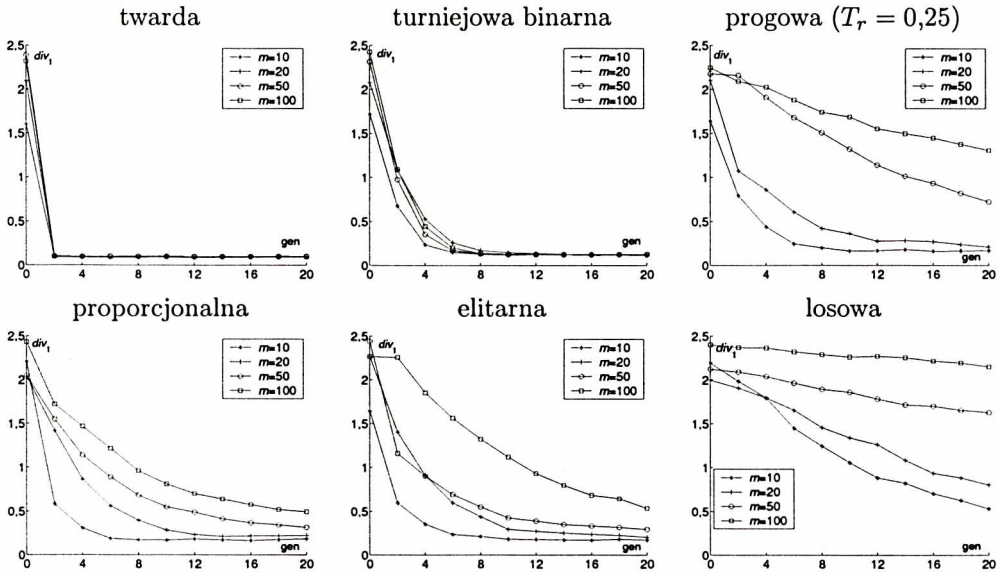


Rys. 6.1. Położenie elementów populacji ewoluującej z selekcją proporcjonalną i mutacją gaussowską. Funkcja celu $q(x) = \exp(-5x^2)$, $\sigma = 0,1$

Fig. 6.1. Locations of the elements of population evolves according to proportional selection and Gaussian mutation. Fitness function $q(x) = \exp(-5x^2)$, $\sigma = 0.1$

Na rysunku 6.1 przedstawiono położenie osobników w kolejnych 20 generacjach ewolucji z selekcją proporcjonalną dla populacji o liczebności $m = 10, 20, 50, 100$. Początkowo rozproszone populacje tracą swą różnorodność już po kilku generacjach. Osobniki podobnych typów tworzą zwarty klaster wokół optimum funkcji przystosowania. Proces unifikacji zachodzi szybciej dla mniejszych populacji, a rozproszenie populacji jest mniejsze niż dla liczniejszych populacji. W tym kontekście nie dziwią porady dla użytkowników algorytmów ewolucyjnych, aby populacje były duże [129, 172]. Wtedy ujednolicenie jest wolniejsze, a i różnorodność zwartej populacji większa.

Wielkość zmian różnorodności zależy od rodzaju selekcji (jej „miękości”). Im mocniejsza presja selekcyjna, tym szybsza unifikacja populacji. Na rysunku 6.2 przedstawiono wartość średnią (ze 100 niezależnych przebiegów) rozproszenia populacji (6.1) dla wymienionych typów selekcji. Twarda selekcja, która reprodukuje tylko najlepszego osobnika, powoduje bardzo szybką koncentrację populacji wokół niego. Jedynym źródłem zmienności jest mutacja, a średni poziom różnorodności jest w przybliżeniu równy odchyleniu standardowemu σ . Podobnie jak w twardej

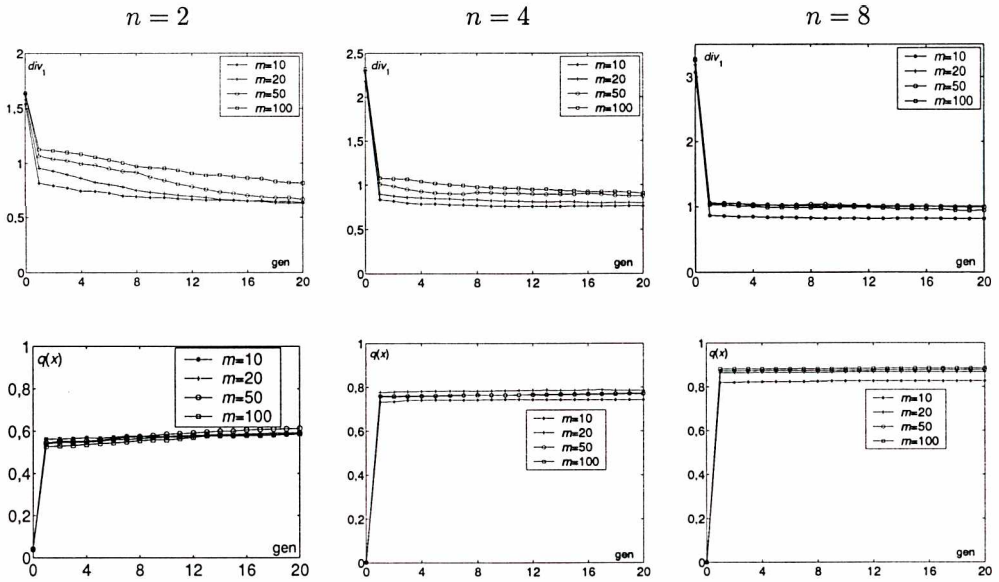


Rys. 6.2. Średnia różnorodność populacji div_1 (6.1) w 20 generacjach dla różnych rodzajów selekcji oraz różnych rozmiarów populacji, $\sigma = 0,1$

Fig. 6.2. The average diversity div_1 (6.1) of population in 20 generations for different selections and various population sizes, $\sigma = 0.1$

selekcji, niski poziom zróżnicowania utrzymuje się dla selekcji turniejowej binarnej. Jedyną różnicą jest nieco dłuższy proces dochodzenia do tego poziomu. Selekcja proporcjonalna i elitarna także redukują rozproszenie populacji, z szybkością odwrotnie proporcjonalną do wielkości populacji. Chociaż oba rodzaje selekcji różnią się jedynie zachowywaniem najlepszego osobnika w selekcji elitarnej, to utrata różnorodności przy elitaryzmie jest większa, szczególnie dla mniejszych populacji. Różnorodność populacji poddanej działaniu selekcji progowej jest stosunkowo duża. Chociaż tylko jedna czwarta populacji jest wybierana do selekcji ($T_r = 0,25$), to zróżnicowanie szczególnie dużych populacji jest zachowane. Nawet w przypadku selekcji losowej zaobserwowano tendencję do zmniejszania się różnorodności populacji o mniejszej liczebności, co może być spowodowane dryfem genetycznym, gdyż nie wszystkie osobniki są wybierane podczas selekcji [55, 37, 83, 38].

Dla selekcji proporcjonalnej analizowano także wpływ wymiarowości przestrzeni poszukiwań na ujednolicanie się populacji (rys. 6.3). W badaniach wykorzystano n -wymiarową jednomodalną funkcję celu: $q(\mathbf{x}) = \exp(-5 \sum_{j=1}^n x_j^2)$. Współrzędne typów populacji początkowych losowano dla każdego wymiaru z przedziału $[-2; 2]$. Małe populacje w wielowymiarowych przestrzeniach koncentrują się nawet szybciej niż dla $n = 1$. Wraz ze wzrostem liczby wymiarów, wpływ rozmiaru populacji na jej rozproszenie maleje. Przedstawione w dolnym wierszu rys. 6.3 średnie przysto-

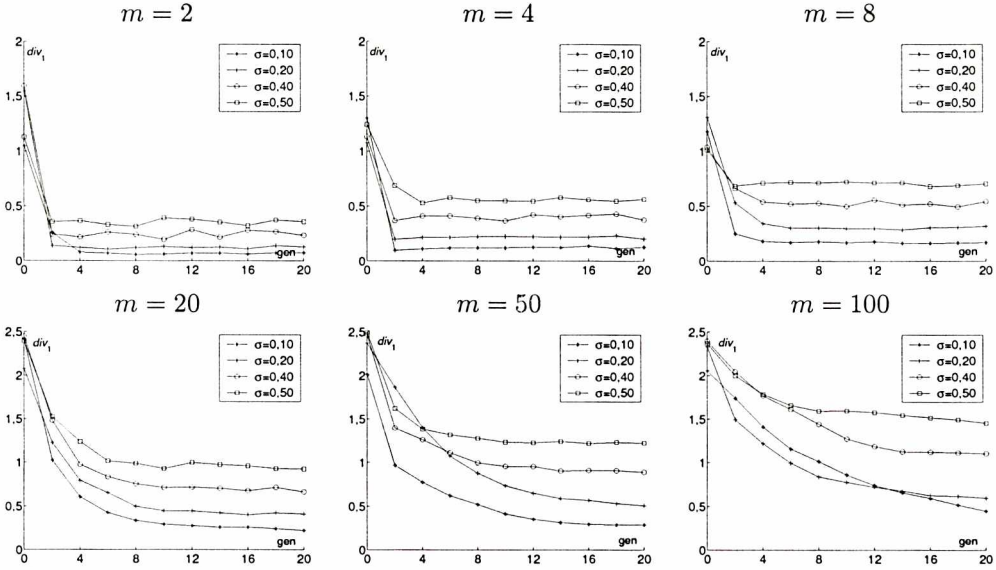


Rys. 6.3. Średnia różnorodność div_1 (6.1) (wiersz górny) oraz średnie przystosowanie populacji (wiersz dolny) w wielowymiarowych przestrzeniach poszukiwań, $\sigma = 0,1$
 Fig. 6.3. The average diversity of population (upper figures) and mean population adaptation (bottom figures) in multidimensional search spaces, $\sigma = 0.1$

sowanie populacji wskazuje, że jest ono większe dla większych n . Efekt ten może być spowodowany właśnie większą koncentracją populacji.

6.1.3. Operatory mutacji

Mutacja jest operatorem eksploracyjnym, przeszukującym nowe obszary przestrzeni poszukiwań. Jest głównym czynnikiem podtrzymującym różnorodność populacji i odpowiednią pulę genową dla rekombinacji. Warto w tym miejscu wspomnieć o długoletnich dyskusjach dotyczących korzyści ze stosowania mutacji. Niektórzy uważają, że mutacja jest podrzędnym operatorem, który nie wpływa na proces poszukiwań optimów, a jego zadaniem jest jedynie utrzymywanie odpowiedniego poziomu różnorodności w populacji i zapobieganie przedwczesnej zbieżności. Argumentuje się, że mutacje mogą dokonywać tylko niewielkich zmian cech, podczas gdy rekombinacje, wymieniając cechy pomiędzy rodzicami, mogą spowodować duże zmiany typów (co jednak nie implikuje ani wzrostu różnorodności, ani przystosowania). Badaczom przedstawiającym korzyści z użycia rekombinacji [40, 154, 160, 193] nie wtórują adwersarze, którzy wyciągają zgoła odmienne wnioski [60, 186, 189]. Obecnie rzadko traktuje się mutację jedynie jako operator pomocniczy. W wielu przypadkach jest to operator bardziej efektywny, szybszy



Rys. 6.4. Średnia różnorodność populacji div_1 (6.1) w kolejnych 20 generacjach dla różnych wartości σ i różnych liczebności populacji

Fig. 6.4. The average diversity div_1 (6.1) of population in 20 consecutive generations for different values of σ and various population size

i zwiększający różnorodność populacji [32, 53, 59]. Wydaje się, że coraz powszechniejsze jest przekonanie, iż wybór operatorów powinien zależeć od rozwiązywanego problemu. W wielu zastosowaniach mutacja może być wystarczająca do znalezienia dobrego rozwiązania. W innych, należy dobierać operator rekombinacji może znacznie przyspieszyć obliczenia [160, 193].

Badania procesu ewolucji dla różnych operatorów selekcji z mutacją opisane w rozdziale 6.1.2 pokazały, że operator mutacji raczej podtrzymuje różnorodność osobników w populacji niż ją zwiększa. Na rysunku 6.4 przedstawiono średnie rozproszenie populacji (6.1) uzyskane ze 100 niezależnych przebiegów algorytmu dla różnych wielkości odchylenia standardowego mutacji σ . Symulacje przeprowadzono w krajobrazie jednomodalnej funkcji przystosowania $q(x) = \exp(-5x^2)$. Populacje początkowe były losowane z rozkładem jednostajnym na odcinku $[-4; 4]$ dla $m = 20, 50, 100$ i $[-2; 2]$ dla $m = 2, 4, 8$. Proces ewolucji przebiegał z selekcją proporcjonalną. Wstępnie różnorodne populacje szybko się ujednolicały, a ich rozproszenie wahało się w granicach σ dla małych rozmiarów populacji i dwa lub trzy σ dla większych.

Wcześniejsze badania dużych populacji [62] pokazują, że nawet gdy zasięg mutacji jest mały, początkowo jednorodne populacje mogą łatwo osiągnąć wyższy

poziom różnorodności. W tym przypadku rekombinacja nie jest potrzebna do podtrzymania różnorodności.

6.1.4. Operatory rekombinacji

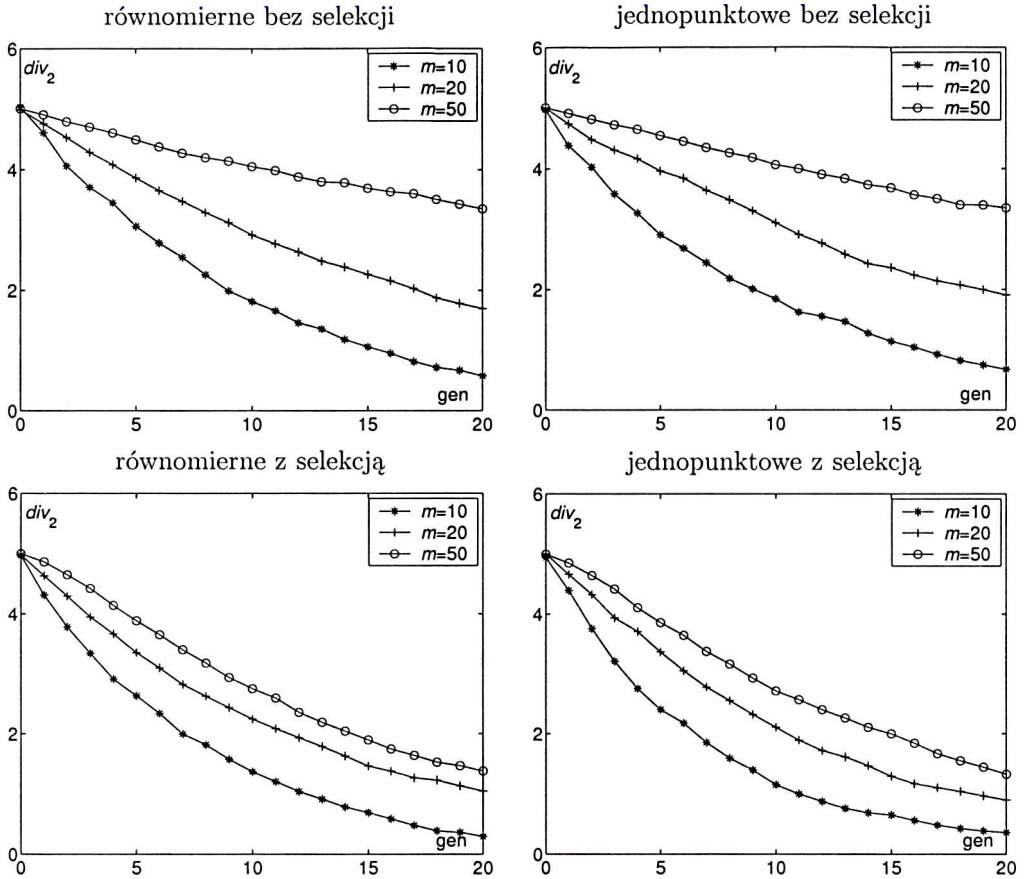
W wielu algorytmach ewolucyjnych operatory rekombinacji traktuje się jako podstawowe, ze względu na ich ważną rolę w konstrukcji korzystnych kombinacji genów, zwiększających przystosowanie populacji [193]. Mówiąc o najbardziej popularnych operatorach rekombinacji (wymieniających fragmenty genów osobników rodzicielskich), podkreśla się podwójny efekt ich działania: dziedziczenie wspólnego materiału genetycznego oraz eksplorację konsekwentnie redukowanej przestrzeni poszukiwań. Korzystne schematy mogą być zachowane i utrwalone przez rekombinację. Operator generuje potomków w tym samym obszarze przestrzeni poszukiwań co populacja bieżąca⁴. W połączeniu z selekcją, obserwuje się zmniejszenie różnorodności populacji. W rezultacie szanse na znalezienie optimum w danym obszarze rosną, ale rosną także szanse na zatrzymanie poszukiwań na tym optimum, co jest zjawiskiem niepożądanym, jeżeli znalezione rozwiązanie nie jest globalnie optymalne. Korzyści ze stosowania rekombinacji w algorytmach genetycznych z kodowaniem binarnym tłumaczy się przez hipotezę bloków tworzących (patrz rozdz. 3.4) [77]. W strategiach ewolucyjnych wyjaśnia się je raczej reperacją genetyczną (ang. *genetic repair*), czyli korekcją błędów zmniejszającą wpływ niekorzystnych mutacji [12].

Zdania badaczy o wpływie operatorów rekombinacji na rozproszenie populacji są podzielone. Część badaczy, mylnie interpretując idee ewolucji naturalnej⁵ uważa, że operacja rekombinacji zwiększa zróżnicowanie populacji poprzez „tasowanie” genów typów rodzicielskich [74, 164]. Inni dowodzą, że rola rekombinacji, zarówno w efektywnym znajdowaniu optimum jak i w utrzymywaniu różnorodności populacji, jest „przereklamowana”. Rola rekombinacji w podtrzymywaniu zróżnicowania zależy głównie od rodzaju stosowanego operatora [3, 18]. Krzyżowanie wymieniające, składające typy potomne z genów rodziców na przykład jedno- i wielopunktowe, nie zwiększa różnorodności populacji. Wniosek ten potwierdza analiza teoretyczna metod ewolucyjnych z wykorzystaniem łańcuchów Markowa [129]. Krzyżowanie jednorodne czy krzyżowanie znane pod anglojęzyczną nazwą jako *disrespectful*⁶ mogą zwiększać różnorodność populacji. Także w przestrzeniach rzeczywistych krzyżowanie równomierne gra ważną rolę w utrzymywaniu różnorod-

⁴ Żadna z wartości genu nie znika z populacji ani nie pojawia się nowa wartość, następuje jedynie ich „przetasowanie” [3].

⁵ Uwagi dotyczące roli rekombinacji w populacjach biologicznych zamieszczono w pracy [117].

⁶ Operator eliminujący dziedziczenie podobieństwa między rodzicami, które występują w innych wersjach operatora krzyżowania. Kreuje się równocześnie dwóch potomków, którzy dziedziczą cechy odpowiednio po rodzicach, jeśli allele na danej pozycji u rodziców są różne. Jeżeli rodzice mają identyczne geny na danej pozycji, to wartości cechy u potomków są wybierane



Rys. 6.5. Średnia różnorodność populacji div_2 (6.2) w 20 generacjach dla kodowania binarnego i różnych rekombinacji, $n = 10$
 Fig. 6.5. The average diversity div_2 (6.2) of population in 20 generations for binary coded vectors and different recombinations, $n = 10$

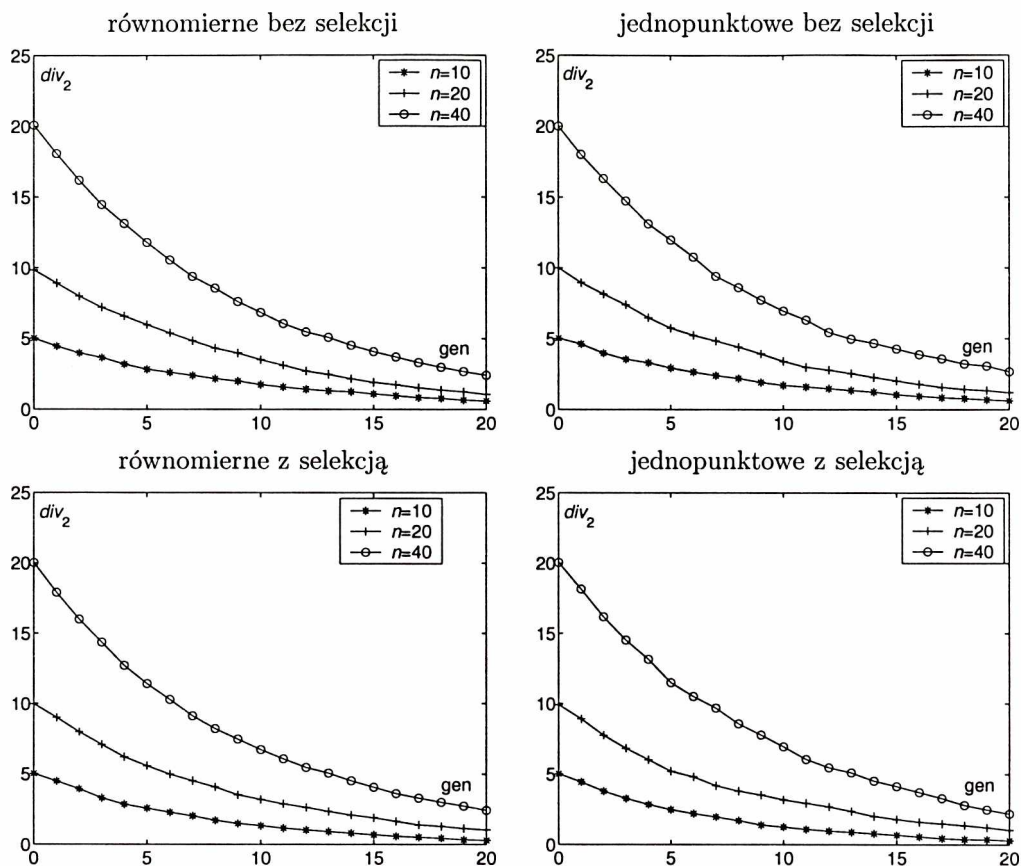
ności populacji [164]. Krzyżowanie uśredniające, gdzie typy potomków leżą na odcinku łączącym pary typów rodzicielskich lub we wnętrzu hiperprostokądnianu generowanego przez typy rodzicielskie, powoduje zmniejszenie różnorodności populacji. Generalnie, coraz częściej dominuje pogląd, że to jednak mutacja zwiększa różnorodność, a rekombinacja ją zmniejsza lub jedynie podtrzymuje [89, 160, 189]. Badania z biologii populacyjnej i symulowanej ewolucji pokazują, że w krajobrazie pojedynczego wzgórza adaptacyjnego rekombinacja powoduje większe skupienie populacji, jednocześnie podwyższając jej średnie przystosowanie [17, 96, 154].

losowo. Operator jest bardzo efektywny i utrzymuje wysoki poziom różnorodności w populacji [206].

W celu zbadania wpływu operatora rekombinacji na różnorodność populacji przeprowadzono badania symulacyjne dla reprezentacji binarnej oraz rzeczywistoliczbowej [118]. Populacje początkowe wybierano losowo z n wymiarowych hipersześcianów: $x_i \in \{0, 1\}^n$ dla wektorów binarnych oraz $x_i \in [-1; 1]$, $i = 1, 2, \dots, n$ dla wektorów rzeczywistoliczbowych. Przebadano najpopularniejsze dwa typy operatorów krzyżowania: jednopunktowe i równomierne. Eksperymenty przeprowadzono z selekcją proporcjonalną (wybrane w procesie selekcji osobniki były łączone w pary i krzyżowane) lub bez selekcji (wszystkie osobniki bieżącej populacji były krzyżowane). Jak w badaniach poprzednich, dla rzeczywistej przestrzeni poszukiwań przyjęto jednomodalną funkcję przystosowania $q(x) = \exp(-5x^2)$. Dla kodowania binarnego wykorzystano funkcję *one-max* (D.7), zliczającą jedynki w kodzie binarnym. Wartości średnie obliczano z 50 uruchomień algorytmu działającego przez 20 generacji.

Na rysunkach 6.5 i 6.6 przedstawiono średnie rozproszenie populacji div_2 (6.2) dla kodowania binarnego i funkcji celu (D.7). Różnorodność populacji maleje dla obu typów krzyżowania, niezależnie od wielkości populacji (rys. 6.5). Im mniejsze populacje, tym szybciej zmniejsza się ich różnorodność i jej poziom. Także długość wektora binarnego wpływa na poziom rozproszenia populacji: im krótsze wektory, tym szybsze zmniejszenie różnorodności (rys. 6.6). Oba rodzaje krzyżowania generują podobne wyniki. Selekcja dodatkowo przyspiesza utratę różnorodności. Na rysunku 6.7 przedstawiono porównanie rozproszenia populacji dla algorytmów działających z krzyżowaniem lub z mutacją. Wykorzystanie mutacji powoduje podtrzymywanie różnorodności, szczególnie dla małych populacji. Oczywiście, im więcej bitów zostało zmutowanych, tym różnorodność populacji była większa. W dużych populacjach i dla długich wektorów binarnych krzyżowanie równomierne powoduje podobne rozproszenie jak mutacja dla jednego genu, a zwiększenie różnorodności przez mutację jest widoczne dopiero dla większej liczby zmutowanych genów.

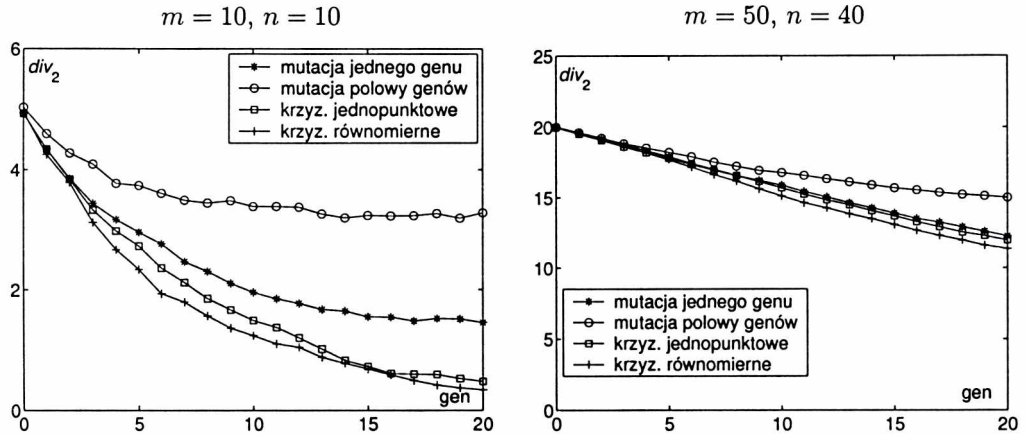
Rezultaty badań dla wektorów rzeczywistoliczbowych przedstawiono na rysunkach 6.8, 6.9, 6.10. Rozproszenie obliczone według miary div_1 (6.1) po każdej operacji rekombinacji przedstawiono dla różnych rozmiarów populacji na rys. 6.8, a obliczone zgodnie z wzorem div_2 (6.2) na rys. 6.9. Symulacje przeprowadzono włączając lub wyłączając selekcję. Aby sprawdzić, czy różnorodność po rekombinacji wzrasta, obliczano także wartość rozproszenia bezpośrednio po selekcji (ze względu na czytelność, nie uwzględniono tych danych na rysunkach). Selekcja zmniejsza różnorodność populacji, a wzrost rozproszenia po rekombinacji był nieznaczny. Zachodził głównie dla większych populacji i to tylko w kilku początkowych generacjach. Gdy rozważamy drugą miarę różnorodności div_2 (6.2) okazuje się, że rekombinacja może nieznacznie zwiększyć rozproszenie populacji (rys. 6.9). Jednakże wzrost ten zmniejsza się w kolejnych generacjach, wraz z ujednolicaniem populacji. Podobnie jak dla kodowania binarnego, mniejsza redukcja różnorodności jest obserwowana



Rys. 6.6. Wpływ długości wektora binarnego n na średnią różnorodność populacji div_2 (6.2) w 20 generacjach dla kodowania binarnego i różnych rekombinacji, $m = 10$
 Fig. 6.6. The influence of the binary vector length n on an average population diversity div_2 (6.2) in 20 generations for binary coding and various recombinations, $m = 10$

dla większych populacji. Z porównania rezultatów wynika, że głównie selekcja jest odpowiedzialna za zmniejszenie różnorodności.

Na rysunku 6.10 przedstawiono wpływ długości wektora n (wymiarowości przestrzeni poszukiwań) na różnorodność populacji. Dla długich wektorów proces utraty różnorodności jest dłuższy, podobnie jak dla kodowania binarnego. Gdy selekcja była włączona, wpływ n był praktycznie niewidoczny (wszystkie linie się pokrywały). Dodanie do algorytmu mutacji (rys. 6.11) powoduje, że choć różnorodność początkowo spada, to począwszy od pewnego poziomu stabilizuje się (lub nawet nieznacznie wzrasta). W symulacjach sprawdzano wartości rozproszenia po rekombinacji i po mutacji. Okazało się, że obniżeniu różnorodności bezpośrednio po rekombinacji odpowiada jej wzrost po mutacji.



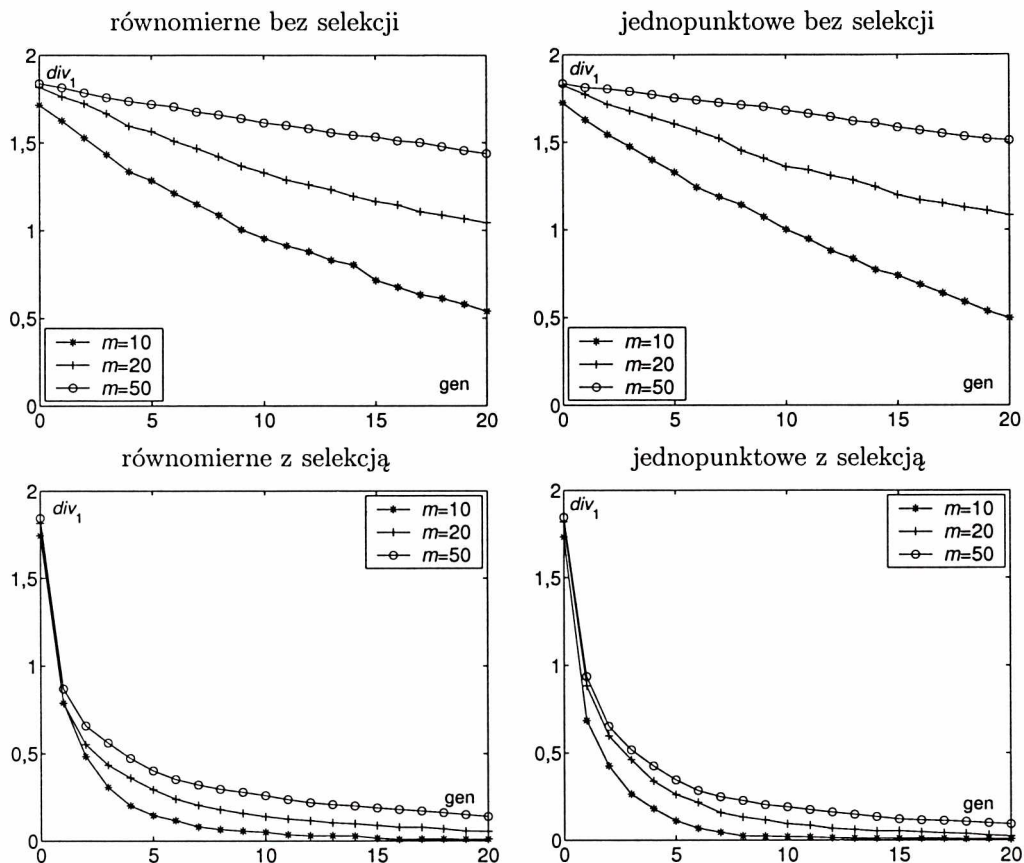
Rys. 6.7. Średnia różnorodność populacji div_2 (6.2) dla kodowania binarnego z selekcją po krzyżowaniu lub mutacji

Fig. 6.7. The average diversity div_2 (6.2) of population for binary coded vectors with selection after crossover or mutation

Przedstawione wyniki symulacji pokazują, że generalnie krzyżowanie (nawet równomierne) nie zwiększa różnorodności populacji. Jedynie dla wektorów rzeczywistych i miary div_2 (6.2) można zaobserwować nieznaczny wzrost rozproszenia. W przypadku kodowania binarnego różnorodność populacji silnie zależy od jej rozmiarów i długości typu. Najmniejszy spadek różnorodności odnotowano dla dużych populacji opisywanych długimi wektorami binarnymi. Ponieważ większość badaczy wykorzystujących algorytmy ewolucyjne z kodowaniem binarnym stosuje duże populacje i długie wektory binarne, to może być przyczyną popularności przekonania o korzystnym działaniu rekombinacji na rozproszenie populacji. Dla wektorów rzeczywistych zależności te są znacznie słabsze.

6.1.5. Podsumowanie

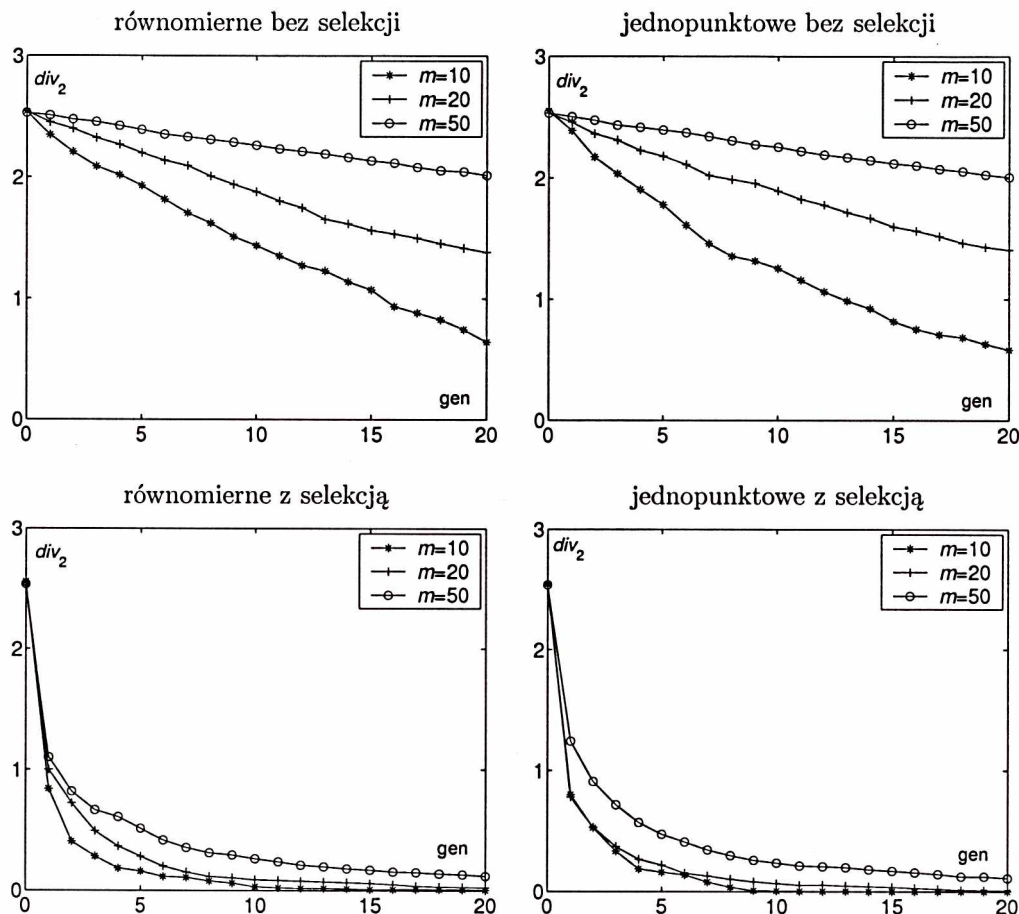
Przedstawione w niniejszym podrozdziale wyniki badań pokazują, że zmniejszanie się początkowo dużej różnorodności populacji, a następnie jej stabilizacja jest specyficzną cechą procesów ewolucyjnych, związaną z ich losowym charakterem oraz działającymi mechanizmami. Utrata różnorodności następuje szybko, przede wszystkim za sprawą operatora selekcji. Intensywność selekcji jest ważnym czynnikiem ograniczającym rozproszenie. Twardsza selekcja powoduje szybszą i ściślejszą koncentrację populacji. Ponieważ pewien poziom różnorodności zapewniają populacji losowe mutacje, więc ewoluje ona jako grupa osobników o zbliżonych typach. Na wielkość klastra wpływa zarówno liczebność populacji, jak i odchylenie standardowe mutacji.



Rys. 6.8. Średnia różnorodność populacji div_1 (6.1) dla kodowania rzeczywistoliczbowego i różnych operatorów krzyżowania, $n = 10$, $\sigma = 0,1$

Fig. 6.8. The average diversity div_1 (6.1) of population for real-valued vectors and various crossover operators, $n = 10$, $\sigma = 0.1$

Zwiększenie wartości tych parametrów powoduje wzrost różnorodności. Prezentowane symulacje wskazują, że rekombinacja generalnie nie zwiększa rozproszenia populacji. Niewielki przyrost różnorodności obserwowano jedynie dla kilku początkowych generacji przy osobnikach opisywanych rzeczywistoliczbowymi cechami fenotypowymi (i to dla jednej tylko miary zróżnicowania). Głównym mechanizmem podtrzymującym lub zwiększającym różnorodność populacji jest oczywiście mutacja. Nie chcąc rezygnować z rekombinacji, bardziej korzystne jest jednocześnie stosowanie operatorów mutacji i rekombinacji (zwłaszcza przy kodowaniu binarnym [30]): tylko krzyżowanie może łączyć pożyteczne

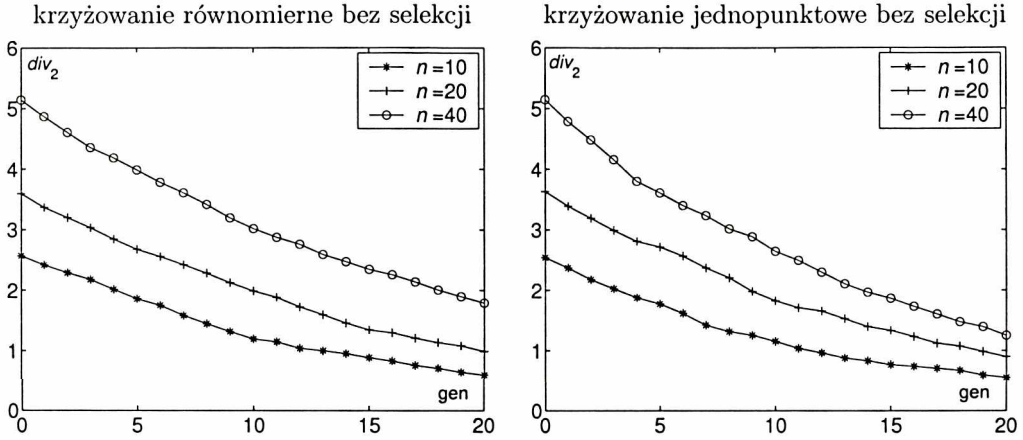


Rys. 6.9. Średnia różnorodność populacji div_2 (6.2) w 20 generacjach dla kodowania rzeczywistoliczbowego i różnych operatorów krzyżowania, $n = 10$

Fig. 6.9. The average diversity div_2 (6.2) of population in 20 generations for real-valued vectors and various crossover operators, $n = 10$

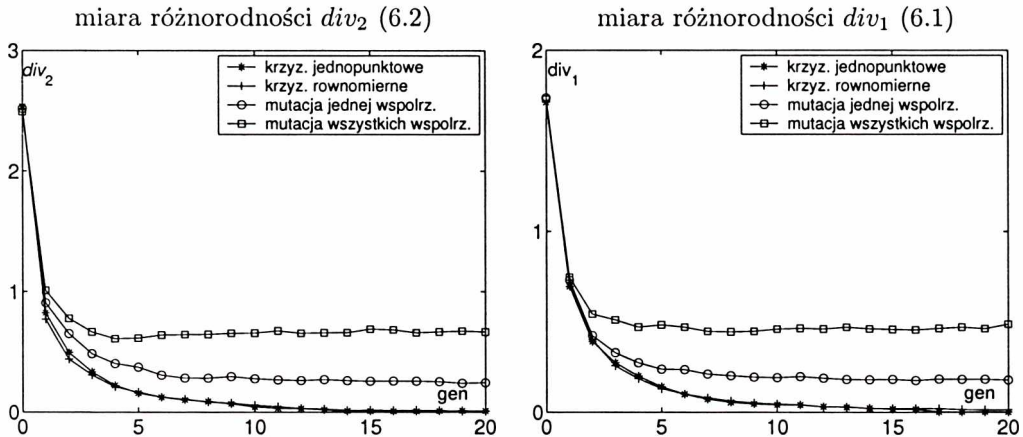
fragmenty z typów rodzicielskich, tylko mutacja może wprowadzić nowe cechy.

W przypadku analizowanego modelu ewolucji fenotypowej, zmniejszenie różnorodności nie jest zjawiskiem niekorzystnym i nie trzeba mu specjalnie zapobiegać. Proces poszukiwań jest tu procesem zależnym od ścieżki, a wielkość klastra nie wpływa zasadniczo na sposób przeszukiwania przestrzeni. Chociaż bardziej skupiona populacja może ewoluować dłużej, to znalezione optimum może być dokładniej zlokalizowane.



Rys. 6.10. Wpływ długości wektora n na średnią różnorodność populacji div_2 (6.2) dla kodowania rzeczywistoliczbowego, $m = 10$

Fig. 6.10. The influence of the vector length n on an average population diversity div_2 (6.2) for real-valued vectors, $m = 10$



Rys. 6.11. Średnia różnorodność populacji w 20 generacjach dla kodowania rzeczywistoliczbowego z selekcją, $m = 10$, $n = 10$, $\sigma = 0,1$

Fig. 6.11. An average diversity of population in 20 generations for real-valued vectors with selection, $m = 10$, $n = 10$, $\sigma = 0.1$

6.2. Określenie parametrów procesu ewolucyjnego i funkcji celu

Przedstawiona w poprzednim rozdziale analiza układu dynamicznego generowanego przez proces ewolucyjny może być pomocna w zastosowaniach metod ewolu-

cyjnych do zadań optymalizacyjnych. Informacje o zachowaniu asymptotycznym układu dynamicznego generowanego przez proces ewolucyjny z miękką selekcją mogą ułatwić wybór najlepszych parametrów algorytmu ewolucyjnego bądź posłużyć do identyfikacji nieznanej funkcji celu. W niniejszym podrozdziale zaproponowano heurystykę doboru wartości odchylenia standardowego mutacji oraz sposób wnioskowania o postaci funkcji przystosowania (w jednowymiarowych przestrzeniach poszukiwań) na podstawie liczby punktów stałych ich stabilności oraz wartości krytycznych σ_c .

6.2.1. Dobór parametrów algorytmu

Dobór optymalnych parametrów metody optymalizacyjnej jest poważnym problemem dla ich użytkowników. Dotyczy to zwłaszcza zastosowań metod ewolucyjnych, w których operuje się zwykle większą liczbą parametrów niż ich mają algorytmy „klasyczne”. Nieodpowiednie wartości parametrów mogą spowodować nieefektywne działanie metody, zwiększenie czasu obliczeń, a także problemy ze znalezieniem rozwiązania optymalnego. Dodatkowo, dobrane na początku obliczeń wartości parametrów mogą się okazać nieodpowiednie na ich końcu (stąd między innymi pomysły na samoadaptację wartości parametrów). Tradycyjna metoda polega zwykle na przeprowadzeniu poprzedzających testów, w których sprawdza się wiele wartości parametrów, zanim zostaną wybrane te właściwe, i dopiero wtedy przeprowadza się rzeczywiste eksperymenty. Często proces doboru parametrów trwa wielokrotnie dłużej niż właściwe badania. Trudno też formułować ogólne zasady doboru parametrów, gdy poszczególne aplikacje różnią się szczegółami implementacyjnymi (kodowaniem, operatorami, kryteriami itp.), chociaż takie próby są oczywiście podejmowane [7, 37, 90, 186].

W poprzednich rozdziałach metodę poszukiwań ewolucyjnych z miękką selekcją analizowano dla jednowymiarowej przestrzeni poszukiwań i ustalonych rozmiarów populacji, tak więc jedynym parametrem metody jest odchylenie standardowe mutacji σ , co znacznie ułatwia rozważania⁷. Analiza oczekiwanego czasu zbieżności prezentowana w rozdziale 5.7 pokazuje, że wartość odchylenia standardowego mutacji nie może być zbyt duża, ponieważ optima będą lokalizowane szybko, lecz niedokładnie lub ujawnią się zachowania niestabilne. Zbyt małe wartości σ zwiększają znacząco koszty obliczeniowe. Jeżeli się weźmie pod uwagę także analizę stanów ustalonych układu dynamicznego (5.16), można zaproponować racjonalną heurystykę dotyczącą doboru parametru σ : *wartość odchylenia standardowego mutacji powinna być od 5 do 10 razy mniejsza niż wartość krytyczna σ_c , dla której układ traci stabilność*. Takie podejście, w którym optymalna wielkość mutacji jest skorelowana

⁷ W ogólnym przypadku dochodzi jeszcze wybór liczności populacji, ale tutaj także wyniki badań teoretycznych i symulacyjnych mogą pomóc we właściwym doborze parametru (patrz rozdział 6.2.3).

z wartością krytyczną (ang. *error threshold*), jest stosowane w niektórych pracach [90, 154].

Badania populacji nieskończonych przedstawione w rozdziale 4 mogą być także pomocne w określeniu optymalnego zasięgu mutacji. Trudno jest podać formułę ilościową, natomiast dokonana analiza potwierdza, że duże wartości σ^2 znacząco przyspieszają zbieżność populacji do optimum, natomiast nie pozwalają na jego dokładną lokalizację. Rozsądny kompromis pomiędzy szybkością a dokładnością poszukiwań zachodzi wtedy, gdy wartości wariancji mutacji są od kilku do kilkunastu razy mniejsze od odległości pomiędzy optimumami funkcji dwumodalnych (zobacz rys. 4.4). Taka zasada jest oczywiście możliwa do wykorzystania tylko wtedy, gdy posiadamy wiedzę o funkcji celu.

6.2.2. Identyfikacja funkcji celu

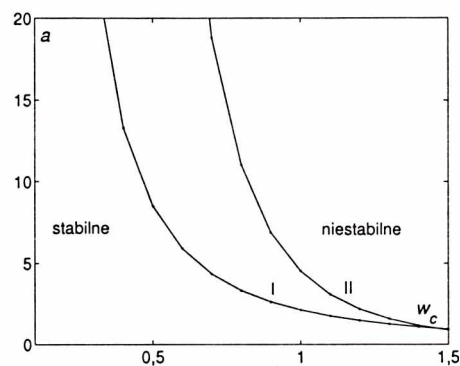
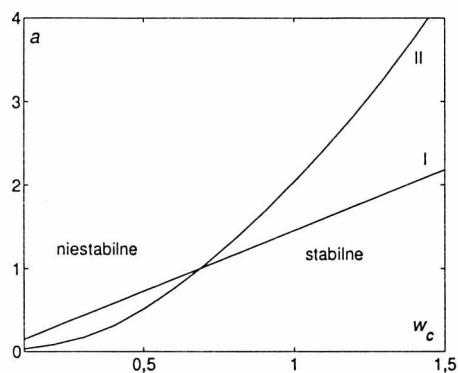
W wielu rzeczywistych problemach optymalizacyjnych funkcja celu nie jest znana *explicite* w postaci wzoru matematycznego, lecz jej wartość uzyskiwana jest jako odczyt pomiaru lub oszacowanie, czy jako wynik symulacji komputerowych (tzw. problemy *black-box*). Jako metody wymagające jedynie wartości funkcji celu, a nie jej postaci analitycznej, algorytmy ewolucyjne są często wykorzystywane do rozwiązywania tego typu zadań. Analiza zachowań niestabilnych układu dynamicznego przedstawiona w rozdziale 5.6 może pomóc w identyfikacji nieznanego jedno-wymiarowej funkcji przystosowania. Ponieważ liczba punktów stałych jest ściśle związana z modalnością funkcji (funkcja o k optimumach ma do $2k + 1$ punktów stałych), może ona posłużyć do oszacowania liczby optimumów. Zachowania chaotyczne dla dużych wartości σ wskazują na asymetryczność funkcji celu (różne nachylenia zboczy funkcji jednomodalnych lub różne szerokości wzgórz funkcji dwumodalnych).

W przypadku funkcji jednomodalnych (trójkątnej (D.5) czy gaussowskiej (D.3)), dla określenia wielkości nachylenia zboczy A lub a można skorzystać z wyników analitycznych bądź wykresów. Punkt stały traci stabilność dla $w_c = 2/3A$ – funkcja trójkątna lub $w_c = \sqrt{2/a}$ – funkcja gaussowska. Znając wartość parametru w_c wielkości nachyleń obliczymy jako: $A = 3/2w_c$ lub $a = 2/w_c^2$ odpowiednio dla obu funkcji. Parametr funkcji przystosowania można także odczytać z diagramów, pokazujących wartości krytyczne σ_c dla różnych parametrów nachylenia (rys. 6.12).

W przypadku gaussowskich funkcji dwumodalnych o wzgórzach różnej szerokości punkty stałe związane z tymi optimumami tracą stabilność dla $w_c \simeq \sqrt{2/a}$, gdzie a jest nachyleniem danego wzgórza. Dla wzgórz o jednakowej szerokości punkt stały związany z optimum globalnym traci stabilność dla $w_c \simeq 2\sqrt{2/a}$, stąd można oszacować odpowiednie wartości a . Im większa różnica w wysokości wzgórz równej szerokości, tym szybciej zanika punkt stały związany z optimum lokalnym (zobacz rys. 5.22, 5.23). Punkty stałe zanikają, gdy $\sigma_d > 0,73d$, gdzie d jest odle-

a) funkcje trójkątne (D.5) ($B = 1,0$)

b) funkcje gaussowskie (D.3)



Rys. 6.12. Zakresy parametrów A , a oraz wartości w_c , dla których jednodymodalnej funkcje celu są stabilne. Funkcje: I – symetryczne, II – asymetryczne
 Fig. 6.12. The ranges of the parameters A , a and the values w_c , for which the unimodal fitness functions are stable. Functions: symmetrical (I), asymmetrical (II)

głością między optimami funkcji gaussowskich lub szerokością lokalnego wzniesienia ($A + B$) funkcji trójkątnej. Wtedy funkcje $q_1(z)$ i $q_2(z)$ (zobacz rys. 5.8) rozsuwają się i zmniejsza się liczba miejsc przecięcia, a więc i liczba punktów stałych. Znając zatem wartość odchylenia standardowego mutacji, dla którego pozostaje tylko jeden punkt stały, można ocenić odległość między optimami dla funkcji gaussowskich lub szerokość wzniesienia funkcji trójkątnej. W przypadku funkcji trójkątnej wartość σ_{d1} , dla której globalny punkt stały zanika, pozwala na oszacowanie szerokości całej funkcji, ponieważ $(A + C) \simeq 1,37\sigma_{d1}$. Bardziej skomplikowane zachowania układu dynamicznego: utrata stabilności punktu stałego związanego z optimum globalnym dla stosunkowo niewielkich wartości σ , ponowne odzyskiwanie stabilności dla pewnego zakresu σ , orbity o różnych okresach wskazują na duże różnice w szerokościach wzniesień i mogą świadczyć o tym, że wzniesienie globalne jest wąskie i jego wpływ na populację jest nikły.

Prezentowane podejście dotyczy jednowymiarowych i specyficznych funkcji celu. Oczekuje się jednak, że można będzie je zastosować dla innych funkcji celu, ponieważ wiele funkcji można aproksymować wykorzystując liniowe segmenty funkcji trójkątnej lub zbocza funkcji gaussowskich.

6.2.3. Uwagi o liczebności populacji

Ponieważ wielkość populacji jest ważnym parametrem algorytmów ewolucyjnych [46, 64, 129, 188], w podrozdziale tym przedstawimy kilka spostrzeżeń uży-

skanych w trakcie badań nad metodami fenotypowymi, a które nie wynikają bezpośrednio z prezentowanej analizy teoretycznej.

Ustalanie wielkości populacji jest w większości metod ewolucyjnych podyktowane równowagą między efektywnością poszukiwań a wydajnością metody. W literaturze można znaleźć aplikacje metod ewolucyjnych, w których wielkość populacji waha się od kilku do kilku tysięcy osobników. Różna liczebność populacji przyczynia się niekiedy do błędnych interpretacji wyników eksperymentalnych, gdy porównuje się liczby generacji potrzebnych do znalezienia optimum zamiast liczby ewaluacji funkcji przystosowania, która jest właściwą miarą złożoności metody. Panuje przekonanie, że zbyt małe populacje nie będą mogły skutecznie przeszukiwać przestrzeni poszukiwań, ponieważ zachowują się zbyt chaotycznie, a duże populacje mogą być mało wydajne obliczeniowo.

W algorytmach genetycznych z kodowaniem binarnym proponowano powiązać liniowo wielkość populacji z długością wektora binarnego l i zalecano przyjmować rozmiar populacji pomiędzy l i $2l$ [76]. Teoretyczną analizę wielkości populacji na podstawie twierdzenia o schematach i hipotezy bloków tworzących przedstawiono w [81, 86]⁸. Okazało się jednak, że wielkość populacji powinna wzrastać eksponencjalnie z długością wektora binarnego. Reeves [173] sugeruje, że populacja powinna być tak liczna, aby każdy punkt w przestrzeni poszukiwań mógł być osiągnięty z populacji początkowej, wykorzystując tylko operator rekombinacji. Warunek ten może być spełniony, gdy przynajmniej jeden egzemplarz każdego allelu w każdym lokusie jest obecny w całej początkowej populacji typów. Można wtedy podać wzór określający szacunkową wielkość populacji dla kodowania binarnego [175]. Wydaje się, że spełnienie tego warunku powoduje stosowanie bardzo licznych populacji w AG i podwyższa efektywność rekombinacji w niektórych zadaniach. Teoria Reevesa została rozszerzona na wektory z alfabetu o mocy $q > 2$. Pokazano, że wtedy populacje powinny być liczniejsze niż w przypadku binarnym. Mniej ostrożne oszacowanie wielkości populacji otrzymano za pomocą modelu ruiny gracza z teorii gier [86]. Badania empiryczne AG wykazały, że populacje wielkości 30 elementów są w wielu przypadkach wystarczające [37, 186], inni sugerują, że powinny być 50-elementowe [175], a dla AG z kodowaniem rzeczywistoliczbowym mniejsze niż 16-elementowe [88].

Optymalna wielkość populacji jest kompromisem między eksploracją i eksploatacją. W dużych populacjach, tworzących zwarty klaster, jest więcej zbliżonych typów, a więc i lepsza eksploatacja lokalna krajobrazu adaptacyjnego [64, 96]. W małych populacjach istnieje znacznie większa szansa, że najlepszy osobnik się nie zreprodukuje, średnia jakość populacji jest niższa, a obszar poszukiwań rozlegle eksploatowany. Badania symulacyjne modelu ewolucji fenotypowej z miękką selekcją pokazały jednak, że najefektywniejsze w sprawnym przekraczaniu siodeł między optimami i znajdowaniu optimum globalnego są małe populacje [29, 62, 64, 69].

⁸ Należy pamiętać, że twierdzenia te opracowano przy założeniu nieskończonych populacji.

Małe populacje sprawnie przekraczają nawet bardzo szerokie siodła w wielowymiarowych przestrzeniach poszukiwań także wtedy, gdy presja selekcyjna jest silniejsza. Przykładowo, większe populacje ewoluujące z selekcją elitarną potrafią przekroczyć siodła o szerokości około 5σ , podczas gdy dwuelementowa populacja szersze niż 10σ nawet dla 10-wymiarowej przestrzeni poszukiwań [29]. Okazuje się, że małe populacje są także efektywniejsze, gdy rozpatrujemy koszt przekraczania siodła (czyli jak długo należy czekać, aby zwróciły się nakłady poniesione na przejście siodła, liczone w kategoriach obniżonej jakości, które to przejście powoduje) [72]. Nakłady potrzebne na przejście siodła dla małych populacji są mniejsze i szybciej mogą być zwrócone niż w przypadku populacji liczniejszych. Znaczne zwiększenie szybkości obliczeń i efektywności poszukiwań optimów globalnych po zastosowaniu bardzo małych populacji zaobserwowano również dla programowania genetycznego, gdzie zjawisko to nazwano nawet „efektem małych populacji” [5].

Populacje duże są efektywniejsze w optymalizacji lokalnej, gdyż potrafią bardziej precyzyjnie zidentyfikować optimum, są mniej czułe na zakłócenia funkcji celu, a także czynią algorytm w pewnym sensie odpornym na rodzaj selekcji, zwiększając prawdopodobieństwo umieszczenia choćby jednego osobnika w obszarze przyciągania optimum globalnego [4]⁹.

⁹ Stosowanie algorytmów ewolucyjnych w zadaniach optymalizacji lokalnej jest zasadne tylko wtedy, gdy nie można zastosować efektywniejszych metod klasycznych (np. funkcja celu typu *black box*).

7. Podsumowanie

Efektywność algorytmów ewolucyjnych w rozwiązywaniu trudnych problemów optymalizacyjnych wykazano wieloma udanymi aplikacjami, a także badaniami empirycznymi. Zastosowania praktyczne nie korespondują jednak z pełnym zrozumieniem sposobu działania metod ewolucyjnych. Można w tym miejscu zacytować wypowiedź jednego z ekspertów: *Poziom zrozumienia istoty tych metod jest niski. Czasami dziedzina jest traktowana jak „czarna magia”, a nie nauka. Każdy wynik teoretyczny, uogólnienie, które pozwala doprowadzić do lepszego poznania, jest cenny*¹.

W niniejszej pracy podjęto próbę analizy teoretycznej prostej metody fenotypowej, bazującej na dwóch mechanizmach ewolucyjnych: selekcji proporcjonalnej i mutacji gaussowskiej. W obliczu trudności formalnych, jakie powstają podczas analizy metod ewolucyjnych, rozpoczęcie badań od najprostszych modeli wydaje się metodologicznie uzasadnione. Nawet dla tak prostego algorytmu dokonano dodatkowych założeń umożliwiających opracowanie jego modelu, który poddano analizie teoretycznej. Rozpatrywano mianowicie nieskończone oraz dwuelementowe populacje w jednowymiarowej przestrzeni poszukiwań.

Populacje nieskończone opisano w terminach rozkładów prób w następnej generacji. Pokazano, że kolejne rozkłady są rozkładami normalnymi lub ich mieszaninami, w zależności od postaci funkcji celu. Udowodniono, że rozkłady nieskończonych populacji w krajobrazie jednomodalnej funkcji przystosowania są zbieżne do lokalnego optimum według średniej oraz z prawdopodobieństwem. Dla multimodalnych funkcji celu rozkłady, stanowiące mieszaniny rozkładów normalnych, koncentrują się na optimach, zwiększając wraz z wpływem kolejnych generacji „masę” rozkładu składowego wokół optimum globalnego, a zmniejszając składowe wokół optimów lokalnych. Przedstawiono sposób generacji parametrów mieszaniny na podstawie wartości parametrów (średniej i wariancji) rozkładu początkowego. Badaniami symulacyjnymi sprawdzano wpływ różnych postaci i parametrów funkcji przystosowania, wariancji mutacji oraz rozkładów początkowych na rozkłady nieskończonych populacji. Nieskończone populacje sprawnie lokalizują optima globalne nawet wtedy, gdy nachylenia zboczy są ostre, siodła bardzo głębokie, a różnice w wysokości pomiędzy optimami znaczne. Co ciekawsze, lokalizacja ekstremum dla tak, wydawałoby się, trudnych przypadków jest szybsza i dokładniejsza niż dla funkcji o łagodnych zboczach i płytkich siodłach. Symulacje pokazały, że przyjęcie

¹ Z. Michalewicz, dyskusja podczas publicznej obrony pracy doktorskiej, Politechnika Wrocławska, 6.06.2007 r.

do analizy rozkładów normalnych i funkcji gaussowskich nie ogranicza ogólności wyników. Zarówno inne rozkłady mutacji (jednostajny, trójkątny), różne rozkłady populacji początkowej (prawie homogeniczna), jak i niegaussowskie funkcje celu nie miały znaczącego wpływu na rozkłady nieskończonych populacji, które pozostały mieszaniną rozkładów gaussowskich. Pokazano także właściwości fraktalne analizowanego modelu.

Analiza populacji dwuelementowych była możliwa dzięki rozpatrywaniu populacji w przestrzeni jej stanów a nie typów. Przyjęcie relacji równoważności, porządkującej elementy w populacji, oraz zamiana zmiennych ułatwiła analizę wartości oczekiwanych i ich trajektorii. Badania analityczne wartości oczekiwanych potwierdziły zaobserwowane w symulacjach dwie fazy ewolucji o różnych prędkościach: szybkie tworzenie się klastra elementów o zbliżonych typach i wolny dryf populacji ku optimum. Początkowo znacznie rozproszone populacje skupiają się, a jednorodnie zwiększają swą różnorodność do poziomu odchylenia standardowego mutacji (różnice w wartościach cech osobników są rzędu σ). W drugiej fazie rozproszenie populacji waha się już tylko nieznacznie, a skoncentrowana populacja przesuwa się wolno w kierunku optimum. Analiza wartości oczekiwanej opisującej różnorodność populacji pozwoliła na określenie jej granic zmienności. Nie może być ona mniejsza niż $\sqrt{2/\pi}\sigma$ i jest także ograniczona od góry.

Wartości oczekiwane współrzędnych stanu populacji generują dyskretny, dwuwymiarowy układ dynamiczny o współrzędnych (w, z) , dla którego określono własności asymptotyczne. Od postaci funkcji przystosowania zależy liczba punktów stałych układu, wartości współrzędnej z tych punktów, a także ich stabilność. Wartości współrzędnej w punktów stałych zależą jedynie od odchylenia standardowego mutacji, co potwierdza obserwacje, że proces ujednolicania populacji zachodzi w każdym krajobrazie adaptacyjnym. Stabilność punktów stałych zależy od drugiej wartości własnej macierzy aproksymacji liniowej, która związana jest z funkcją celu. Udowodniono, że punkty stałe leżące w pobliżu siodła są niestabilne. Szczegółową analizę punktów stałych i ich stabilności przeprowadzono dla jednomodalnych i dwumodalnych funkcji przystosowania, w wersji symetrycznej i asymetrycznej. Dla funkcji jednomodalnych, gaussowskich i trójkątnych podano wzory analityczne opisujące współrzędną z punktu stałego, drugą wartość własną oraz wartość krytyczną odchylenia standardowego mutacji, powyżej której układ staje się niestabilny. W przypadku funkcji dwumodalnych przedstawienie postaci analitycznej współrzędnej z punktu stałego możliwe jest tylko dla funkcji trójkątnej. Analiza obszaru przyciągania punktów stałych dla wielomodalnych funkcji przystosowania pokazała, że położenie i kształt tych zbiorów zależą od obszarów wpływów poszczególnych optimów oraz postaci funkcji celu. Potwierdziły się wcześniejsze obserwacje, że populacja jest przyciągana przez grzbiety funkcji przystosowania i ewoluuje wzdłuż nich. Punkty stałe układu są stabilne dla małych wartości odchylenia standardowego mutacji. Przy zwiększających się wartościach σ punkty

tracą stabilność, obserwuje się bifurkacje i orbity o różnych okresach. Dla funkcji asymetrycznych (zarówno jedno- jak i dwumodalnych) w układzie pojawiają się bifurkacje podwajania okresu prowadzące do chaosu. Zachowania chaotyczne studiowano, korzystając ze znanych z teorii układów dynamicznych metod: wykresów fazowych, widm mocy, funkcji autokorelacji oraz wykładników Lapunowa. Wszystkie metody potwierdziły występowanie w układzie dynamicznym generowanym przez wartości oczekiwane stanów populacji orbit o różnych okresach oraz chaosu.

Przeprowadzono badania symulacyjne oczekiwanego czasu zbieżności do punktów stałych. Jak można się było spodziewać, najdłużej stabilizują się populacje, których stan początkowy leży w obszarach małych wartości średniego przystosowania, a najszybciej – populacje znajdujące się w pobliżu optimum lub grzbietów odpowiadających stanom, w których jeden z osobników ma dużą jakość. Szybkość zbiegania populacji do stanów w pobliżu atraktora skoku jest duża i nieznacznie zależna od stanu początkowego. Otrzymane rezultaty ponownie potwierdzają tezę o przebiegu ewolucji w dwóch fazach o zróżnicowanych prędkościach. Badania wpływu odchylenia standardowego mutacji na oczekiwany czas zbieżności do punktów stałych mogą być pomocne w wyborze odpowiednich wartości odchylenia standardowego mutacji w rzeczywistych zastosowaniach.

Analizę dwuelementowych populacji uogólniono, przedstawiając wyniki symulacji populacji kilkuelementowych. Także i w tym przypadku zaobserwowano zachowania charakterystyczne dla ewolucji populacji dwuelementowych: skok populacji do kanału ewolucyjnego oznaczający ujednolicanie się populacji i wolne fluktuacje wzdłuż osi identyczności w kierunku optimum. Liczniejsze populacje są bardziej zwarte, a stany zajmowane przez populację są ulokowane dalej od optimum niż dla populacji mniejszych.

Rezultaty badań dynamiki nieskończonych i dwuelementowych populacji wykorzystano do określenia innych aspektów ewolucji: różnorodności populacji oraz określania nieznanymi parametrów procesu ewolucyjnego i funkcji przystosowania. Analiza wartości oczekiwanej współrzędnej w pokazała, że mechanizm szybkiego ujednolicania się rozproszonej populacji jest immanentną cechą procesu ewolucji, zależną jedynie od parametru operacji mutacji. Przedstawiono badania symulacyjne wpływu operatorów selekcji, mutacji i krzyżowania na rozproszenie populacji. Jak oczekiwano, wszystkie rozpatrywane metody selekcji (nawet czysto losowa) powodowały zmniejszenie się różnorodności populacji. Stopień ujednolicania populacji zależy od „miękości” selekcji i był większy, a także szybszy, dla selekcji „twardych”, o silniejszej presji selekcyjnej. Proces unifikacji zachodzi szybciej dla populacji mniejszych i większej wymiarowości przestrzeni poszukiwań. Badania symulacyjne wpływu operatorów mutacji i rekombinacji pokazują, że krzyżowanie nie zwiększa w sposób istotny różnorodności, a głównym mechanizmem podtrzymu-

jącym lub zwiększającym różnorodność w populacji jest mutacja. Jeżeli stosuje się metody ewolucyjne z kodowaniem binarnym, wskazane jest użycie obu operatorów.

Korzystając z wyników analizy oczekiwanego czasu zbieżności oraz stanów ustalonych, zaproponowano heurystykę doboru wartości parametru σ . Pokazano jak informacje o liczbie punktów stałych, ich stabilności i wartościach krytycznych odchylenia standardowego mutacji pozwalają na wnioskowanie o postaci nieznannej jednowymiarowej funkcji celu. Przeprowadzone w ramach niniejszej monografii oraz wcześniejsze badania analityczne i symulacyjne metody poszukiwań ewolucyjnych z miękką selekcją pozwoliły na sformułowanie wskazówek dotyczących wyboru kluczowego parametru algorytmów ewolucyjnych – liczebności populacji.

Reasumując, najważniejsze wnioski z przeprowadzonej analizy teoretycznej i badań symulacyjnych fenotypowej metody poszukiwań ewolucyjnych z miękką selekcją są następujące:

- ujednolicanie się rozproszonej populacji jest immanentną cechą procesu ewolucji, niezależną od funkcji przystosowania,
- małe populacje efektywnie przeszukują przestrzeń poszukiwań;
- bardzo ważnym czynnikiem wpływającym na efektywność metody jest wartość odchylenia standardowego mutacji. Zbyt mała wydłuża proces obliczeń pozwalając jednocześnie na dokładniejszą lokalizację optimum, a zbyt duża może prowadzić do zachowań niestabilnych;
- mutacja jest jedynym czynnikiem podtrzymującym różnorodność w populacji;
- krzyżowanie nie jest konieczne w metodach fenotypowych dla poszukiwań optimum globalnych, powoduje jednak lepszą koncentrację populacji na już zlokalizowanym optimum;
- analiza oczekiwanych zachowań populacji może być pomocna przy doborze parametrów oraz identyfikacji (jednowymiarowej) funkcji przystosowania.

Autorka ma nadzieję, że przedstawione w niniejszej dysertacji wyniki analizy teoretycznej i badań symulacyjnych prostej metody ewolucyjnej z miękką selekcją pozwolą na pełniejsze zrozumienie mechanizmów działających w algorytmach ewolucyjnych, a także przyczynią się do efektywniejszego ich działania, choćby tylko przez wybór odpowiednich do zadania operatorów i ich parametrów.

Dodatek. Funkcje celu

W pracy wykorzystano funkcje gaussowskie oraz ich sumy jako przykłady jedno- oraz wielomodalnych funkcji celu.

Suma funkcji gaussowskich

$$q(x) = \sum_i^k h_i \exp(-a_i(x - d_i)^2), \quad h_i, a_i, d_i \in R, a_i > 0. \quad (D.1)$$

Jednomodalna symetryczna funkcja gaussowska, rys. D.1a

$$q(x) = \exp(-ax^2). \quad (D.2)$$

Jednomodalna asymetryczna funkcja gaussowska, rys. D.1a

$$q(x) = \begin{cases} \exp(-a_1x^2) & \text{dla } x \leq 0, \\ \exp(-a_2x^2) & \text{dla } x > 0, \end{cases} \quad (D.3)$$

gdzie parametry a_1, a_2 opisują nachylenia poszczególnych zboczy funkcji.

Dwumodalna funkcja gaussowska, rys. D.1b

$$q(x) = \exp(-a_1x^2) + h \exp(-a_2(x - d)^2). \quad (D.4)$$

Dla funkcji dwumodalnych parametry $\zeta_1^2 = 1/(2a_1)$, $\zeta_2^2 = 1/(2a_2)$ warunkują szerokości wzgórz. Gdy zbocza wzgórz mają jednakowe nachylenia (są o identycznej szerokości), podano tylko jeden parametr $a = a_1 = a_2$. Jako przykłady funkcji asymetrycznych stosowano też funkcje trójkątne.

Jednomodalna funkcja trójkątna, rys. D.1c (symetryczna dla $A = B$)

$$q(x) = \begin{cases} x/A + 1 & \text{dla } x \in [-A, 0), \quad A > 0, \\ -x/B + 1 & \text{dla } x \in [0, B], \quad B \geq A, \\ 0 & \text{poza.} \end{cases} \quad (D.5)$$

Bimodalna funkcja trójkątna, rys. D.1d

$$q = \begin{cases} x/A + 1 & \text{dla } x \in [-A, 0), \quad A > 0, \\ -(1 - h_1)x/B + 1 & \text{dla } x \in [0, B], \quad B \geq A, \\ (h - h_1)(x - B)/(1 - B) + h_1 & \text{dla } x \in [B, 1), \\ -h(x - 1)/(C - 1) + h & \text{dla } x \in [1, C], \\ 0 & \text{poza.} \end{cases} \quad (D.6)$$

Funkcja *one-max* zlicza jedynki w łańcuchu binarnym $\pm x$ o długości l .

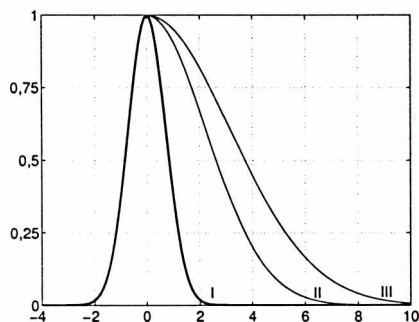
$$q(x) = \sum_{i=1}^l x_i. \quad (D.7)$$

a) funkcje jednomodalne:

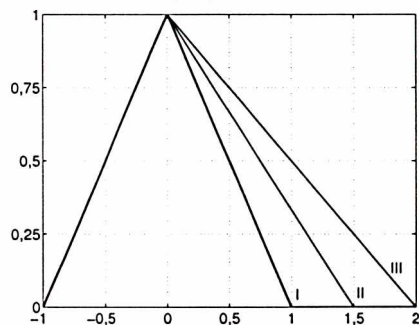
I. (D.2), $a = 1$;

II. (D.3), $a_1 = 1$, $a_2 = 0,1$;

III. (D.3), $a_1 = 1$, $a_2 = 0,05$;



c) funkcje jednomodalne (D.5), $A = 1$,
 $B = 1; 1,5; 2$, I/II/III;

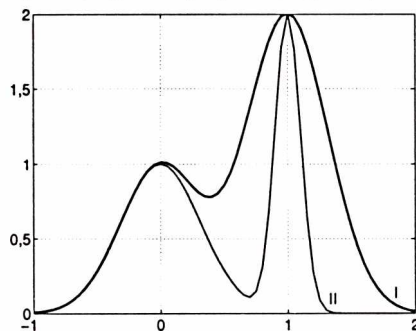


b) funkcje dwumodalne:

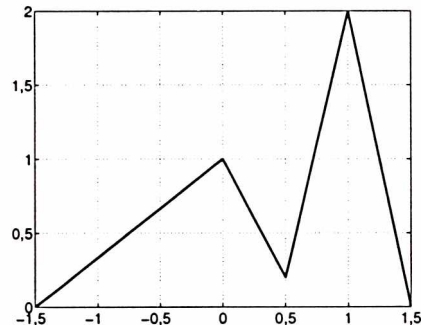
I. (D.4) $h = 2$, $a = 5$, $d = 1$;

II. (D.1) $h_1 = 1$, $h_2 = 2$,

$a_1 = 5$, $a_2 = 50$, $d_1 = 0$, $d_2 = 1$;



d) funkcje bimodalne (D.6), $A = 1,5$,
 $B = 0,5$, $C = 1,5$, $h = 2$, $h_1 = 0,2$



Rys. D.1. Funkcje przystosowania: gaussowskie a) i b); trójkątne c) i d)

Fig. D.1. Fitness functions: Gaussian a) and b); tent c) and d)

Bibliografia

- [1] Altenberg L., *The schema theorem and Price's theorem*, [w:] Proc. Foundations of Genetic Algorithms, pod red. D. Whitley, M. Vose, Morgan Kaufmann Publ., San Francisco 1995, p. 23–49.
- [2] Antonisse J., *A new interpretation of schema notation that overturns the binary encoding constraint*, [w:] Conf. on Evolutionary Computation, Morgan Kaufmann Publ. 1989, p. 86–91.
- [3] Arabas J., Algorytmy ewolucyjne, WNT, Warszawa 2001.
- [4] Arabas J., Metoda ewolucyjna jako narzędzie przeszukiwania przestrzeni Banacha, No. 152 [w:] Prace Naukowe, Oficyna Wyd. Polit. Warszawskiej, Warszawa 2005.
- [5] Ashlock W., *Using very small population sizes in genetic programming*, [w:] IEEE Congress on Evolutionary Computation, 2006, p. 319–326.
- [6] Aytug H., Koehler G., *New stopping criteria for genetic algorithms*, European Journal of Operational Research 2000, Vol. 126, p. 662–674.
- [7] Baeck T., *Optimal mutation rates in genetic search*, [w:] Int. Conf. on Genetic Algorithms, 1993, p. 2–8.
- [8] Baeck T., *An overview of parameter control methods by self-adaptation in evolutionary algorithms*, Fundamenta Informaticae 1998, Vol. 35, No. 1–4, p. 51–66.
- [9] Baeck T., Fogel D., Michalewicz Z., red., Handbook of evolutionary computation, Oxford Univ. Press, NY 1997.
- [10] Barker A., Martin W., *Dynamics of a distance-based population diversity measure*, [w:] Congress on Evolutionary Computation, Vol. 2, IEEE Press 2000, p. 1002–1009.
- [11] Barnsley M., Fractals Everywhere, Academic Press Inc. 1988.
- [12] Beyer H.G., *Toward a theory of evolution strategies: on the benefit of sex*, Evolutionary Computation 1995, Vol. 3, p. 81–111.
- [13] Beyer H.G., *Toward a theory of evolution strategies: Self-adaptation*, Evolutionary Computation 1996, Vol. 3, No. 3, p. 311–347.
- [14] Beyer H.G., The theory of evolution strategies, Springer-Verlag 2001.
- [15] Beyer H.G., Deb K., *On self-adaptive features in real-parameter evolutionary algorithms*, IEEE Trans. Evol. Comp. 2001, Vol. 5, No. 3, p. 250–270.
- [16] Bishop C., Neural networks for pattern recognition, Clarendon Press, Oxford 1995.
- [17] Boerlijst M.C., Bonhoeffer S., Nowak M.A., *Viral quasi-species and recombination*, Proc. Royal Society London B 1996, Vol. 263, p. 1577–1584.
- [18] Booker L., Fogel D., Whitley D., Angeline P., *Recombination*, [w:] Baeck i in. [9], p. C3.3.
- [19] Bose K., Kahng A., Muddu S., *A new adaptive multistart technique for combinatorial global optimizations*, Operations Research Letters 1994, Vol. 16, p. 101–113.
- [20] Branke J., Jin Y., red., *Special issue on evolutionary computation in the presence of uncertainty*, IEEE Trans. Evol. Comp. 2006, Vol. 10, No. 4.

- [21] Bremermann H.J., *Optimization through evolution and recombination*, Spartan Books, Washington DC 1962.
- [22] Bruce D., Simpson R.J., *Evolution determined by trajectory of expected populations: sufficient conditions, with application to crossover*, *Evolutionary Computation* 1999, Vol. 7, No. 2, p. 151–171.
- [23] Bulmer M.G., *The mathematical theory of quantitative genetics*, Clarendon Press, Oxford 1980.
- [24] Chakraborty U.K., Chakraborty M., *Analysis of selection algorithms: a Markov chain approach*, *Evolutionary Computation* 1996, Vol. 4, No. 2, p. 133–167.
- [25] Chorażyczewski A., *Analiza własności adaptacyjnych ewoluujących populacji ze względu na zastosowania w optymalizacji globalnej*, Rozprawa doktorska, Politechnika Wrocławska 2001.
- [26] Chorażyczewski A., Galar R., *The impact of nonheritable variance on simulated evolutionary adaptation*, [w:] *KAEOiG*, Łągów 2003, p. 49–55.
- [27] Chorażyczewski A., Galar R., *Adaptive saddle crossings in diverse landscapes*, [w:] *KAEOiG*, Murzasichle 2006, p. 77–85.
- [28] Chorażyczewski A., Galar R., Karcz-Dulęba I., *Considering phenotypic evolution in the space of population states*, [w:] *Proc. 5th Conf. Neural Networks and Soft Computing*, pod red. R.T. L. Rutkowski, Zakopane, Poland 2000, p. 615–620.
- [29] Chorażyczewski A., Karcz-Dulęba I., *On crossing saddles by small populations*, [w:] *KAEOiG*, Łągów 2003, p. 57–64.
- [30] Culberson J.C., *Mutation-crossover isomorphism and the construction of discriminating functions*, *Evolutionary Computation* 1995, Vol. 2, No. 3, p. 279–311.
- [31] Darwin C., *On the origin of species by means of natural selection or the preservation of favoured races in the struggle for life*, John Murray, London 1859.
- [32] Davis L., red., *Handbook of Genetic Algorithms*, Van Nostrand Reinhold, NY 1991.
- [33] Davis T.E., Principe J.C., *A simulated like convergence theory for SGA*, [w:] *Int. Conf. on Genetic Algorithms*, 1991, p. 174–181.
- [34] Davis T.E., Principe J.C., *A Markov chain framework for the simple genetic algorithm*, *Evolutionary Computation* 1993, Vol. 1, No. 3, p. 269–288.
- [35] De Jong K., *Genetic algorithms are not function optimizers*, [w:] *Proc. Foundations of Genetic Algorithms*, 1992, p. 5–17.
- [36] De Jong K., Spears W., Gordon D., *Using Markov chains to analyze GAFOs*, [w:] *Proc. Foundations of Genetic Algorithms*, pod red. L. Whitley, M. Vose, Vol. 3, Morgan Kaufmann Publ. 1995, p. 115–137.
- [37] De Jong K.A., *An analysis of the behavior of a class of genetic adaptive systems*, Rozprawa doktorska, Univ. of Michigan, Ann Arbor, MI 1975.
- [38] Del Moral P., Miclo L., Patras F., Rubenthaler S., *The convergence to equilibrium of neutral genetic models*, *ArXiv Mathematics e-prints* 2007.
- [39] Devaney R.L., *An introduction to chaotic dynamical systems*, Addison-Wesley, Reading, M. A. 1989.
- [40] Droste S., Jansen T., Wegener I., *Perhaps not a free lunch but at least a free appetizer*, [w:] *Genetic and Evolutionary Computation Conf.*, Orlando, FL 1999, p. 833–839.
- [41] Dua R., Hart P., *Pattern classification and scene analysis*, Wiley 1973.

- [42] Duleba I., Karcz-Duleba I., *Zastosowania algorytmów genetycznych w robotyce*, [w:] *IV Krajowa Konferencja Robotyki*, 1993, p. 243–250.
- [43] Duleba I., Karcz-Duleba I., *Trajectory planning of robots: A GA approach*, [w:] *Cybernetics and Systems. Proc. 12 Europ. Meeting on Cybern. and Syst. Res.*, pod red. R. Trappl, Vol. 2, Vienna 1994, p. 1467–1474.
- [44] Duleba I., Karcz-Duleba I., *Analiza dyskretnego układu dynamicznego generowanego przez pewien proces ewolucyjny*, [w:] *Proc. 9 Symp. Symulacja Procesów Dynamicznych*, Zakopane 1996, p. 351–356.
- [45] Duleba I., Karcz-Duleba I., *Algorithms of trajectory planning with constrained deviation from a given end-effector path*, *Robotica* 2004, Vol. 22, No. 6, p. 633–642.
- [46] Eiben A.E., Hinterding R., Michalewicz Z., *Parameter control in evolutionary algorithms*, *IEEE Trans. Evol. Comp.* 1999, Vol. 3, No. 2, p. 124–141.
- [47] Eigen M., *Self-organization of matter and the evolution of biological macromolecules*, *Naturwissenschaften* 1971, Vol. 58, p. 465–523.
- [48] Eigen M., *Macromolecular evolution: dynamical ordering in sequence space*, [w:] *Emerging Synthesis in Science: Proc. Founding Workshops of the Santa Fe Institut*, pod red. D. Pines, Addison Wesley, Reading MA 1985, p. 25–69.
- [49] Eigen M., Schuster P., *The hypercycle – a principle of natural self-organization. Part A: emergency of the hypercycle*, *Naturwissenschaften* 1977, Vol. 64, p. 541–565.
- [50] Eldredge N., Gould S., *Punctuated equilibria: an alternative to phylogenetic gradualism*, [w:] *Models in Paleobiology*, pod red. T.J.M. Schopf, Freeman, Cooper and Co., San Francisco 1972, p. 82–115.
- [51] English T.M., *On the structure of sequential search: Beyond “no free lunch”*, [w:] *Evolutionary Computation in Combinatorial Optimization – EvoCOP 2004*, pod red. J. Gottlieb, G. Raidl, LNCS 3004, Springer-Verlag, Coimbra 2004, p. 95–103.
- [52] Eshelman L., Schaffer D., *Real-coded genetic algorithms and interval schemata*, [w:] *Proc. Foundations of Genetic Algorithms*, pod red. D.L. Whitley, Morgan Kaufmann Publ., San Mateo, CA 1993, p. 187–202.
- [53] Eshelman L.J., Schaffer D.J., *Crossover’s niche*, [w:] *Int. Conf. on Genetic Algorithms*, pod red. S. Forest, Morgan Kaufmann Publ. 1993, p. 9–14.
- [54] Feller W., *Wstęp do rachunku prawdopodobieństwa*, PWN, Warszawa 1981.
- [55] Fisher R.A., *On the dominance ratio*, *Proceedings of the Royal Society of Edinburgh* 1922, No. 42, p. 321–341.
- [56] Fogel D.B., *An evolutionary approach to the traveling salesman problem*, *Biol. Cybern.* 1988, Vol. 60, p. 139–144.
- [57] Fogel D.B., *Evolving artificial intelligence*, Rozprawa doktorska, Univ. of California, San Diego 1992.
- [58] Fogel D.B., *Principles of evolutionary processes*, [w:] Baeck i in. [9], p. A2.1.
- [59] Fogel D.B., Atmar J., *Comparing genetic operators with Gaussian mutations in simulated evolutionary processes using linear systems*, *Biol. Cybern.* 1990, Vol. 63, p. 111–114.
- [60] Fogel D.B., Stayton L., *On the effectiveness of crossover in simulated evolutionary optimization*, *BioSystems* 1994, Vol. 32, p. 171–182.
- [61] Fogel L., Owens A., Walsh M., *Artificial Intelligence through Simulated Evolution*, John Wiley, New York 1966.

- [62] Galar R., *Handicapped individua in evolutionary processes*, Biol. Cybern. 1985, Vol. 51, p. 1–9.
- [63] Galar R., *Evolutionary search with soft selection*, Biol. Cybern. 1989, Vol. 60, p. 357–364.
- [64] Galar R., *Miękka selekcja w losowej adaptacji globalnej w R^n . Próba biocybernetycznego ujęcia rozwoju*, Prace Naukowe Instytutu Cybernetyki Technicznej nr 84 Monografie nr 17, Wyd. Politechniki Wrocławskiej 1990.
- [65] Galar R., *Knowledge economy and evolutionary traps*, [w:] *The knowledge based economy. The global challenges of 21st century*, pod red. A. Kukliński, W. Orłowski, *Science and Government Series*, Vol. 4, Rewasz, Warszawa 2000, p. 32–61.
- [66] Galar R., *Knowledge management from the evolutionary perspective*, [w:] *Evolutionary computation and global optimization*, pod red. J. Arabas, Oficyna Wyd. Polit. Warszawskiej 2002, p. 9–20.
- [67] Galar R., *An adaptive approach to the regional strategy and its perception.*, [w:] *Education and the future. Knowledge, education and territorial development in Central and Eastern European Countries*, pod red. F. Russo, S. Arnaldi, Trieste, Italia 2003, p. 135–148.
- [68] Galar R., *Freedom, perfection and adaptive saddles*, [w:] *Intellectual entrepreneurship through or against institutions*, pod red. S. Kwiatkowski, P. Houdayer, Wyd. Wyższej Szkoły Przeds. i Zarz. im. L. Koźmińskiego 2004, p. 53–70.
- [69] Galar R., Chorażyczewski A., *Evolutionary dynamics in space of states*, [w:] *Congress on Evolutionary Computation*, Seoul, Korea 2001, p. 1366–1373.
- [70] Galar R., Karcz-Dulęba I., *The evolution of two: An example of space of states approach*, [w:] *Proc. 3rd Annual Conf. on Evolutionary Programming*, pod red. A. Sebald, L. Fogel, World Scientific, San Diego CA 1994, p. 261–268.
- [71] Galar R., Karcz-Dulęba I., *Najprostszy przypadek ewolucji: dynamika adaptacyjna dwuelementowych populacji*, [w:] *Systemy sterowania*, pod red. W. Greblicki, C. Smutnicki, WKŁ, Warszawa 2005, p. 65–76.
- [72] Galar R., Karcz-Dulęba I., Łojek R., *Adaptation cost: what is the repayment time for the evolutionary saddle crossings?*, [w:] *KAEOiG*, Będlewo 2007, p. 91–97.
- [73] Gallagher M., Freon M., Downs T., *Real-valued evolutionary optimization using a flexible probability density estimator*, [w:] *Genetic and Evolutionary Computation Conf.*, Morgan, San Francisco, CA 1999, p. 840–846.
- [74] Geard N., Wiles J., *Diversity maintenance on neutral landscapes: an argument for recombination*, [w:] *Congress on Evolutionary Computation*, Vol. 1, Honolulu 2002, p. 211–213.
- [75] Goertzel B., *A convergence theorem for the simple GA with population size tending to infinity*, [w:] *IEEE Int. Conf. on Evolutionary Computation*, 1995, p. 74–79.
- [76] Goldberg D., *Sizing populations for serial and parallel GA*, [w:] *Int. Conf. on Genetic Algorithms*, Morgan Kaufmann Publ. 1989, p. 215–221.
- [77] Goldberg D.E., *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley 1989.
- [78] Goldberg D.E., *A note on Boltzmann tournament selection for genetic algorithms and population oriented simulated annealing*, *Complex Systems* 1990, Vol. 4, p. 445–460.

- [79] Goldberg D.E., *Algorytmy genetyczne w zastosowaniach*, WNT, Warszawa 1995.
- [80] Goldberg D.E., Deb K., *A comparative analysis of selection schemes used in genetic algorithms*, [w:] *Proc. Foundations of Genetic Algorithms*, pod red. G. Rawlins, Morgan Kaufmann Publ., San Mateo, CA 1990, p. 69–93.
- [81] Goldberg D.E., Deb K., Clark J.H., *Genetic algorithms, noise, and the sizing of populations*, *Complex Systems* 1992, Vol. 6, No. 4, p. 333–362.
- [82] Goldberg D.E., Lingle R., *Alleles, loci and the traveling salesman problem*, [w:] *ICGA and their applications*, Lawrence Erlbaum, Mahwah, NJ 1985, p. 154–159.
- [83] Goldberg D.E., Segrest P., *Finite Markov chain analysis of genetic algorithms*, [w:] *Genetic Algorithms and their Applications: Proc. Int. Conf. on Genetic Algorithms*, pod red. J.J. Grefenstette, Lawrence Erlbaum, London 1987, p. 1–8.
- [84] Gould S.J., Eldredge N., *Punctuated equilibria: the tempo and mode of evolution reconsidered*, *Paleobiology* 1977, Vol. 3, p. 115–251.
- [85] Grygiel K., *On asymptotic properties of a selection-with-mutation operator*, [w:] *KAEOiG*, Murzasichle 1996, p. 50–56.
- [86] Harik G.R., Cantu-Paz E., Miller B.L., Goldberg D.E., *The gambler's ruin problem, genetic algorithms, and the sizing of populations*, [w:] *IEEE Int. Conf. on Evolutionary Computation*, IEEE Press, Indianapolis IN 1997, p. 7–12.
- [87] Hart W.E., *Rethinking the design of real-coded evolutionary algorithms: Making discrete choices in continuous search domains*, *Soft Computing Journal* 2005, Vol. 9, p. 225–235.
- [88] Haupt R.L., Haupt S.E., *Optimum population size and mutation rate for a simple real genetic algorithm that optimizes array factors*, *Applied Computational Electromagnetics Society Journal* 2000, Vol. 15, No. 2, p. 1034–1037.
- [89] Herrera F., Lozano M., Verdegay J.L., *Tackling real-coded genetic algorithms: Operators and tools for behavioural analysis*, *Artificial Intelligence Review* 1998, Vol. 12, p. 265–319.
- [90] Hesser J., Manner R., *Towards an optimal mutation probability for genetic algorithms*, [w:] *Conf. on Parallel Problem Solving from Nature*, pod red. H.P. Schwefel, LNCS 496, Springer-Verlag 1991, p. 23–32.
- [91] Holland J.H., *Adaptation in natural and artificial systems*, Univ. of Michigan Press, Ann Arbor 1975.
- [92] Horn J., *Finite Markov chain analysis of genetic algorithms with niching*, [w:] *Int. Conf. on Genetic Algorithms*, pod red. S. Forrest, Morgan Kaufmann Publ., San Mateo, CA 1993, p. 110–117.
- [93] Horn J., Goldberg D., *Genetic algorithm difficulty and the modality of fitness landscapes*, [w:] *Proc. Foundations of Genetic Algorithms*, pod red. L. Whitley, M. Vose, Morgan Kaufmann Publ., San Francisco, CA 1995, p. 243–269.
- [94] Iosifescu M., *Skończone procesy Markowa i ich zastosowania*, PWN, Warszawa 1988.
- [95] Jones J., *Evolutionary algorithms, fitness landscapes and search*, Rozprawa doktorska, Univ. Of New Mexico 1995.
- [96] Karcz I., *Badania symulacyjne stochastycznych procesów równowagi selekcyjno-mutacyjnej*, Praca magisterska, Politechnika Wrocławska 1988.
- [97] Karcz-Dulęba I., *Symulowane procesy ewolucyjne jako narzędzie optymalizacji globalnej w R^n* , Rozprawa doktorska, Politechnika Wrocławska 1992.

- [98] Karcz-Dulęba I., *Analiza właściwości adaptacyjnych dużych populacji*, [w:] *Symposium Symulacja procesów dynamicznych SPD-8*, Polana Chochołowska 1994, p. 185–192.
- [99] Karcz-Dulęba I., *Soft selection in d-optimal designs*, [w:] *Conf. on Parallel Problem Solving from Nature*, pod red. Y. Davidor, et al, LNCS 866, Springer-Verlag, Jerusalem 1994, p. 608–616.
- [100] Karcz-Dulęba I., *Niektóre aspekty zbieżności metody poszukiwań ewolucyjnych z miękką selekcją*, [w:] *KAEOiG*, Ryto 1997, p. 113–119.
- [101] Karcz-Dulęba I., *Dynamics of two element populations in the space of population states: the case of symmetrical quality functions*, [w:] *KAEOiG*, Łądek Zdrój 2000, p. 115–122.
- [102] Karcz-Dulęba I., *Dynamics of infinite populations evolving in a landscape of uni- and bimodal fitness functions*, *IEEE Trans. Evol. Comp.* 2001, Vol. 5, No. 4, p. 398–409.
- [103] Karcz-Dulęba I., *Evolution of two-element population in the space of population states: equilibrium states for asymmetrical fitness functions*, [w:] *KAEOiG*, Jastrzębia Góra 2001, p. 106–113.
- [104] Karcz-Dulęba I., *Domains of attraction of dynamical system generated by simple evolutionary model*, [w:] *Proc. Workshop on Genetic Algorithms*, Bielsko-Biała – Szyndzielnia 2002, p. 20–22.
- [105] Karcz-Dulęba I., *Evolution of two-element population in the space of population states: Equilibrium states for asymmetrical fitness functions*, [w:] *Evolutionary Algorithms and Global Optimization*, pod red. J. Arabas, Oficyna Wyd. Polit. Warszawskiej 2002, p. 35–46.
- [106] Karcz-Dulęba I., *Asymptotic behavior of discrete dynamical system generated by simple evolutionary process*, *Journ. Appl. Math. and Comp. Sc.* 2004, Vol. 14, No. 1, p. 79–90.
- [107] Karcz-Dulęba I., *Concentration of population in phenotypic evolution*, [w:] *Artificial Intelligence and Soft Computing*, pod red. L. Rutkowski, et al, LNAI 3070, Springer-Verlag 2004, p. 420–425.
- [108] Karcz-Dulęba I., *Period-doubling bifurcations in discrete dynamical system generated by evolutionary process*, [w:] *KAEOiG*, Kazimierz Dolny 2004, p. 83–88.
- [109] Karcz-Dulęba I., *Time to the convergence of evolution in the space of population states*, *Journ. Appl. Math. and Comp. Sc.* 2004, Vol. 14, No. 3, p. 279–287.
- [110] Karcz-Dulęba I., *Bifurkacje i chaos w ewolucji fenotypowej dla unimodalnych funkcji jakości*, [w:] *Systemy sterowania*, pod red. W. Greblicki, C. Smutnicki, WKŁ, Warszawa 2005, p. 41–50.
- [111] Karcz-Dulęba I., *Unstable behavior in a dynamical system generated by phenotypic evolution*, [w:] *Computer methods and systems*, pod red. R. Tadeusiewicz, et al, Oprogramowanie Naukowo-Techniczne, Kraków 2005, p. 261–269.
- [112] Karcz-Dulęba I., *Chaos detection by Lyapunov exponents in dynamical system generated by evolutionary process*, [w:] *Artificial Intelligence and Soft Computing*, pod red. L. Rutkowski, LNCS, LNAI 4029, Springer-Verlag 2006, p. 380–389.
- [113] Karcz-Dulęba I., *Dynamics of two-element populations in the space of population states*, *IEEE Trans. Evol. Comp.* 2006, Vol. 10, No. 2, p. 199–209.

- [114] Karcz-Duleba I., *Identification of cyclic and chaotic behavior in a dynamical system generated by phenotypic evolution*, [w:] KAE OiG, Kazimierz Dolny 2006, p. 195–202.
- [115] Karcz-Duleba I., *On use of unstable behavior of a dynamical system generated by phenotypic evolution*, [w:] *Knowledge-Driven Computing. Knowledge Engineering and Intelligent Computations*, pod red. C. Cotta, et al, Studies in Computational Intelligence, Springer-Verlag 2008, p. 119–134.
- [116] Karcz-Duleba I., Chorążyczewski A., *Fraktalna metoda analizy granicznych rozkładów nieskończonych populacji*, [w:] KAE OiG, Potok Złoty 1999, p. 141–148.
- [117] Karcz-Duleba I., Dziemieszkiewicz P., *Recombination and maintaining population diversity: a comparative study*, [w:] KAE OiG, Korbielów 2005, p. 95–100.
- [118] Karcz-Duleba I., Dziemieszkiewicz P., *Recombination and maintaining population diversity: simulations results*, [w:] KAE OiG, Korbielów 2005, p. 101–108.
- [119] Kimura M., *A stochastic model concerning the maintenance of genetic variability in quantitative characters*, Proc. Natl. Acad. USA 1965, Vol. 54, p. 731–736.
- [120] Kirkpatrick M., *Quantum evolution and punctuated equilibria in continuous genetic characters*, American Naturalist 1982, Vol. 119, No. 6, p. 833–848.
- [121] Kirkpatrick S., Gelatt C., Vecchi M., *Optimization by simulated annealing*, Science 1983, Vol. 220, No. 4598, p. 671–680.
- [122] Koehler G., *A proof of the Vose-Liepins conjecture*, Journal Annals of Mathematics and Artificial Intelligence 1994, Vol. 10, No. 4, p. 409–422.
- [123] Koehler G., Bhattacharyya S., Vose M., *General cardinality genetic algorithms*, Evolutionary Computation 1998, Vol. 5, No. 4, p. 439–459.
- [124] Kołodziej J., *Modeling hierarchical genetic strategy as a Lindenmayer system*, [w:] *Proc. of Int. Conf. on Parallel Computing in Electrical Engineering PARELEC '2002*, 2002, p. 409–415.
- [125] Kołodziej J., *How accurately the infinite population model can approximate the behaviour of an evolutionary algorithm*, [w:] KAE OiG, Kazimierz 2004, p. 89–97.
- [126] Kołodziej J., Ogiński L., *The estimation of the population distribution by the mixture of Erlang distributions in evolutionary algorithms*, [w:] KAE OiG, Korbielów 2005, p. 109–116.
- [127] Kwaśnicka H., Szerszon P., *NetGen- evolutionary algorithms in designing artificial neural networks*, [w:] *Conf. Neural Networks and Their Applications*, Kule 1997, p. 671–676.
- [128] Lande R., *Natural selection and random genetic drift*, Evolution 1976, Vol. 30, p. 314–334.
- [129] Leung Y., Gao Y., Xu Z., *Degree of population diversity – a perspective on premature convergence in Genetic Algorithms and its Markov chain analysis*, IEEE Trans. Neural Networks 1996, Vol. 8, No. 5, p. 1165–1176.
- [130] Lindgren K., *Evolutionary phenomena in simple dynamics*, [w:] *Artificial Life II*, pod red. C.G. Langton, et al, Addison-Wesley, Reading, MA 1992, p. 295–312.
- [131] Macdonald P., *Analysis of length-frequency distributions*, [w:] *Age and Growth of fish*, pod red. R. Summerfelt, G. Hall, Iowa State Univ. Press, Ames 1987, p. 371–384.
- [132] Makuch M., Patan K., Obuchowicz A., *Evolutionary search with soft selection in artificial neural network learning process*, [w:] KAE OiG, Murzasichle 1996, p. 87–93.

- [133] Mathias E., Whitley L., *Genetic operators, the fitness landscape and the TSP*, [w:] *Conf. on Parallel Problem Solving from Nature*, Elsevier 1992, p. 219–228.
- [134] Mattiussi K., Waibel M., Floreano D., *Measures of diversity for populations and distances between individuals with highly reorganizable genomes*, *Evolutionary Computation* 2004, Vol. 12, No. 4, p. 495–515.
- [135] Maynard Smith J., *Evolutionary genetics*, Oxford Univ. Press 1989.
- [136] McLachlan G.J., Peel D., *Finite Mixture Models*, Wiley, New York 2000.
- [137] Merz P., Freisleben B., *Memetic algorithms and the fitness landscape of the graph bi-partitioning problem*, [w:] *Conf. on Parallel Problem Solving from Nature*, Springer-Verlag, Berlin 1998, p. 765–774.
- [138] Michalewicz Z., *A GA for the linear transportation problem*, *IEEE Trans. Syst. Man Cybernetics* 1991, Vol. 21, No. 2, p. 445–452.
- [139] Michalewicz Z., *Algorytmy genetyczne + struktury danych = programy ewolucyjne*, WNT, Warszawa 1996.
- [140] Michalewicz Z., Fogel D.B., *Jak to rozwiązać, czyli nowoczesna heurystyka*, WNT, Warszawa 2006.
- [141] Mirosław L., Chorążyczewski A., Buchholz F., Kittler R., *EA validation method in detection of mitotic cells*, [w:] *KAE OiG*, Korbielew 2005, p. 157–162.
- [142] Moey C.C.J., Rowe J.E., *A reduced Markov model of GAs without the exact transition matrix*, [w:] *Conf. on Parallel Problem Solving from Nature*, 2004.
- [143] Mühlenbein H., *Towards a theory of organisms and evolving automata – open problems and ways to explore frontiers of evolutionary computation*, [w:] *Genetic Algorithms and Evolutionary Computation*, pod red. A. Menon, Vol. 11, Kluwer Acad. Publ. 2004, p. 1–39.
- [144] Mühlenbein H., Mahnig T., *Mathematical analysis of evolutionary algorithms for optimization*, [w:] *Proc. 3rd Int. Symposium on Adaptive Systems*, pod red. A. Rodriguez, et al, Cuba 2001, p. 166–185.
- [145] Nimwegen E., Crutchfield J., *Optimizing epochal evolutionary search: Population-size dependent theory*, *Machine Learning* 2001, Vol. 45, No. 1, p. 77–114.
- [146] Nimwegen E., Crutchfield J., Mitchell M., *Finite populations induce metastability in evolutionary search*, *Physics Letters A* 1997, Vol. 229, No. 3, p. 144–150.
- [147] Nimwegen E., Crutchfield J., Mitchell M., *Statistical dynamics of the royal road genetic algorithm*, *Theor. Computer Science* 1999, Vol. 214, p. 41–102.
- [148] Nix A.E., Vose M.D., *Modeling genetic algorithms with Markov chains*, *Annals of Mathematics and Artificial Intelligence* 1992, Vol. 5, p. 79–88.
- [149] Obuchowicz A., *Evolutionary search with soft selection and deterioration of the objective function*, [w:] *Intelligent Information Systems, Symposium*, pod red. M. Kłopotek, et al, Wyd. PAN, Warszawa, Zakopane 1997, p. 288–295.
- [150] Obuchowicz A., *Adaptation in a time-varying landscape using an evolutionary search with soft selection*, [w:] *KAE OiG*, Wyd. Polit. Warszawskiej, Potok Złoty 1999, p. 245–251.
- [151] Obuchowicz A., *Evolutionary search with soft selection in training a network of dynamic neurons*, [w:] *Intelligent Information Systems, Symposium*, pod red. M. Kłopotek, M. Michalewicz, Wyd. PAN, Warszawa, Ustroń 1999, p. 214–223.

- [152] Obuchowicz A., *Multi-dimensional Gaussian and Cauchy mutations*, [w:] *Intelligent Information Systems*, pod red. M. Kłopotek, et al, Physica-Verlag, Heidelberg 2002, p. 133–142.
- [153] Obuchowicz A., *Evolutionary algorithms for global optimization and dynamic system diagnosis*, Lubusky Sc. Society, Zielona Góra 2003.
- [154] Ochoa G., Harley I., Buxton H., *On recombination and optimal mutation rates*, [w:] *Genetic and Evolutionary Computation Conf.*, Morgan Kaufmann Publ. 1999, p. 13–17.
- [155] Ochoa G., Harvey I., *Recombination and error thresholds in finite populations*, [w:] *Proc. Foundations of Genetic Algorithms*, pod red. W. Banzhaf, C. Reeves, Morgan Kaufmann Publ. 1998, p. 245–264.
- [156] Peitgen H., Jurgens H., Saupe D., *Granice chaosu. Fraktale, cz.1 i cz. 2*, PWN, Warszawa 1996.
- [157] Poli R., Langdon W., *Schema theory for genetic programming with one-point crossover and point mutation*, *Evolutionary Computation* 1998, Vol. 6, No. 3, p. 231–252.
- [158] Poli R., McPhee N., Rowe J., *Exact schema theory and Markov chain models for genetic programming and variable-length genetic algorithms with homologous crossover*, *Genetic Programming and Evolvable Machines* 2004, Vol. 5, No. 1, p. 31–70.
- [159] Prügel-Bennett A., *Modeling finite populations*, [w:] *Proc. Foundations of Genetic Algorithms*, 2003, p. 99–114.
- [160] Prügel-Bennett A., *Symmetry breaking in population-based optimization*, *IEEE Trans. Evol. Comp.* 2004, Vol. 8, No. 1, p. 63–79.
- [161] Prügel-Bennett A., Rogers A., *Modelling genetic algorithm dynamics*, [w:] *Theoretical Aspects of Evolutionary Computing*, pod red. L. Kallel, et al, Springer-Verlag 2001, p. 59–86.
- [162] Prügel-Bennett A., Shapiro J., *The dynamics of a genetic algorithm for simple random ising systems*, *Physica D* 1997, Vol. 104, p. 75–114.
- [163] Qi X., Palmieri F., *Theoretical analysis of evolutionary algorithms with an infinite population size in continuous space. Part I: Basic properties of selection and mutation*, *IEEE Trans. Neural Networks* 1994, Vol. 5, p. 102–119.
- [164] Qi X., Palmieri F., *Theoretical analysis of evolutionary algorithms with an infinite population size in continuous space. Part II: Analysis of the diversification role of crossover*, *IEEE Trans. Neural Networks* 1994, Vol. 5, p. 120–128.
- [165] R. M., Jong K.D., *Measurement of population diversity*, [w:] *Selected Papers from the 5th European Conference on Artificial Evolution*, Springer-Verlag, London, UK 2002, p. 31–41.
- [166] Radcliffe N., Surry P., *Fundamental limitations on search algorithms: Evolutionary computing in perspective.*, [w:] *Computer Science Today*, pod red. J. van Leeuwen, LNCS 1000, Springer-Verlag 1995, p. 275–291.
- [167] Radcliffe N.J., *Forma analysis and random respectful recombination*, [w:] *Proc. 4th Int. Conf. on Genetic Algorithms*, pod red. R. Belew, L. Booker, Morgan Kaufmann Publ. 1991, p. 222–229.

- [168] Radcliffe N.J., *Genetic set recombination*, [w:] *Proc. Foundations of Genetic Algorithms*, pod red. D. Whitley, Morgan Kaufmann Publ., San Mateo, CA 1992, p. 203–219.
- [169] Rattray L.M., *The dynamics of a genetic algorithm under stabilizing selection*, *Complex Systems* 1995, Vol. 9, p. 213–234.
- [170] Rechenberg I., *Evolutionsstrategie: Optimierung technischer Systeme nach Prinzipien der biologischen Evolution*, Frohmann-Holzboog, Stuttgart 1973.
- [171] Rees J., Koehler G.J., *An investigation of GA performance results for different cardinality alphabets*, [w:] *Evolutionary Algorithms*, pod red. L.D. Davis, et al, Springer-Verlag, NY 1999, p. 191–206.
- [172] Reeves C.R., red., *Modern heuristic techniques for combinatorial problems*, Blackwell Scien. Publ., Oxford, UK 1993.
- [173] Reeves C.R., *Using GA with small populations*, [w:] *Proc. 5 ICGA*, Morgan Kaufmann Publ. 1993, p. 92–99.
- [174] Reeves C.R., *Landscapes, operators and heuristic search*, *Annals of Operations Research* 1999, Vol. 86, p. 473–490.
- [175] Reeves C.R., Rowe J., *Genetic Algorithms: Principles and perspectives*, Kluwer Acad. Publ. 2003.
- [176] Reeves C.R., Steele N., *Problem solving by simulated genetic process. a review and application to neural networks*, [w:] *Applied Informatics, 10th Int. Conf. IASTED*, 1992, p. 269–272.
- [177] Rocha M., Rocha I., Ferreira E.C., *A new representation in evolutionary algorithms for the optimization of bioprocesses*, [w:] *IEEE Conf. on Evolutionary Computation*, IEEE Press 2005, p. 484–490.
- [178] Rogers A., Prügel-Bennett A., *The dynamics of a GA on a model hard*, *Complex Systems* 2000, Vol. 11, No. 6, p. 437–464.
- [179] Rudolph G., *Convergence analysis of canonical genetic algorithm*, *IEEE Trans. Neural Networks* 1994, Vol. 5, No. 1, p. 96–101.
- [180] Rudolph G., *Convergence properties of evolutionary algorithms*, Verlag Dr. Kovac, Hamburg 1997.
- [181] Rudolph G., *Stochastic processes*, [w:] Baeck i in. [9], p. B2.2:1–8.
- [182] Rudolph G., *Finite Markov chain results in evolutionary computation: A tour d’horizon*, *Fundamenta Informaticae* 1998, Vol. 35, No. 1–4, p. 67–89.
- [183] Schaefer R., *Podstawy genetycznej optymalizacji globalnej*, Wyd. UJ, Kraków 2002.
- [184] Schaefer R., Jabłoński Z.J., *On the convergence of sampling measures in the global genetic search*, [w:] *PPAM Conf.*, LNCS 2328, 2002, p. 593–600.
- [185] Schaefer R., Kołodziej J., *Genetic search reinforced by population hierarchy*, [w:] *Proc. Foundations of Genetic Algorithms*, Morgan Kaufmann Publ. 2003, p. 383–401.
- [186] Schaffer J., Caruna R., Eshelman L., Das R., *A study of control parameters affecting online performance of genetic algorithms for function optimization*, [w:] *Int. Conf. on Genetic Algorithms*, pod red. J. Schaffer, Morgan Kaufmann Publ. 1989, p. 51–60.
- [187] Schwefel H.P., *Numerical optimization of computer models*, Wiley 1981.
- [188] Smith R., *Population size*, [w:] Baeck i in. [9], p. E1.1:1–5.

- [189] Spears W., *Crossover or mutation*, [w:] *Proc. Foundations of Genetic Algorithms*, pod red. D. Whitley, Morgan Kaufmann Publ. 1992, p. 221–237.
- [190] Spears W.M., *Evolutionary Algorithms: The Role of Mutation and Recombination*., Springer-Verlag, Heidelberg 2000.
- [191] Spears W.M., Jong K.A.D., *Analysing GAs using Markov models with sematically ordered and lumped states*, [w:] *Proc. Foundations of Genetic Algorithms*, pod red. R.K. Belew, M.D. Vose, Morgan Kaufmann Publ., San Francisco, CA 1997, p. 85–100.
- [192] Strogatz S.H., *Nonlinear Dynamics and Chaos*, Addison-Wesley, Reading, MA 1994.
- [193] Suzuki H., Iwasa Y., *Crossover accelerates evolution in gas with a babel-like fitness landscape: mathematical analyses*, *Evolutionary Computation* 1999, Vol. 7, No. 3, p. 275–310.
- [194] Suzuki J., *A Markov chain analysis on a genetic algorithm*, [w:] *Int. Conf. on Genetic Algorithms*, 1993, p. 146–154.
- [195] Suzuki J., *A Markov chain analysis on simple genetic algorithm*, *IEEE Trans. Syst. Man Cybernetics* 1995, Vol. 25, No. 4, p. 655–659.
- [196] Syswerda G., *A study of reproduction in generational and steady-state genetic algorithms*, [w:] *Proc. Foundations of Genetic Algorithms*, Morgan Kaufmann Publ. 1991, p. 94–101.
- [197] Titterington D.M., Smith A.F., Makov U.E., *Statistical analysis of finite mixture distributions*, Wiley 1985.
- [198] Vazquez M., Whitley L.D., *A comparison of genetic algorithm for the static job shop scheduling problem*, [w:] *Conf. on Parallel Problem Solving from Nature*, 2000, p. 303–345.
- [199] Vitanyi P., *Genetic fitness optimization using rapidly mixing Markov chains*, [w:] *7th Int. Workshop on Algorithmic Learning Theory*, LNAI 1160, Springer-Verlag, Heidelberg 1996, p. 67–82.
- [200] Voget S., *Theoretical analysis of GA with infinite population size*, *Complex Systems* 1996, Vol. 10, No. 3, p. 167–183.
- [201] Vose M., Wright A.H., *The simple genetic algorithm and the Walsh transform: Part I, theory*, *Evolutionary Computation* 1998, Vol. 6, No. 3, p. 253–273.
- [202] Vose M., Wright A.H., *The simple genetic algorithm and the Walsh transform: Part II, the inverse*, *Evolutionary Computation* 1998, Vol. 6, No. 3, p. 275–289.
- [203] Vose M.D., *The Simple Genetic Algorithm. Foundations and Theory*, MIT Press 1999.
- [204] Vose M.D., Liepins G.E., *Punctuated equilibria in genetic search*, *Complex Systems* 1991, Vol. 5, p. 31–44.
- [205] Vose M.D., Wright A.H., *Simple genetic algorithms with linear fitness*, *Evolutionary Computation* 1994, Vol. 4, No. 2, p. 347–368.
- [206] Watson R.A., Pollack J.B., *Recombination without respect: Schema combination and disruption in genetic algorithm crossover*, [w:] *Genetic and Evolutionary Computation Conf.*, pod red. D. Whitley, et al, Morgan Kaufmann Publ. 2000, p. 112–119.
- [207] Wolpert A.H., Macready W.G., *No free lunch theorems for optimization*, *IEEE Trans. Evol. Comp.* 1997, Vol. 1, p. 67–82.

- [208] Wright A., Zhao Y., *Markov chain models of genetic algorithms*, [w:] *Genetic and Evolutionary Computation Conf.*, Vol. 1, Morgan Kaufmann Publ., Orlando, Florida 1999, p. 734–741.
- [209] Wright A.H., *Genetic algorithms for real parameter optimization*, [w:] *Proc. Foundations of Genetic Algorithms*, pod red. G.J.E. Rawlins, Morgan Kaufmann Publ., San Mateo, CA 1991, p. 205–218.
- [210] Wright A.H., Agapie A., *Cyclic and chaotic behavior in genetic algorithms*, [w:] *Proc. Genetic and Evolutionary Computation Conf.*, Morgan Kaufmann Publ., San Francisco 2001, p. 718–724.
- [211] Wright A.H., Bidwell G., *A search for counterexamples to two conjectures on the simple genetic algorithm*, [w:] *Proc. Foundations of Genetic Algorithms*, pod red. R.K. Belew, M.D. Vose, Morgan Kaufmann Publ., San Francisco, CA 1997, p. 73–84.
- [212] Wright A.H., Rowe J., Neil J., *Analysis of the simple genetic algorithm on the single-peak and double-peak landscapes*, [w:] *Congress on Evolutionary Computation*, IEEE Press 2002, p. 214–219.
- [213] Wright A.H., Rowe J., Stephens C., Poli R., *Bistability in a gene pool GA with mutation*, [w:] *Proc. Foundations of Genetic Algorithms*, Morgan Kaufmann Publ. 2003, p. 63–80.
- [214] Wright S., *Evolution in Mandelian populations*, *Genetics* 1931, Vol. 16, p. 97–159.
- [215] Wright S., *Character change, speciation, and the higher taxa*, *Evolution* 1982, Vol. 36, No. 3, p. 427– 443.
- [216] Yao X., Liu Y., Lin G., *Evolutionary programming made faster*, *IEEE Trans. Evol. Comp.* 1999, Vol. 3, No. 2, p. 82–102.
- [217] Zieliński R., Neumann P., *Stochastyczne metody poszukiwania minimum funkcji*, WNT, Warszawa 1986.

Dynamics of adaptation of phenotypic evolution methods

The results of a theoretical analysis of an evolutionary search with the soft selection method are presented. The method is based on a simple model of Darwinian phenotypic evolution operating with proportional selection and Gaussian (normally distributed) mutation. In effect of superposition of both processes, populations could explore unbounded real-valued spaces, “searching” for better adaptations. Computer implementations of the method seem to be the key to the global optimisation in unbounded spaces of search. Especially, they have the ability to cross saddles between optima of a fitness function (thus capability of finding global optima), they display little sensitivity to an increase of the search space dimensionality and the robustness of inaccuracies and noise in a fitness function evaluation.

The studied evolutionary process follows a Markov process with uncountable states, thus it is difficult to analyse it theoretically. In order to obtain theoretical results, some simplifying assumptions should be made. Two extremes of the population size are analysed: infinite and two-element ones. The search spaces are one-dimensional and without constraints. The local dynamics – in search of the only optimum of an unimodal fitness, and global dynamics – when a fitness function is multimodal, are studied. In the case of global dynamics, the studies are restricted to bimodal functions and the process of crossing saddle is a main issue. Then, some results are extended to fitness functions with more than two optima.

Infinite populations are characterised by the distributions of traits and described by density of trait’s distribution functions. It is shown that for unimodal fitness functions, the distribution is normal and a population converges toward a selection-mutation equilibrium state close to the optimum. For multimodal fitness functions, the distribution of population is a mixture of Gaussian distributions. Parameters of the mixture and component distributions are described. The influence of a fitness function topology and the parameters of evolutionary process on infinite population distribution is studied.

Two-element populations are analysed in the space of population states, where each point corresponds to the whole population. The expected values of population states are calculated. These values generate a two-dimensional discrete dynamical system. Trajectories and asymptotic behaviour of the system are studied. The number and location of fixed points are analysed and the conditions of their stability are given. The number of fixed points depend on the modality of the fitness and on the standard deviation of mutation. Fixed points are usually located in the vicinity of

optima and saddles of fitness functions. The basins of attraction of fixed points and the speed of convergence toward the points are studied. Fixed points being close to saddles are unstable. Stable fixed points close to optima can lose their stability at high values of the standard deviation of mutation and then periodical orbits appear. For an asymmetrical fitness, the sequences of double-period bifurcations leading to chaos are observed. Cyclic and chaotic behaviour is analysed using techniques for the dynamical systems' theory: phase space portraits, auto-correlation function, power spectrum and Lyapunov exponents.

The results obtained from the studies of the dynamics of infinite and two-elements populations are applied to investigate different aspects of evolutionary methods. The influence of various evolutionary operators on population diversity is studied. Simulation results confirmed our earlier observations that a population diversity does not change significantly during the evolutionary process. In initially widely distributed populations, the diversity is decreasing while in almost homogeneous populations, the diversity is increasing. As a result, a compact cluster of individuals with similar types drifts through a search space. Critical parameters derived from the dynamical system, for which fixed points change their stability, are used to set parameters of the evolutionary process and to identify unknown one-dimensional fitness functions.

The survey of the known literature methods devoted to the theoretical analysis of evolutionary algorithms is also presented in the monograph. The models of phenotypic evolution, models exploiting the Markov chains, the theory of dynamical systems, schemes, etc., are described.

Spis treści

Wykaz ważniejszych skrótów i oznaczeń	5
1. Wstęp	7
1.1. Tło historyczne i motywacja	7
1.2. Cel i zakres pracy	10
1.3. Układ pracy	12
2. Metody ewolucyjne	13
2.1. Podstawowe idee	13
2.1.1. Ogólny schemat algorytmu ewolucyjnego	13
2.1.2. Algorytmy genetyczne (AG)	15
2.1.3. Strategie ewolucyjne (SE)	16
2.2. Mechanizmy ewolucji	16
2.2.1. Metody selekcji	17
2.2.2. Mutacja	18
2.2.3. Operatory rekombinacji (krzyżowania)	18
2.3. Metody fenotypowe. Metoda poszukiwań ewolucyjnych z miękką selekcją .	19
2.3.1. Modele ewolucji fenotypowej	19
2.3.2. Poszukiwania ewolucyjne z miękką selekcją	21
2.4. Dynamika poszukiwań ewolucyjnych	24
2.4.1. Krajobraz adaptacyjny funkcji przystosowania	24
2.4.2. Dynamika populacji	26
2.4.3. Twierdzenie <i>No free lunch</i>	29
3. Teoria metod ewolucyjnych	31
3.1. Analiza metod z kodowaniem rzeczywistoliczbowym	31
3.2. Modele, w których stosowane są łańcuchy Markowa	33
3.3. Modele, w których stosuje się teorię układów dynamicznych	36
3.4. Inne modele	38
4. Dynamika modelu nieskończonej populacji	43
4.1. Model poszukiwań ewolucyjnych z miękką selekcją dla nieskończonych populacji	44
4.2. Dynamika lokalna	45
4.3. Dynamika globalna	48
4.3.1. Rozkłady populacji	48
4.3.2. Parametry rozkładu mieszaniny	49
4.3.3. Badania symulacyjne rozkładów	54
4.3.4. Dyskusja	62
4.3.5. Zbiór Cantora generowany przez wartości oczekiwane rozkładu . . .	63

5. Dynamika bardzo małych populacji	65
5.1. Przestrzeń stanów populacji	66
5.2. Dwuelementowa populacja w przestrzeni stanów	67
5.3. Analiza wartości oczekiwanych stanów populacji	70
5.4. Wartości oczekiwane stanu populacji jako dyskretny układ dynamiczny	80
5.5. Analiza punktów stałych i ich stabilności	82
5.5.1. Punkty stałe	82
5.5.2. Obszary przyciągania punktów stałych	85
5.5.3. Stabilność punktów stałych	89
5.6. Badanie zachowań chaotycznych	101
5.7. Oczekiwany czas zbieżności	106
5.8. Uwagi dotyczące kilkuelementowych populacji	110
6. Przykłady wykorzystania analizy dynamiki populacji	113
6.1. Różnorodność populacji	113
6.1.1. Miary zróżnicowania populacji	114
6.1.2. Operatory selekcji	115
6.1.3. Operatory mutacji	118
6.1.4. Operatory rekombinacji	120
6.1.5. Podsumowanie	124
6.2. Określenie parametrów procesu ewolucyjnego i funkcji celu	127
6.2.1. Dobór parametrów algorytmu	128
6.2.2. Identyfikacja funkcji celu	129
6.2.3. Uwagi o liczebności populacji	130
7. Podsumowanie	133
Dodatek. Funkcje celu	137
Bibliografia	139
Dynamics of adaptation of phenotypic evolution methods	151

Contents

Notational convention 5

1. Introduction 7

1.1. Historical background and motivation 7

1.2. Aim and scope of the dissertation 10

1.3. Dissertation layout 12

2. Evolutionary methods 13

2.1. Basic ideas 13

2.1.1. General outline of evolutionary algorithms 13

2.1.2. Genetic algorithms (GA) 15

2.1.3. Evolutionary strategies (ES) 16

2.2. Evolutionary mechanisms 16

2.2.1. Selection methods 17

2.2.2. Mutation 18

2.2.3. Recombination (crossover) operators 18

2.3. Phenotypic methods. Method of evolutionary search with soft selection . 19

2.3.1. Models of phenotypic evolution 19

2.3.2. Evolutionary search with soft selection 21

2.4. Dynamics of evolutionary search 24

2.4.1. Adaptive landscape of a fitness function 24

2.4.2. Dynamics of population 26

2.4.3. No free lunch theorem 29

3. Theory of evolutionary methods 31

3.1. Analysis of methods with real-valued encoding 31

3.2. Models using Markov chains 33

3.3. Models using dynamical systems' theory 36

3.4. Other models 38

4. Dynamics of infinite population model 43

4.1. Model of evolutionary search with soft selection for infinite populations . . 44

4.2. Local dynamics 45

4.3. Global dynamics 48

4.3.1. Distributions of populations 48

4.3.2. Parameters of mixed distributions 49

4.3.3. Simulation studies of distributions 54

4.3.4. Discussion 62

4.3.5. Cantor set generated by expected values of distributions 63

5. Dynamics of very small populations	65
5.1. Space of population states	66
5.2. Two-element population in space of states	67
5.3. Analysis of expected values of population states	70
5.4. Expected values of population states as a discrete dynamical system	80
5.5. Analysis of fixed points and their stability	82
5.5.1. Fixed points	82
5.5.2. Basins of attraction of fixed points	85
5.5.3. Stability of fixed points	89
5.6. Studies on chaotic behaviour	101
5.7. Expected time to convergence	106
5.8. Notes on populations of a few	110
6. Examples of exploitation of population dynamics analysis	113
6.1. Population diversity	113
6.1.1. Diversity measures	114
6.1.2. Selection operators	115
6.1.3. Mutation operators	118
6.1.4. Recombination operators	120
6.1.5. Summary	124
6.2. Determination of parameters of evolutionary process and fitness function	127
6.2.1. Choice of algorithm parameters	128
6.2.2. Identification of fitness function	129
6.2.3. Notes about population size	130
7. Summary	133
Appendix – fitness functions	137
References	139
Dynamics of adaptation of phenotypic evolution methods	151

PRACE NAUKOWE INSTYTUTU INFORMATYKI, AUTOMATYKI I ROBOTYKI
(wydane w latach 1999–2008)

- Nr 100, Monografie nr 26, K. Nitka-Styczeń, *Optymalne sterowanie procesów z opóźnieniami, teoria, algorytmy i zastosowania*, Wrocław 1999
- Nr 101, Monografie nr 27, E. Nowicki, *Metoda tabu w problemach szeregowania zadań produkcyjnych*, Wrocław 1999
- Nr 102, Konferencje nr 46, *VII Krajowa konferencja robotyki*, Tom 1, Wrocław 2001
- Nr 103, Konferencje nr 47, *VII Krajowa konferencja robotyki*, Tom 2, Wrocław 2001
- Nr 104, Monografie nr 28, E. Skubalska-Rafajłowicz, *Krzywe wypełniające w rozwiązywaniu wielowymiarowych problemów decyzyjnych*, Wrocław 2001
- Nr 105, Monografie nr 29, O. Unold, *Ewolucyjne wnioskowanie gramatyczne*, Wrocław 2006

Algorytmy ewolucyjne są chętnie stosowanymi heurystycznymi technikami optymalizacyjnymi. W monografii przedstawiono analizę teoretyczną fenotypowych metod ewolucyjnych na przykładzie metody poszukiwań ewolucyjnych z miękką selekcją, bazującej na selekcji proporcjonalnej i mutacji gaussowskiej. Prezentowane wyniki dotyczą badań dynamiki lokalnej i globalnej populacji nieskończonych i dwuelementowych w jednowymiarowych rzeczywistych przestrzeniach poszukiwań bez ograniczeń. Nieskończone populacje scharakteryzowano za pomocą rozkładów cech. Do analizy populacji dwuelementowych zastosowano teorię układów dynamicznych. Rezultaty teoretyczne uzupełniono i potwierdzono wynikami badań symulacyjnych.



Wydawnictwa Politechniki Wrocławskiej
są do nabycia w księgarni „Tech”
plac Grunwaldzki 13, 50-377 Wrocław
budynek D-1 PWr., tel. 071 320 29 35
Prowadzimy sprzedaż wysyłkową

ISSN 0324-9786