

Praca magisterska

Analiza czułości niestacjonarnych układów turbulentnych

Adam Marek

28 listopada 2019

Spis treści

1	Wprowadzenie	3
1.1	Ogólna charakterystyka	3
1.2	Tło literaturowe badanego problemu	3
1.3	Układ pracy	4
2	Elementy teorii gładkich układów dynamicznych	5
2.1	Pojęcia podstawowe	5
2.2	Stabilność rozwiązań	11
2.3	Chaos deterministyczny	17
2.4	Lemat cienia	30
3	Wybrane metody optymalizacji lokalnej stosowane w technice	33
3.1	Metoda różnic skończonych	33
3.2	Metoda Tangent	34
3.3	Metoda Operatora Sprzężonego	34
4	Metoda Trajektorii Cienia	36
4.1	Wstęp	36
4.2	Metoda Operatora Cienia	37
4.3	Metoda Trajektorii Cienia	40
5	Implementacja metody LSS	45
5.1	Scenariusz I - koncektualne preliminarium	45
5.1.1	Model matematyczny	45
5.1.2	Wyniki symulacji numerycznej	46
5.2	Scenariusz II - oblicze chaosu	56
5.2.1	Słowo wstępne	56
5.2.2	Metoda Siatek Boltzmanna	57
5.2.3	Model matematyczny	58
5.2.4	Aspekty numeryczne	59
5.2.5	Wyniki symulacji	61
6	Dyskusja wyników oraz wnioski	80
6.1	Scenariusz pierwszy	80
6.2	Scenariusz drugi	80
6.3	Konkluzje	81

1 Wprowadzenie

1.1 Ogólna charakterystyka

Istotę niniejszej pracy stanowi efektywna implementacja metody wyznaczania wrażliwości dla pewnej klasy funkcji celu stanowiących długookresowe wartości średnie w przypadku układów dynamicznych, których rozwiązania przejawiają skomplikowane zachowania. Analiza taka jest niezbędna w wielu praktycznych zastosowaniach, takich jak optymalizacja kształtu elementów maszyn i urządzeń ciepłno-przepływowych [19], sterowanie z przesuwnym horyzontem czasowym (predykcyjne) [2][3], adaptacja siatek obliczeniowych [42], szacowanie błędów i niepewności [8], asymilacja danych związana z prognozami pogody [37]. Mimo że praca ta skupia się głównie na zastosowaniach związanych z mechaniką płynów, przedstawione algorytmy mogą być pomocne do rozwiązywania szerszej klasy problemów związanych z układami dynamicznymi.

W ostatnich latach powstało kilka prac sugerujących sposób rozwiązania postawionego problemu, jednakże wspólną ich cechą jest wysoki koszt numeryczny. W niniejszej pracy wykorzystuje się podstawowe ich idee - w szczególności zastosowana została Metoda Trajektorii Cienia (ang. *Least Squares Shadowing*) w formie wariacyjnej. Całość została zaimplementowana w języku skryptowym *MATLAB* oraz w języku wysokiego poziomu: *C++*. Nowość stanowi sposób dyskretyzacji równań mechaniki płynów - zastosowana została Metoda Siatkowa Boltzmanna.

1.2 Tło literaturowe badanego problemu

Opracowane do tej pory klasyczne metody estymacji czułości bazujące na wyznaczaniu gradientu funkcji celu pozostają nieskuteczne wobec istotnych z punktu widzenia techniki układów niestacjonarnych, których dynamika (silnie nieliniowa) przekłada się na praktycznie nieprzewidywalne zachowanie. Mówi się wówczas o układach chaotycznych (patrz Rozdział 2). Świetną pozycję opisującą historię odkryć takich układów, związane z nimi problemy oraz pierwsze próby ich matematycznego opisu stanowi książka [15]. Trudności polegające na załamaniu się klasycznych metod wyznaczania wrażliwości zotały po raz pierwszy zauważone w pracy [22] przy badaniu układu Lorenza (1963) stanowiącego skrajnie uproszczony model termicznej konwekcji naturalnej w atmosferze. Ci sami autorzy w pracy [23] zaproponowali sposób regularyzacji poprzez uśrednianie całkowe w przestrzeni fazowej dla skal czasowych zdeterminowanych poprzez największą wartość wykładników Lyapunowa. Uzasadnienie takiego postępowania stanowi pozycja [31], w której autor dowodzi poprawności sformułowania problemu (tj. pokazuje, że poszukiwana pochodna istnieje i jest ograniczona przy pewnych ogólnych założeniach) oraz podaje *explicite* jej wzór. Podejście takie bez znaczących modyfikacji jest niepraktyczne ze względu na koszt numeryczny, co zostało pokazane w pracy [9].

Innym podejściem jest próba rozwiązania równań Fokkera-Plancka odnoszących się do funkcji gęstości prawdopodobieństwa zdefiniowanej na przestrzeni fazowej ujęte w pracy [38]. Kontynuację tego podejścia można znaleźć w pracy [5]. Dużym problemem jest regularność wspomnianej funkcji rozkładu - uzyskuje się ją poprzez dodanie sztucznego członu

do równań wprowadzając tym samym błędy rozwiązań. Podobna procedura wykorzystuje Twierdzenie o Dyssypacji Fluktuacji [1], jednakże zмага się ona z analogicznymi problemami.

Pomysł Metody Trajektorii Cienia, stanowiącej główny obiekt zainteresowań w ramach niniejszej pracy, narodził się w ramach zdefiniowania Operatora Cienia w pracy [39], w której autor rozwiązuje postawiony problem poprzez dekompozycję przestrzeni fazowej na wektory kowariantne Lyapunova. Takie rozwiązanie jest kosztowne w przypadku obecności wielu dodatnich wykładników Lyapunova, co jest cechą charakterystyczną układów chaotycznych. Pierwszą pracą, w której wspomniana metoda została formalnie zdefiniowana jest [41]. Analizę zbieżności zregulowanego w ten sposób rozwiązania ilustruje [40]. Przykład jej zastosowania w ramach równań Kuramoto-Shivashinsky zawiera praca [7]. Wyznaczenie wrażliwości dla dwuwymiarowego układu przepływowego zdyskretyzowanego równaniami RANS zawiera [6].

Uproszczona wersja Metody Trajektorii Cienia po raz pierwszy pojawiła się w artykule [12], w której rolę funkcji odpowiadającej za dylatację czasu (zob. Rozdział 4) pełni z góry określona funkcja okienka, stanowiąca odpowiednio regularne jądro całkowe o zwartym nośniku.

1.3 Układ pracy

Rozdział 2 zawiera przegląd definicji oraz niezbędnych faktów dotyczących teorii układów dynamicznych i chaosu deterministycznego wykorzystywanych w dalszych rozdziałach. W Rozdziale 3 przedstawiono wybrane metody klasyczne wyznaczania gradientu funkcji celu, na których bazują algorytmy optymalizacji lokalnej stosowane w technice. Rozdział 4 prezentuje metodę najmniejszych kwadratów dla trajektorii cienia (ang. *Least Squares Shadowing*, oznaczana dalej również jako LSS). W rozdziale 5 zestawiono wyniki dla dwóch przypadków:

- układu Lorenza (1963),
- dwuwymiarowego zagadnienia przepływowego, w którym geometrię stanowi kanał o skokowo zmiennej szerokości.

Rozdział 6 przedstawia omówienie wyników, a także konkluzje

2 Elementy teorii gładkich układów dynamicznych

Dla zrozumienia problemów związanych ze stabilnością układów ewoluujących w czasie konieczne jest zapoznanie się z nomenklaturą obowiązującą w matematycznym opisie takich układów. W rozdziale tym przedstawione zostały podstawowe definicje i fakty dotyczące zagadnień stabilności równań różniczkowych zwyczajnych oraz chaosu deterministycznego. Na wstępie należy zaznaczyć, że nie należy mylić tego pojęcia z terminem chaotyczności przywołanym na wstępie pracy, który oznacza skomplikowane zachowanie rozwiązań i w ogólności nie musi zależeć od wartości wykładników Lyapunowa stanowiących formalne kryterium chaotyczności w sensie układów dynamicznych. Mimo że w tej pracy przedstawiona została między innymi metoda analizy wrażliwości na przykładzie układu Lorentza, który w istocie jest chaotyczny (dodatnie wykładniki Lyapunowa), celem zainteresowań będzie szersza klasa problemów, o których (w większości) nie ma informacji związanych z ich przynależnością do tej klasy problemów (w szczególności nie jest udowodniona chaotyczność układu równań Naviera-Stokesa). Zaproponowana metoda jest jednak dogodna do dużej klasy układów przejawiających skomplikowane zachowania, które w dalszej części pracy określa się mianem chaotycznych.

2.1 Pojęcia podstawowe

Niech M oznacza rozmaitość różniczkową z metryką Riemanna oznaczaną symbolem d . W dalszej części pracy (w przypadku braku konfliktu oznaczeń) przetrzeń metryczna (M, d) będzie oznaczana przez M . Niech dany będzie homeomorfizm $\mathbf{f}: M \mapsto M$. Załóżmy również, że dany jest układ autonomicznych równań różniczkowych zwyczajnych postaci

$$\frac{d\mathbf{u}}{dt} = \mathbf{f}(\mathbf{u}; \mathbf{s}), \quad (1)$$

z warunkiem początkowym

$$\mathbf{u}(0) = \mathbf{u}_s \quad (2)$$

gdzie $t \in (0, T]$ oznacza czas ($T > 0$ - rozpatrywany okres), \mathbf{s} oznacza wektor zmiennych decyzyjnych (w problemach optymalizacyjnych nazywany wektorem parametrów projektowych - nie stanowi on zmiennej niezależnej, należy go traktować jako parametr), zaś \mathbf{u}_0 oznacza funkcję wektorową, której regularność uzależniona jest od struktury równania (1) (przykładowo w analizie równań Naviera-Stokesa przyjęlibyśmy założenie, że $\mathbf{u}_0 \in V$, gdzie V oznacza domknięcie funkcji gładkich o zerowej dywergencji w odpowiedniej przestrzeni Sobolewa¹). Autonomiczność układu oznacza, że jego prawa strona nie zależy *explicite* od czasu t .

W tym momencie poczynione zostaną uwagi o charakterze technicznym, które pozostają w mocy do końca bieżącego rozdziału:

- Równanie (1) jest w ogólności określone na przestrzeni funkcji (tj. na pewnej nieskończonej wymiarowej przestrzeni Banacha). Jednakże jakościowe warunki dotyczące

¹Więcej informacji na ten temat Czytelnik znajdzie na przykład w [36]

istnienia i jednoznaczności rozwiązań takich równań zachodzą i w takich przypadkach (tj. twierdzenia: Peano oraz Picarda-Lindelöfa mają analogiczną postać do swoich skończenie wymiarowych odpowiedników).

- Funkcję określoną na przestrzeni nieskończenie wymiarowej określa się często mianem operatora. W tej pracy konwencja ta nie będzie obowiązywała, aby nie zaciemniać istoty rzeczy.
- Poprzez rozmaitość różniczkową M generowaną na pewnej przestrzeni Hilberta H rozumie się przestrzeń topologiczną Hausdorffa, w której każdy punkt ma otoczenie dyfeomorficzne z pewnym zbiorem otwartym w H (w przypadku skończenie wymiarowym każdy punkt ma otoczenie dyfeomorficzne z pewną kulą).
- Metryka Riemanna jest w ogólności pojęciem bardzo abstrakcyjnym. Jej precyzyjna definicja wymaga pojęcia kowariantnego pola tensorowego, którego wprowadzanie pozostaje poza zakresem tej pracy. W związku z tym przedstawiona zostanie intuicja stojąca za tym pojęciem odsyłając zainteresowanego Czytelnika do literatury [4], [28].
- Pełne dookreślenie układu dynamicznego wymaga zdefiniowania niezmienniczej względem homeomorfizmu \mathbf{f} miary. Oczywiście 'zwykła' miara Lebesgue'a nie jest odpowiednia ze względu na brak niezmienniczości (w większości przypadków objętości, pod wpływem dynamiki, redukują się do tworów niższej wymiarowych). W związku z czym wprowadzona została miara SRB (nazwa pochodzi od nazwisk jej twórców: Sinai, Ruelle, Bowen). Miara ta konieczna jest jedynie do zdefiniowania własności układu zwaną ergodycznością, która zostanie wytłumaczona na gruncie intuicji, wobec czego formalna definicja miary SRB nie zostanie sformułowana. Można ją znaleźć na przykład w pracy [31]. W pracy tej przyjmuje ona oznaczenie μ

Równanie (1) można interpretować jako ewolucję w czasie wielkości wektorowej, na przykład pola prędkości przy zadanych warunkach brzegowych, które w ogólności zależą od czasu, współrzędnych przestrzennych oraz od wektora zmiennych decyzyjnych:

$$\mathbf{u} = \mathbf{u}(t, \mathbf{x}; \mathbf{s}). \quad (3)$$

Realizacja numeryczna równania (1) z warunkiem początkowym (2) wyraża się poprzez procedurę iteracyjną postaci

$$\mathbf{u}_{i+1} = \mathbf{f}^s(\mathbf{u}_i) \quad (4)$$

$$\mathbf{u}_0 = \mathbf{u}_s, \quad (5)$$

gdzie \mathbf{f}^s oznacza zdyskretyzowany homeomorfizm \mathbf{f} sparametryzowany wektorem \mathbf{s} . Wobec postaci równań (1) i (4) można opisywane przez nie układy dynamiczne identyfikować z funkcjami prawych stron (zadającymi dynamikę). W dalszej części pracy o powyższych zagadnieniach zakłada się, iż mają jednoznaczne rozwiązanie dla $t \in (0, T]$, tzn. spełniają one założenia twierdzenia Picarda-Lindelöfa (dla przestrzeni M). Jednym z najistotniejszych pojęć pozwalających na jakościową analizę układów dynamicznych jest *przestrzeń fazowa*.

Definicja 1. Przestrzenią fazową $M_{ph} \subseteq M$ równania (1) nazywamy zbiór wszystkich stanów tego układu traktowanych jako punkty tej przestrzeni.

W dalszym ciągu pracy o przestrzeni fazowej M_{ph} zakłada się, iż ma odpowiednią strukturę topologiczną. Aby uściślić to pojęcie założymy o \mathbf{f} , iż jest dyfeomorfizmem określonym na przestrzeni fazowej ($\mathbf{f}: M_{ph} \mapsto M_{ph}$) która jest zbiorem zwartym, bez brzegu². Przy tych założeniach można przyjąć następujące definicje.

Definicja 2. Przestrzenią styczną $T_x M_{ph}$ do przestrzeni fazowej w punkcie $x \in M_{ph}$ nazywamy zbiór $\{x\} \times \phi'(u) \cdot E \subset \{x\} \times M$, gdzie $\phi(U)$ jest parametryzacją otoczenia punktu $x \in M_{ph}$ (istniejąca na mocy założeń o M) taką, że $\phi(u) = x$

Definicja 3. Wiązką styczną TM_{ph} do przestrzeni fazowej $M_{ph} \subseteq M$ nazywamy rodzinę przestrzeni stycznych $TM_{ph} = \{(x, y) | x \in M_{ph}, y \in T_x M_{ph}\}$.

Tak zdefiniowana wiązka stanowi dziedzinę pochodnej odwzorowania \mathbf{f} , tj. $D\mathbf{f}: TM_{ph} \mapsto TM_{ph}$. Wizualizację wiązki stycznej dla prostych kształtów geometrycznych ilustruje Rysunek (1).

²Zwartość jest stosunkowo mocnym założeniem, które często będzie pomijane. Niemniej dla ścisłości rozumowania jest ono konieczne



Rysunek 1: Wizualizacja pojęcia wiązki stycznej dla okręgu

Jednym z dwóch najważniejszych założeń o rozpatrywanych układach dynamicznych jest możliwość dekompozycji ich przestrzeni fazowych na sumę prostą dwóch podprzestrzeni, których wektory cechują się podobnymi właściwościami dynamicznymi.

Definicja 4. Powiemy, że zbiór M_{ph} ma strukturę hiperboliczną, o ile istnieją zbiory $E^s, E^u \subseteq TM_{ph}$ (nazywane wiązką stabilną i niestabilną, odpowiednio) takie, że $TM_{ph} = E^s \oplus E^u$ i spełniające warunki:

- $Df|_{E^s}$ jest kontrakcją,
- $Df|_{E^u}$ jest ekspansją.

Istnieją wówczas stałe $\lambda \in (0, 1), c \in (0, +\infty)$ takie, że:

- $\|Df^i x_s\|_M \leq c\lambda^i \|x_s\|_M$ dla dowolnych $x_s \in E^s$ oraz $0 < i \in \mathbb{N}$,
- $\|Df^{-i} x_u\|_M \leq c\lambda^i \|x_u\|_M$ dla dowolnych $x_u \in E^u$ oraz $0 < i \in \mathbb{N}$.

Układ dynamiczny, którego przestrzeń fazowa ma strukturę hiperboliczną będzie nazywany układem hiperbolicznym. Ostatnie dwa warunki w Definicji 4 grupują wektory przestrzeni stycznej ze względu na jakościowe zachowanie ich norm wraz z rozwojem układu. Jak się przekonamy w późniejszych rozdziałach jest to niezbędne założenie warunkujące użyteczność różnych metod regularyzacji pozwalających na wyznaczanie wrażliwości w rozważanych układach.

Drugie, równie istotne założenie dotyczy pośrednio funkcji, której gradient jest poszukiwany. Istotnymi z punktu widzenia techniki pozostają długookresowe wartości średnie, tj. funkcje postaci

$$\bar{J}(u; s) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T J(u; s) dt, \quad (6)$$

gdzie $J(u; s)$ oznacza chwilową wartość funkcji będącej obiektem zainteresowań. Może ona oznaczać przykładowo opór aerodynamiczny, siłę nośną, ciąg silnika lub współczynnik przejmowania ciepła na powierzchniach wymiennika ciepła. W praktyce dysponujemy oczywiście jedynie skończonymi przybliżeniami granicy (6), tj.

$$\bar{J}(u; s) \approx \frac{1}{\Delta T} \sum_{i=0}^{N-1} \frac{J(u_{i+1}; s) + J(u_i; s)}{2} (T_{i+1} - T_i), \quad (7)$$

gdzie ΔT oznacza przyjętą (odpowiednio dużą) długość okresu uśredniania, $(T_{i+1} - T_i)$ określa długość kroku czasowego w danym kroku (przyjęty został w tym przypadku schemat trapezów dla całkowania numerycznego).

Aby zastosować odpowiedni aparat matematyczny w dalszym ciągu rozpatrywana będzie formuła postaci (6) pozostawiając jej numeryczną realizację (7) na moment opisu implementacji odpowiedniej metody. Poprzez wrażliwość tak zdefiniowanej funkcji będziemy rozumieli jej gradient względem wektora parametrów \mathbf{s} . Klasyczne metody wyznaczania takiej wielkości przedstawione zostały w Rozdziale, zaś problemy z nich wynikające przy aplikacji do układów chaotycznych wyjaśnia Rozdział 4. Należy przy tym zaznaczyć, że aby cały ten aparat mógł być użyty należy przyjąć kolejne, bardzo istotne założenie dotyczące rozpatrywanego układu, które będzie rzutowało na regularność takiej funkcji: ergodyczność systemu.

Definicja 5. Niech dany będzie układ dynamiczny $\mathbf{f}: M_{ph} \mapsto M_{ph}$ z niezmienniczną miarą μ . O przekształceniu \mathbf{f} powiemy, że jest ergodyczne o ile dla każdego zbioru mierzalnego $A \subset M_{ph}$ zachodzi implikacja $T^{-1}(A) = A \implies \mu(A) = 0$ lub $\mu(M_{ph} - A) = 0$.

Zauważmy, że wobec założenia dotyczącego różniczkowalności $\mathbf{f}: M_{ph} \mapsto M_{ph}$ przeciwobraz dowolnego zbioru mierzalnego można podmienić na obraz tegoż zbioru. A więc przekształcenie ergodyczne to takie, dla którego jedynymi zbiorami niezmienniczymi są te pełnej miary lub zbiory miary zero.

Formalna definicja ergodyczności wydawać się może abstrakcyjna, niemniej wprowadzenie tego pojęcia wynikało pierwotnie z potrzeb fizyki, a dokładnie z mechaniki statystycznej i ma ono praktyczną interpretację. Załóżmy bowiem, że chcielibyśmy wyznaczyć temperaturę gazu w danym pomieszczeniu. Patrząc na to z eksperymentalnego punktu widzenia nie ma w tym nic nadzwyczajnego - wystarczy w dowolnym miejscu dokonać pomiaru

za pomocą termometru. Należy przy tym odczekać odpowiednią ilość czasu związaną z termiczną bezwładnością przyrządu. Na ten oczywisty fakt można spojrzeć z punktu widzenia układów dynamicznych - temperatura jest miarą średniej energii kinetycznej ruchu postępowego cząstek gazu w pomieszczeniu (przyjmujemy tutaj model gazu doskonałego jako dobre przybliżenie). Pomiar za pomocą termometru mierzy energie kinetyczne cząstek, których wartości różnią się. Stąd, aby pomiar był poprawny, należy zapewnić odpowiednio długi okres uśredniania. Dodatkowo należy zwrócić uwagę, że nie ma znaczenia w którym miejscu tego pomiaru dokonamy (gaz jest w równowadze termodynamicznej). Mimo, że z mikroskopowego punktu widzenia w każdej chwili cząstki poruszają się chaotycznie ich średnia wartość pewnych wielkości (w tym przypadku energii kinetycznej) nie zmienia się. Analizując to zadanie, ale tym razem na drodze teorii napotyka się pewien problem. Rozwiązanie równań dynamiki dla tak złożonego układu jest zbyt kosztowne w sensie czasu obliczeń. Z mechaniki statystycznej wiadomo jednakże, że istnieje w stanie równowagi funkcja rozkładu prędkości cząstek (rozkład Maxwella-Boltzmanna). Jest to funkcja gęstości prawdopodobieństwa i dzięki niej można bez większych problemów (tj. całkując po przestrzeni fazowej) wyznaczyć poszukiwaną średnią wartość energii kinetycznej przypadającej na cząstkę a w konsekwencji temperaturę gazu. Kryje się za tym jednak pewna subtelność - wykorzystane zostało uśrednianie po przestrzeni fazowej (po wszystkich możliwych prędkościach z wagami odpowiadającymi rozkładowi Maxwella-Boltzmanna). Jest to abstrakcyjny zabieg, który można sobie wyobrazić jako mierzenie parametrów jednej cząstki 'rozmażanej' przez funkcję rozkładu po wszystkich możliwych jej stanach. W tym momencie dochodzimy do istoty problemu: kiedy można powiedzieć, że losowo wybrana cząsta jest reprezentatywna w tym sensie, że (po odpowiednio długim czasie) wartość średnia mierzonej dla niej wielkości będzie odpowiadała wartości oczekiwanej tej samej wielkości dla zespołu cząstek? Kierując się intuicją można dojść do wniosku, że owa cząstka z pewnością winna przyjmować każdą z możliwych wartości prędkości z odpowiednią częstością dyktowaną rozkładem. Innymi słowy: powinna ona przejść przez (prawie) każdy punkt przestrzeni fazowej, przy czym okres 'przebywania' w różnych jej obszarach nie może być jednorodny. Aby to mogło zajść układ powinien wykazywać brak stałych tendencji, a więc przejawiać własności chaotyczne, w powszechnym tego słowa znaczeniu. Jako pierwszy do takiego wniosku doszedł sam Boltzmann, który opisał to słowami: "Duża nieregularność ruchu cieplnego i zmienność sił wewnętrznych działających na ciało stwarza duże prawdopodobieństwo, że atomy w ruchu cieplnym mogą przyjmować wszystkie wartości położeń i pędów zgodnie z ustaloną energią wewnętrzną ciała" [35]. Wniosek taki nie jest ścisły, ale jest bardzo bliski prawdy. Okazuje się bowiem, że warunkiem koniecznym i dostatecznym ergodyczności układu jest brak możliwości podzielenia przestrzeni fazowej na dwa podzbiory o niezerowych miarach, niezmienniczych względem dynamiki układu (jest to nieformalne sformułowanie słynnych twierdzeń ergodycznych autorstwa George'a Birkhoffa). Stąd (i z definicji ergodyczności) można już dojść do wniosku, że brak takich obszarów przestrzeni fazowej wymusza pokrywanie przez μ -prawie każdą trajektorię całej przestrzeni fazowej (nie wszystkie, bo mogą bowiem istnieć orbity cykliczne, o których będzie mowa w następnym rozdziale). Co więcej funkcja rozkładu statystycznego takich układów jest ustalona i niezmienna względem czasu, a więc średnie długookresowe istnieją i są (przynajmniej w przejściu granicznym) sobie równe.

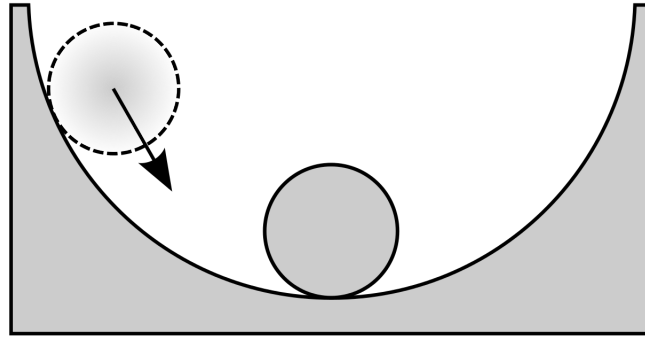
Wobec powyższego rozumowania można dojść do wniosku, że własność ta implikuje istotną cechę rozpatrywanych układów: długookresowe wartości średnie są niezależne od warunków początkowych (uściślając: są one takie same dla prawie wszystkich warunków początkowych) i ich wartość można równoważnie uzyskać poprzez uśrednianie po przestrzeni fazowej układu (względem miary probabilistycznej, która w takich układach jest niezmiennicza). Jak się później przekonamy (Rozdział) całkowanie w przestrzeni możliwych stanów układu ma działanie regularyzujące.

2.2 Stabilność rozwiązań

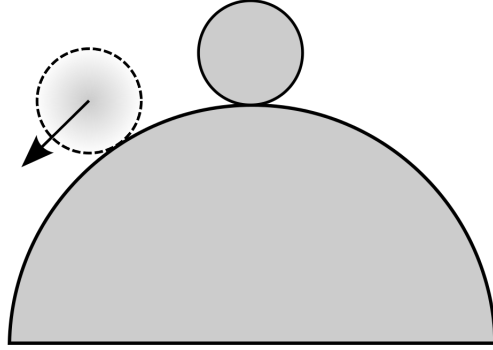
Zagadnienie stabilności związane jest z odpowiedzią na pytanie jak zmieni się globalny przebieg rozwiązania pod wpływem małych (w odpowiednim sensie) perturbacji warunków początkowych lub funkcji określającej dynamikę układu, który jest zidentyfikowany poprzez (wektorowe) równanie różniczkowe zwyczajne.

Definicja 6. Punkt $u_c \in M_{ph}$ nazywamy położeniem równowagi równania (1), o ile jest to punkt krytyczny odwzorowania f , tzn. $f(u_c) = 0$.

Zauważamy, że położenie równowagi jest (stałym) rozwiązaniem równania (1) z takim samym warunkiem początkowym. Punkty takie jakościowo mogą istotnie się różnić ze względu na ewolucję w czasie swojego otoczenia. Ilustrację tej uwagi stanowi Rysunek (2) przedstawiający ideę stabilnych punktów równowagi oraz analogiczny Rysunek (3) dla punktów niestabilnych (pojęcia te i ich znaczenie jest analogiczne dla rodzajów równowagi termodynamicznej: trwałej i chwiejnej, odpowiednio).



Rysunek 2: Ilustracja stabilnego punktu równowagi



Rysunek 3: Ilustracja niestabilnego punktu równowagi

Jak już zostało wspomniane narzucenie warunku początkowego w postaci punktu równowagi implikuje stałe rozwiązanie układu (1). Powstaje naturalne pytanie jak zachowują się rozwiązania owego układu dla warunków początkowych $\mathbf{u}_0 \in M_{ph}$ takich, że

$$\|\mathbf{u}_c - \mathbf{u}_0\|_M \leq \epsilon, \quad (8)$$

gdzie $\epsilon \in \mathbb{R}_+$.

Dla lepszego zrozumienia problemu oraz wytyczenia kierunku dalszych rozważań przytoczony zostanie prosty przykład (w przypadku gdy $M = \mathbb{R}$) ilustrujący zachowanie obrazu otoczenia punktów równowagi w przypadku zaburzenia warunków początkowych oraz przekształcenia \mathbf{f} (poprzez perturbację wektora parametryzującego wspomnianą funkcję).

Przykład 1. *Rozważmy układ postaci*

$$\frac{dx}{dt} = x^2 - 1, \quad (9)$$

z warunkiem początkowym postaci

$$x(0) = x_0 \quad (10)$$

Układ taki można rozwiązać analitycznie. Rodzina krzywych całkowych w tym przypadku określona jest wzorem

$$x(t) = \frac{1 - e^{2C+2t}}{1 + e^{2C+2t}}, \quad (11)$$

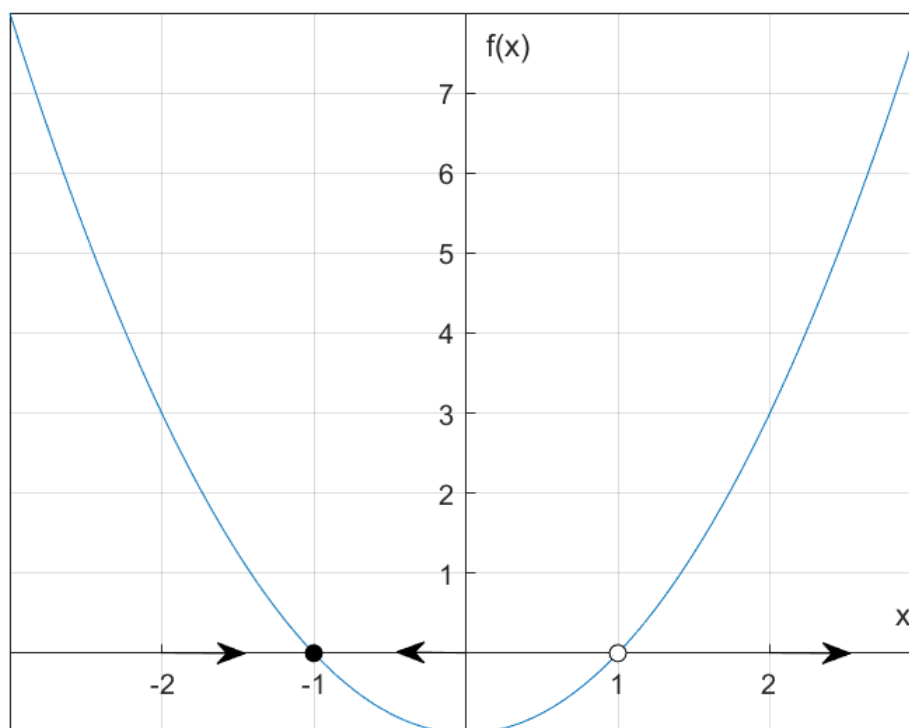
gdzie

$$C = \frac{1}{2} \ln \frac{1 - x_0}{1 + x_0}. \quad (12)$$

Zauważmy jednak, że powyższy wzór ma sens jedynie w przypadku, gdy $x_0 \in (-1, 1)$. Dla $x_0 = -1$ lub $x_0 = 1$ otrzymujemy stałe rozwiązanie równe x_0 (punkt równowagi), zaś przy padek, gdy $\|x_0\| > 1$ należy rozpatrzyć osobno.

Powyższa analiza pokazuje, że jakościowe badanie równań wymaga dodatkowych technik (nawet w przypadku obecności analitycznego rozwiązania). Dużo lepsze efekty uzyskuje się

poprzez podejście geometryczne z użyciem tzw. portretów fazowych. Otrzymuje się je poprzez zwizualizowanie funkcji zadającej dynamikę w rozpatrywanym układzie. Następnie analizuje się ewolucję dowolnego punktu reprezentującego warunek początkowy w zależności od znaku funkcji. Idea ta została przedstawiona na Rysunku (4).



Rysunek 4: Portret fazowy układu z Przykładu 1

Dla dowolnego punktu $x_0 \in (-1, 1)$ odpowiadająca mu wartość funkcji $f(x) = x^2 - 1$ jest ujemna, wobec czego pochodna rozwiązania jest ujemna i w 'następnym kroku czasowym' będzie miało ono niższą wartość. Zostało to przedstawione za pomocą strzałki o grocie skierowanym w lewą stronę. Analogiczne rozumowanie pozwala dojść do wniosku, że dla $x_0 \in (-\infty, -1) \cup (1, \infty)$ rozwiązanie będzie zwiększało swoją wartość wraz z upływem czasu, co ilustrują strzałki skierowane w prawą stronę. Oczywiście dla $x_0 = -1$ lub $x_0 = 1$ rozwiązanie nie będzie się zmieniało, ponieważ są to punkty równowagi. Na koniec zauważmy, że niewielka perturbacja pierwszego z tych punktów 'sprowadzi' rozwiązanie (po pewnym czasie) do warunku początkowego, podczas gdy dowolna zmiana drugiego punktu powoduje rozbieżność rozwiązania. Wobec tego punkty równowagi można podzielić na (lokalnie) stabilne ($x_0 = -1$) i niestabilne ($x_0 = 1$). Formalna definicja tych pojęć podana zostanie później w tym rozdziale.

Powyższy przykład sugeruje następującą definicję.

Definicja 7. Niech $\bar{\mathbf{u}} \in M_{ph}$ będzie rozwiązaniem układu (1). Powiemy, że jest ono stabilne w sensie Lyapunowa o ile $\forall \epsilon > 0 \exists t_0 \geq 0 \exists \eta > 0 \forall t > t_0$ zachodzi implikacja:

$$\|\bar{\mathbf{u}}(t_0) - \mathbf{u}(t_0)\|_M \leq \eta \Rightarrow \|\bar{\mathbf{u}}(t) - \mathbf{u}(t)\|_M \leq \epsilon,$$

gdzie symbolem \mathbf{u} oznaczono dowolne rozwiązanie układu (1).

Ponadto jeśli

$$\lim_{t \rightarrow \infty} \|\bar{\mathbf{u}}(t) - \mathbf{u}(t)\|_X = 0,$$

to rozwiązanie $\bar{\mathbf{u}} \in M_{ph}$ nazywamy asymptotycznie stabilnym.

Taka definicja stabilności jest szeroko stosowana w wielu zastosowaniach technicznym. Dobrym przykładem mogą być układy automatycznej regulacji, w których praktyczne kryteria (takie jak kryterium Rutha lub Hurwitza) wyprowadza się bazując na stabilności określonej w powyższym sensie.

Zazwyczaj oczywiście nie dysponujemy rozwiązaniami. Z tego powodu konieczne są inne techniki służące jakościowemu badaniu równań. Jednym z takich pomysłów jest zdefiniowanie specjalnej funkcji związanej z nazwiskiem Lyapunowa.

Definicja 8. Funkcję Lyapunowa dla układu postaci (1) nazywamy dyfeomorfizm $V : M_{ph} \rightarrow \mathbb{R}$, który spełnia warunki:

- $\forall x \in M_{ph} V(x) \geq 0$,
- $V(u) = 0$ wtedy i tylko wtedy, gdy u jest punktem równowagi,
- dla dowolnego rozwiązania $u(t)$ zachodzi $\frac{d}{dt} V(u(t)) \leq 0$

Zauważmy, że nie trzeba dysponować rozwiązaniem, gdy znaleziona zostanie odpowiadająca mu funkcja Lyapunowa (co zazwyczaj jest zadaniem dużo łatwiejszym). Ostatni warunek Definicji 8 mówi, że funkcja Lyapunowa na krzywych całkowych jest nierosnąca. W związku z tym prawdziwe jest następujące twierdzenie.

Twierdzenie 1. Niech dany będzie układ postaci (1) i niech funkcja prawych stron będzie dyfeomorfizmem. Ponadto niech $0 \in M_{ph}$ i $\mathbf{f}(0) = 0$. Jeśli dla takiego układu istnieje funkcja Lyapunowa, to rozwiązanie $\mathbf{u}(t) \equiv 0$ jest stabilne. Ponadto jeśli $\frac{d}{dt} V(u(t)) < 0$ dla $u \in M_{ph} - \{0\}$, to jest ono asymptotycznie stabilne.

Dowód tego twierdzenia jest dość techniczny. Można go znaleźć w większości podręczników traktujących o równaniach różniczkowych (na przykład [30]). Intuicyjnie jednakże jest ono dość oczywiste. Z założeń poczynionych na początku rozdziału wybór dowolnego warunku początkowego z przestrzeni fazowej determinuje jednoznacznie rozwiązanie tego układu (zależność przyporządkowująca warunkowi początkowemu rozwiązanie jest przekształceniem ciągłym), a funkcja Lyapunowa dla takiego rozwiązania jest nierosnąca. Ponieważ z definicji jest ona nieujemna i przyjmuje zerowe wartości tylko w punktach równowagi, to krzywa całkowita musi pozostać dostatecznie blisko takiego punktu, gdy jest ono stabilnym rozwiązaniem lub też zbiegać do niego w przypadku gdy jest to asymptotycznie stabilne rozwiązanie.

Powyższe uwagi sugerują pewien sposób postrzegania funkcji Lyapunowa. Niech dana będzie funkcja $V(u)$ nazywana dalej potencjałem zdefiniowana równaniem

$$f(u) = -\frac{dV}{du}. \quad (13)$$

Trajktując u jako zmiennej niezależnej t mamy:

$$\frac{dV}{dt} = \frac{dV}{du} \frac{du}{dt}. \quad (14)$$

Referując dalej do układu postaci (1) otrzymujemy:

$$\frac{dV}{dt} = -\left(\frac{dV}{du}\right)^2 \leq 0. \quad (15)$$

Tak zdefiniowany potencjał zachowuje się jak funkcja Lyapunowa. Ta intuicja sugeruje również (zgodnie z oczekiwaniami), że lokalne maksima potencjału odpowiadają niestabilnym punktom równowagi, zaś lokalne minima odpowiadają stabilnym (idea potencjału jest często używana w kontekście fizyki, w szczególności pola sił, które można zdefiniować jako gradient pewnego pola skalarnego określany jest mianem potencjalnego i posiada przy tym szereg specyficznych cech, jak na przykład zachowawczość).

Dotychczas wszelkie rozważania oparte były na rozwiązaniach ustalonego układu, w którym czas jest traktowany jako zmienna niezależna, zaś warunek początkowy jako parametr. Oznaczając rozwiązanie układu (1) z warunkiem początkowym (2) w czasie t przez $u(t; u_s)$ możemy zdefiniować rodzinę dyfeomorfizmów (na mocy założeń dotyczących funkcji $f(u)$ i przestrzeni fazowej M_{ph}) sparametryzowanych zmienną t : $g^t : M_{ph} \rightarrow M_{ph}$ taką, że:

$$g^t(u_s) = u(t; u_s). \quad (16)$$

Fakt, że tak zdefiniowane funkcje są różniczkowalne wynika z gładkiej zależności rozwiązań względem warunków początkowych. Załóżmy od tego momentu, że zmienna $t \in (-\infty; \infty)$, tzn. że zmienna ta może przyjmować dowolne wartości. Wówczas zachodzą następujące definicje.

Definicja 9. *Potokiem układu (1) nazywamy parę (M_{ph}, g^t) , gdzie M_{ph} oznacza przestrzeń fazową, zaś $\{g^t\}_{t \in (-\infty; \infty)}$ oznacza zdefiniowaną wcześniej rodzinę dyfeomorfizmów.*

Definicja 10. *Trajektorię (orbitę) punktu u_s nazywa się obraz przekształcenia $g^t(u_s)$ dla $t \in (-\infty; \infty)$.*

Trajektorię można postrzegać jako ewaluację rozwiązania w konkretnej chwili, traktując przy tym warunek początkowy jako zmienną niezależną. W związku z tym, gdy u_s jest punktem równowagi, to jego orbita jest stała $g^t(u_s) \equiv u_s$.

Przełiśmy od analizy krzywych całkowych jako funkcji czasu do trajektorii układu, które od czasu nie zależą i w naturalny sposób można je zilustrować w przestrzeni fazowej. Jednocześnie mają one wiele wspólnego z rozwiązaniami. Zauważmy najpierw, że sam potok (M_{ph}, g^t) ma strukturę grupy abelowej z działaniem

$$g^{t+s} = g^t g^s. \quad (17)$$

Elementem neutralnym tej grupy jest przekształcenie identycznościowe. Stąd i z jedyności rozwiązań rozpatrywanego układu wynika następujące twierdzenie.

Twierdzenie 2. *Przez każdy punkt przestrzeni fazowej przechodzi dokładnie jedna orbita.*

Takie zachowanie trajektorii upadabnie je do rozwiązań. Ich zaleta przejawia się tym, że z topologicznego punktu widzenia różnych rodzajów orbit jest niewiele, w przeciwieństwie do geometrycznego bogactwa kształtów krzywych całkowych. Mówi o tym następujące twierdzenie.

Twierdzenie 3. *Przy założeniach przyjętych uprzednio orbity potoku generowanego przez układ (1) dzielą się na trzy kategorie:*

- orbity otwarte
- orbity zamknięte
- punkty krytyczne

Dowód tego twierdzenia opiera się istotnie o poprzednim twierdzeniu i o własnościach domkniętych grup addytywnych. Zamiast podawać pełną jego wersję zilustrowane zostaną jego konsekwencje bazując na dwóch przykładach zaczerpniętych z książki [30].

Przykład 2. *Rozważmy równanie*

$$\frac{du}{dt} = u. \quad (18)$$

Trajektoria punktu u_s ma postać $\{e^t u_s\}$, gdzie $t \in (-\infty, \infty)$. W związku z tym dla $u_s > 0$ orbitę stanowi przedział $(0, \infty)$, dla $u_s < 0$ przedział $(-\infty, 0)$, zaś dla $u_s = 0$ trajektoria redukuje się do punktu $\{0\}$

Z przykładu tego można łatwo uogólnić następujący wniosek: w przypadku jednowymiarowej przestrzeni fazowej orbity układów dynamicznych redukują się do punktów lub też do przedziałów (nieograniczonych przynajmniej z jednej strony). W szczególności niemożliwe jest otrzymanie trajektorii okresowej. Sytuację tę zmienia dodatkowy wymiar przestrzeni fazowej, co pokazuje następny przykład.

Przykład 3. *Rozważmy układ postaci*

$$\begin{aligned} \frac{du_1}{dt} &= -u_1 + u_1(1 - u_1^2 - u_2^2) \\ \frac{du_2}{dt} &= u_1 + u_2(1 - u_1^2 - u_2^2) \end{aligned} \quad (19)$$

Rozwiązanie można uzyskać poprzez wprowadzenie współrzędnych biegunowych

$$\begin{aligned} u_1 &= r \cos(\theta) \\ u_2 &= r \sin(\theta), \end{aligned} \quad (20)$$

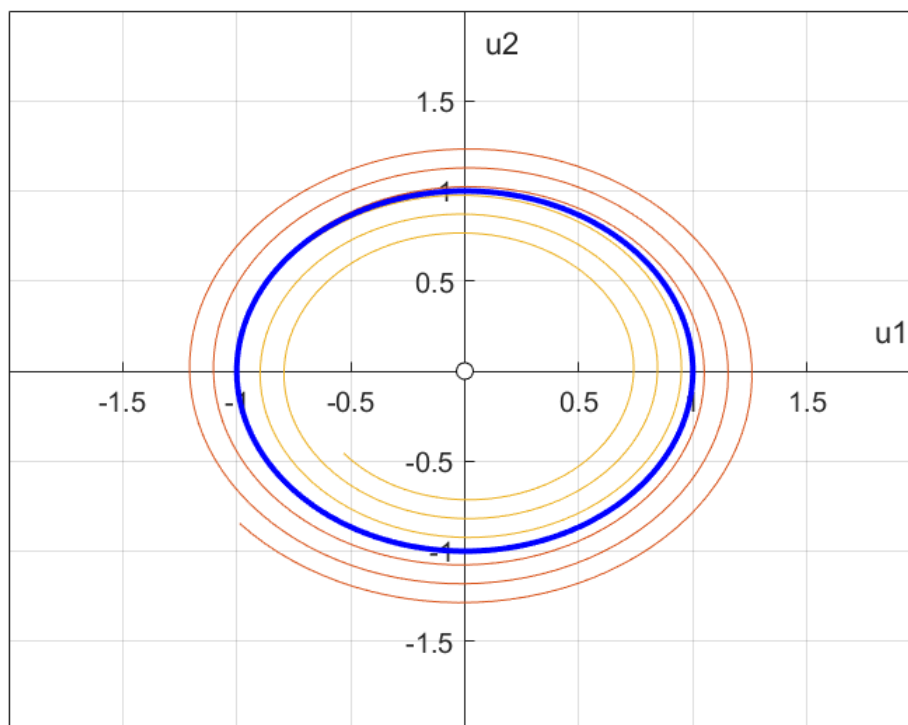
aby otrzymać układ

$$\begin{aligned} \frac{dr}{dt} &= r(1 - r^2) \\ \frac{d\theta}{dt} &= 1 \end{aligned} \quad (21)$$

Trajektorie pierwszego równania mają postać $\left\{ \frac{u_s e^t}{\sqrt{u_s^2 e^{2t} - p^2 + 1}} \right\}$. Dostajemy trzy rodzaje trajektorii:

- punkt równowagi $(0, 0)$,
- orbitę okresową $u_1^1 + u_2^2 = 1$,
- orbity spiralne 'nawijające' się od zewnątrz i od wewnątrz na orbitę okresową.

Zostały one zilustrowane na Rysunku (5), gdzie pogrubioną linią zaznaczono trajektorię okresową.



Rysunek 5: Trajektorie rozważanego układu

Z powyższych przykładów łatwo można dojść do wniosku, że analiza trajektorii w przestrzeni fazowej jest dużo łatwiejsza niż badanie nieskończonej (być może nieprzeliczalnej) liczby rozwiązań danego układu. Dysponując już odpowiednim aparatem pojęciowym możemy przejść do analitycznego ujęcia zjawiska chaosu.

2.3 Chaos deterministyczny

Do tej pory przedstawione zostały pewne ogólne sposoby badania stabilności układów dynamicznych. Obecnie przechodzimy do analizy równań, których rozwiązania (choć nadal są wyznaczone w sposób jednoznaczny) mają bardzo skomplikowane portrety fazowe,

które w praktyce istotnie ograniczają możliwość przewidywania długookresowych zachowań tych rozwiązań. Należy zauważyć, że do tej pory takie układy nie były rozważane, gdyż dotychczas prezentowane były systemy o co najwyżej dwuwymiarowej przestrzeni fazowej, zaś wspomniana cecha ujawnia się gdy liczba wymiarów jest nie mniejsza niż trzy. Dodatkowy wymiar przestrzeni fazowej umożliwia pojawienie się trajektorii o kształtach tak skomplikowanych, że klasyczne metody geometrii pozostają wobec nich bezsilne. Stwierdzą one istotną wskazówkę, iż mamy do czynienia z chaosem (choć nie jest to reguła). Chaos deterministyczny wyróżniają (nieformalnie ujmując) trzy cechy układu:

- determinizm,
- nieliniowość,
- wrażliwość na warunki początkowe.

Pierwsza z nich dotyczy istnienia i jednoznaczności rozwiązania, zaś druga struktury równań zadających dynamikę (suma rozwiązań w ogólności nie musi być rozwiązaniem). Ostatnia cecha dotyczy zachowania się trajektorii wraz z upływem czasu - norma ich różnicy pozostaje nieograniczona pomimo zbliżonych warunków początkowych. Innymi słowy: ustaliliśmy uprzednio warunek początkowy, trajektorie wychodzące z punktów z jego dowolnego otoczenia z czasem pozostają odległe od trajektorii rozpoczynającej się w tymże punkcie³. Wrażliwość układu dynamicznego równie silnie przejawia się w stosunku do zmiany parametrów funkcji $f : M_{ph} \mapsto M_{ph}$ przy ustalonym warunku początkowym. Występują istotne różnice pomiędzy wspomnianymi wrażliwościami, które mają bezpośrednie konsekwencje w analizie układu - zagadnienie to zostanie poruszone przy okazji omawiania *Metody Trajektorii Cienia*. Dodatkowo, często wspomina się o długookresowym, aperiodycznym zachowaniu rozwiązań w kontekście chaosu.

Z powyższych ustaleń można dojść do wniosku, że o takim układzie niemalże nic konkretnego nie można powiedzieć. Należy pamiętać o tym, że z punktu widzenia metod numerycznych zawsze istnieją ograniczenia (precyzja maszynowa, błędy zaokrągleń, itp.), co wobec ostatniej z wyżej wspomnianych cech również nie poprawia sytuacji. W matematyce oraz w naukach przyrodniczych, w sytuacjach podobnych do zainstniałej, dobrym pomysłem jest próba znalezienia niezmienników - stałych elementów w rozważanych systemach. Okazuje się, że dla układów dynamicznych pomocnym jest zdefiniowanie niezmienniczego ze względu na dynamikę zbioru zwanego atraktorem.

Definicja 11. Zbiór $\Lambda \subset M_{ph}$ nazywamy atraktorem układu (1), jeżeli:

- Λ jest niezmienniczy względem f , tj. $f(\Lambda) \subset \Lambda$,
- Istnieje zbiór otwarty $U \subset M_{ph}$ taki, że $\forall u \in M_{ph}$ zachodzi:

$$u(0) \in U \implies \lim_{t \rightarrow \infty} d(\Lambda, u(t)) = 0,$$
gdzie przez $d(,)$ oznaczono metrykę w M ,

³Mówiąc bardziej precyzyjnie - cecha ta zachodzi dla μ -prawie wszystkich punktów z otoczenia.

- Λ jest minimalny w tym sensie, że żaden jego właściwy podzbiór nie spełnia powyższych dwóch warunków⁴.

Zgodnie z powyższą definicją można przyjąć, iż atraktor 'przyciąga' trajektorie zaczynające się z konkretnych punktów przestrzeni fazowej. Zbiór takich warunków początkowych dla danego atraktora nazywa się jego *basenem atrakcji*. Pojęcia te są ze sobą blisko związane i dlatego w dalszej części pracy uwaga zostanie skupiona głównie na tym pierwszym.

Patrząc na dotychczas omówione przykłady można dojść do wniosku, że różne proste postaci atraktorów już się pojawiały (patrz). Różnorodność ich geometrii była jednak mocno ograniczona (punkt lub też pewna regularna krzywa). Jedną z najbardziej zaskakujących cech teorii chaosu jest praktyczny brak ograniczeń na topologiczną złożoność atraktorów (pomimo warunków regularności nałożonych na równanie opisujące dynamikę!). Skomplikowany portret fazowy atraktora nie jest też wyjątkiem ani nawet sytuacją rzadką, a raczej regułą. W pewnym sensie można powiedzieć, że jest to warunek chaotyczności układu (sprecyzujemy to stwierdzenie w dalszej części tego paragrafu). Szukając cech wspólnych takich atraktorów można odnaleźć w literaturze dwa terminy stosowane zamiennie: atraktor dziwny oraz fraktal. Ten drugi w geometrii nazywany jest obiektem o 'ułamkowym' wymiarze, który jest samopodobny w każdej skali. Należy tutaj zaznaczyć, że nie ma jednej, powszechnie przyjętej definicji fraktala. Aby jednak nakreślić pewne intuicje możemy przyjąć, iż fraktal stanowi zbiór, którego wymiar Hausdorffa jest istotnie większy od jego wymiaru geometrycznego. Pojęcia te celowo nie zostają uściślone odsyłając zainteresowanego czytelnika do literatury [35], [21]. Zamiast tego w dalszej części zwizualizowane zostaną popularne fraktale wywodzące się z układów chaotycznych. Zanim jednak to nastąpi potrzebna konieczne jest wprowadzenie jeszcze jednego pojęcia.

Definicja 12. *Bifurkacją układu dynamicznego (1) nazywamy jakościową zmianę zachowania jego rozwiązań. Parametr $s_0 \in \mathbb{R}^n$, dla którego następuje taka zmiana nazywany jest parametrem bifurkacji.*

Przykład 4. *Aby zilustrować powyższą definicję powtórzmy Przykład 1. Tym razem rozważmy sparametryzowane równanie postaci*

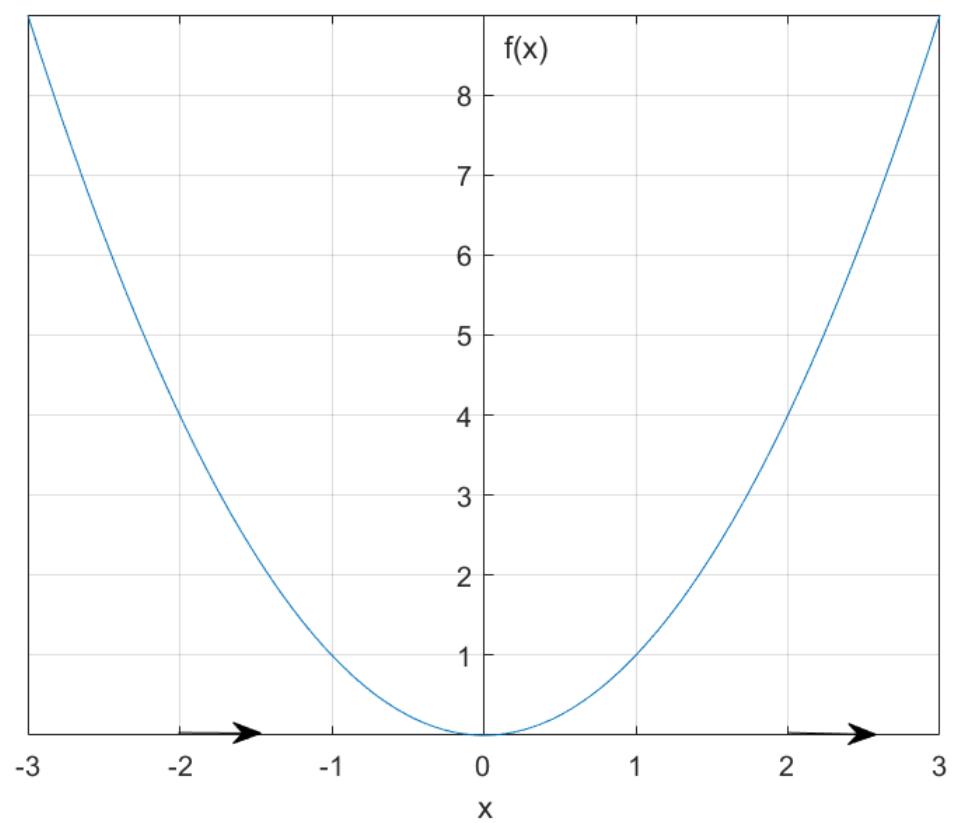
$$\frac{dx}{dt} = x^2 - a, \quad (22)$$

z warunkiem początkowym postaci

$$x(0) = x_0, \quad (23)$$

gdzie $a \in \mathbb{R}$. Zauważmy, że kształt krzywej obrazującej funkcję prawej strony pozostaje niezmienny - zmieniają się jedynie punkty przecięcia z osiami współrzędnych. Dla $a > 0$ sytuacja odpowiada Rysunkowi (z dokładnością do miejsc zerowych). Odpowiadające portrety dla $a = 0$ oraz $a = -1$ przedstawione zostały na Rysunkach (6) oraz (7), odpowiednio.

⁴Ze względu na trudności w dowodzeniu poprawności tego warunku w większości przypadków jest on pomijany.



Rysunek 6: Portret fazowy układu z Przykładu 5 dla $a=0$



Rysunek 7: Portret fazowy układu z Przykładu 5 dla $a < 0$

Zauważmy co stało się z punktami stabilnymi (w świetle nowo wprowadzonej nomenklatury możemy nazwać je również atraktorami, chociaż pojęcie to zarezerwowane jest dla bardziej skomplikowanych geometrycznie zbiorów). Wraz ze zmniejszaniem się parametru a punkty stabilne zbiegały do zera z różnych stron, aż w granicznym przypadku $a = 0$ scalały się w jednym miejscu. Należy zwrócić uwagę, że taki nazywany jest punktem częściowo stabilnym (przyciąga punkty ze swojego lewostronnego otoczenia i odpycha te z prawej strony, co symbolizują strzałki o zgodnie skierowanych grotach). Dalsze zmniejszanie parametru powoduje zanik punktów stabilnych - rozwiązanie z dowolnym warunkiem początkowym pozostaje niestabilne w sensie Lyapunowa.

Aby dać jeszcze lepsze zrozumienie pojęcia bifurkacji (jest to istotne, bo takowa pojawiająca się z coraz większą częstotliwością jest stanowi najpopularniejszą 'drogę' do powstania chaosu) zostaną przedstawione najpopularniejsze układy dynamiczne o skomplikowanym zachowaniu wraz ze zmianą pewnego parametru kontrolnego. Dodatkowo dla wybranych z nich przedstawione zostaną ich atraktory.

Bifurkacja Feigenbauma Numeryczna realizacja równań różniczkowych zwyczajnych polega przeważnie na wyznaczeniu wartości poszukiwanej funkcji w kolejnych krokach

czasowych bazując na jednym lub kilku wartości w krokach poprzednich. Jest to więc procedura postaci

$$u_{n+1} = f(u_n), \quad (24)$$

gdzie u_n oznacza wartość poszukiwanej funkcji (być może wektorowej) w n -tym kroku czasowym, zaś $f()$ oznacza funkcję zadającą dynamikę (przy zadanym warunku początkowym u_0). Okazuje się, że pewne postaci funkcji f implikują skomplikowane zachowanie ciągu $\{u_n\}_{n \in \mathbb{N}}$. Jako przykład rozważmy sytuację, w której u_n jest funkcją skalarną oraz

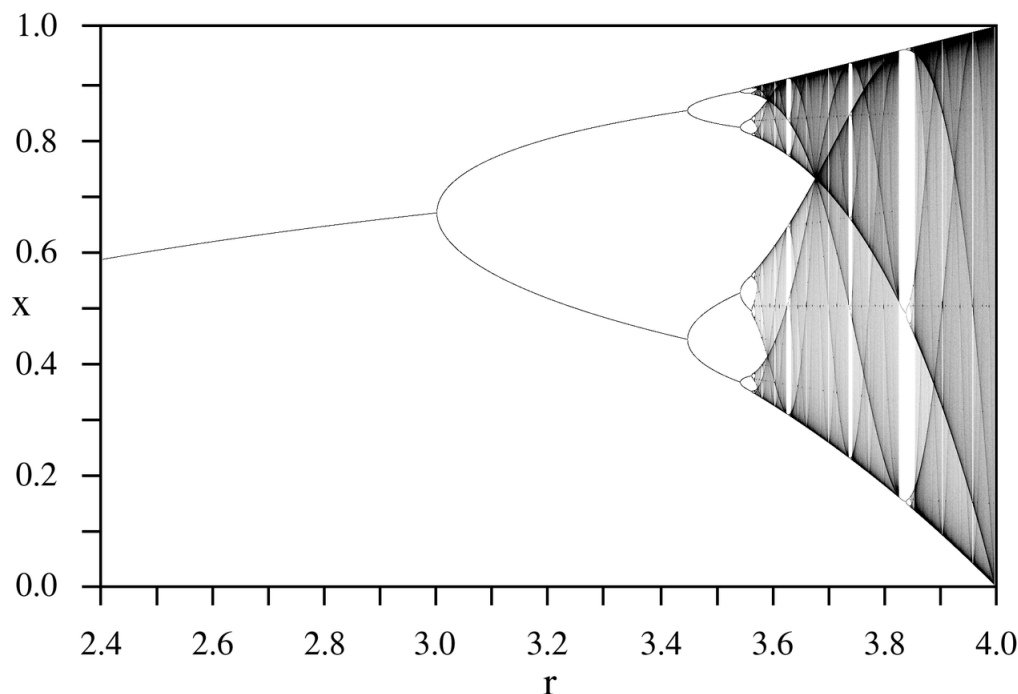
$$f(u_n) = ru_n(1 - u_n), \quad (25)$$

gdzie $r \in \mathbb{R}$. Jest to tzw. *odwzorowanie logistyczne* (zauważmy przy okazji, iż f jest nieliniową funkcją swojego argumentu). Połóżmy dodatkowo $r \in [0, 4]$. Łatwo wówczas zauważyć, że $f([0, 1]) = [0, 1]$. Dodatkowo, dla $r \in (0, 1]$ punkt $u^* = 0$ jest atraktorem, zaś jego basenem atrakcji jest cały przedział $[0, 1]$, w związku z tym wszystkie trajektorie zbiegają do tego punktu, bez względu na wartość warunku początkowego u_0 , co nie stanowi interesującego nas zachowania. Załóżmy więc, że $r > 1$. Zauważmy natychmiast, że $u^* = 0$ stanowi w oczywisty sposób punkt stały tego odwzorowania. Dodatkowo, dla $r > 1$ pojawia się kolejny: $u^{**} = 1 - \frac{1}{r}$. Punkty takie mogą stanowić atraktor o ile 'przyciągają' punkty z pewnego swojego otoczenia, co praktycznie oznacza, iż pochodna w tym punkcie ma moduł mniejszy od jedności. Pochodna funkcji logistycznej wyraża się wzorem $f'(u) = r(1 - 2u)$, w związku z czym $u^* = 0$ nie jest atraktorem ($r > 1$), zaś $u^{**} = 1 - \frac{1}{r}$ spełnia ten warunek dla $1 < r < 3$. Atraktor ten 'przesuwa' się wraz ze zmianą wartości parametru kontrolnego r od $u^{**} = 0$ do wartości $u^{**} = \frac{2}{3}$.

Dla $r > 3$ punkt $u^{**} = 1 - \frac{1}{r}$ nadal jest punktem stałym, jednakże pochodna nie spełnia wspomnianego warunku i punkt ten 'odpycha' punkty z dowolnego swojego otoczenia. Pojawia się wówczas dwuelementowy cykl przyciągający, tj. dwa punktu u_1, u_2 wzajemnie przekształcane na siebie, przy czym każdy punkt w ich otoczeniu po dwóch iteracjach zbliża się do takiego cyklu (oba te punkty stają się atraktorami odwzorowanie złożonego $f(f(u))$). Taka sytuacja jest obecna przy warunku $3 < r < 1 + \sqrt{6}$. Po przekroczeniu prawej nierówności sytuacja się powtarza i powstaje atraktor czteroelementowy. Punkty z sąsiedztwa 'skaczą' po coraz to bliższych otoczeniach elementów tego atraktora, zachowując przy tym zawsze tę samą kolejność (i znów otrzymujemy pojedynczy atraktor dla złożenia $f(f(f(f(u))))$).

Sytuacja ta pojawia się coraz częściej (wraz ze zmianą parametru r), przy czym liczba elementów atraktora podwaja się za każdym razem i zawsze stanowi on jednoelementowy atraktor złożenia funkcji f 2^n razy ze sobą.

Zachowanie takie jest jednak wciąż przewidywalne (w pewnym sensie). Sytuacja zmienia się dla granicznej wartości [35] $r_{\text{inf}} \approx 3,5699456$. Przypadek ten charakteryzuje się cyklem o nieskończonej liczbie elementów, zaś przekroczenie tej wartości powoduje prawdziwy chaos - pojawiają się cykle o dowolnej liczbie elementów, a także odcinki, odcinki pozbawione atraktorów. Pełny diagram bifurkacji ilustruje Rysunek (8).



Rysunek 8: Diagram bifurkacji Feigenbauma

Dla $r = 4$ otrzymujemy jednorodny chaos - iteracje pokrywają odcinek $[0, 1]$. Podsumowując powyższe rozważania można stwierdzić, iż pierwsze pojawienie się bifurkacji ($r = 1$) pozwoliło na 'kontrolowaną' zmianę jakościową rozwiązań rozważanego równania. Kolejne bifurkacje (poczynając od $r = 3$) pojawiające się z coraz większą częstotliwością doprowadziły do chaosu poprzez tzw. podwajanie chaosu (z tym też związana jest etymologia słowa bifurkacja - rozgałęzienie się jakiejś całości na dwie części [przypis]). W ramach prac nad tym (wyjątkowo prosto wyglądającym) procesem iteracyjnym wykryto wiele prawidłowości. Najważniejszą z nich jest tzw. stała Feigenbauma. Wyraża ona stosunek odległości od kolejnych punktów bifurkacji (dla $r > 3$). Interesującym jest fakt, że analogiczna stała (wraz z zachowaniem się układu odwzorowan) o tej samej wartości jest obecna dla innych postaci funkcji f (np dla $f(u_n) = r \sin(u_n)$). Jest to przesłanka, iż mechanizm ten jest w pewnym stopniu uniwersalny. Po dokładniejszym zbadaniu okazuje się, iż w istocie jest to uniwersalna stała gładkich układów dynamicznych, których (skalarna) funkcja dynamiki jest unimodalna w rozważanym przedziale określoności.

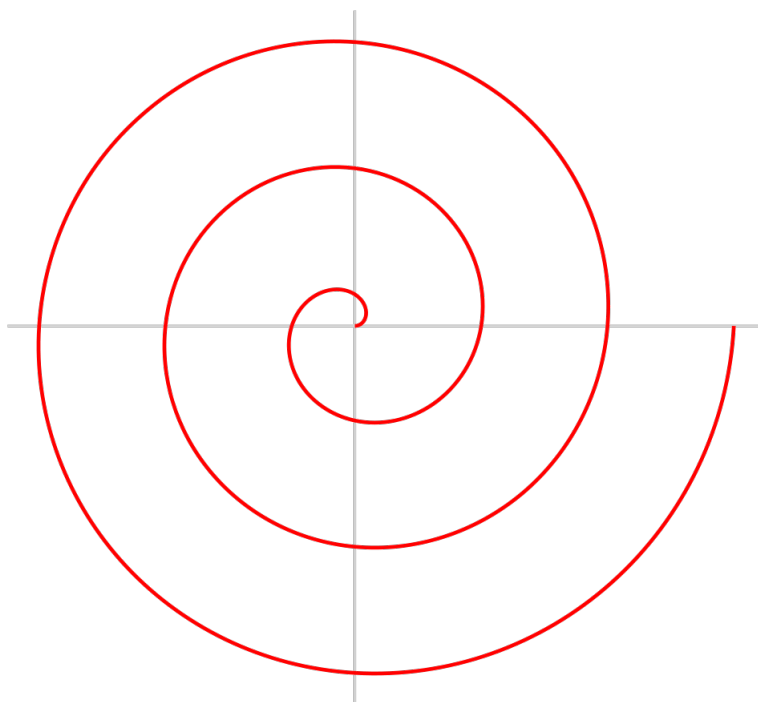
Bifurkacja Hopfa Wspomniane bifurkacje Feigenbauma stanowią klasyczny przykład dla układów dyskretnych. W przypadku systemów ciągłych taki szandarowy układ ma postać

$$\begin{aligned} \frac{dr}{dt} &= \lambda r(1 - r^2) \\ \frac{d\phi}{dt} &= \omega \end{aligned} \quad (26)$$

Jest to równanie ruchu obrotowego punktu w płaszczyźnie (ω stanowi prędkość obrotową, r to promień, ϕ kąt tworzącej z ustaloną półosią układu współrzędnych, zaś λ to parametr

kontrolny). Zauważmy przy okazji, że układ ten to ogólniejszy przypadek dla Przykładu 4, w którym $\omega = 1$ oraz $\lambda = 1$.

Łatwo dostrzec, że dla $\lambda < 0$ punkt $(0, 0)$ jest atraktorem, bez względu na warunek początkowy. Dzieje się tak, gdyż wyrażenie $\lambda r(1 - r^2)$ jest ujemne, a więc promień będzie zmniejszał swoją wartość, przy czym intensywność tych zmian będzie malała. W związku z tym dowolna trajektoria będzie się 'zawijać' ku środkowi układu współrzędnych - Rysunek (9).



Rysunek 9: Przykładowa trajektoria dla układu Hopfa

Sytuacja w przypadku $\lambda > 0$ generuje atraktory cykliczne w postaci okręgów o promieniach $r = \sqrt{\lambda}$. Trajektorie odpowiadające punktom zaczynającym się na zewnątrz takiego okręgu nawijają się nań, zaś te zaczynające się wewnątrz - stopniowo się do niego zbliżają (Rysunek)

Bifurkacja Henona Kolejnym dwuwymiarowym układem posiadającym atraktor fraktalnej natury jest model Henona opisany równaniami dyskretnymi:

$$\begin{aligned} x_{n+1} &= 1 - ax_n^2 + y_n \\ y_{n+1} &= bx_n \end{aligned} \tag{27}$$

gdzie $a, b \in \mathbb{R}$ stanowią parametry modelu. Jakobian przekształcenia wynosi $|b|$, w związku z czym dla $|b| < 1$ otrzymujemy kontrakcję, które na mocy twierdzenia Banacha ma dokładnie jeden punkt stały stanowiący atraktor punktowy

Dla $|b| > 1$ układ ten można rozwiązać rozbijając każdą iterację na trzy kroki stanowiące elementarne operacje geometryczne:

- zginanie powierzchni przy zachowaniu jej miary:

$$\begin{aligned} x' &= x \\ y' &= 1 - ax^2 + y \end{aligned} \quad (28)$$

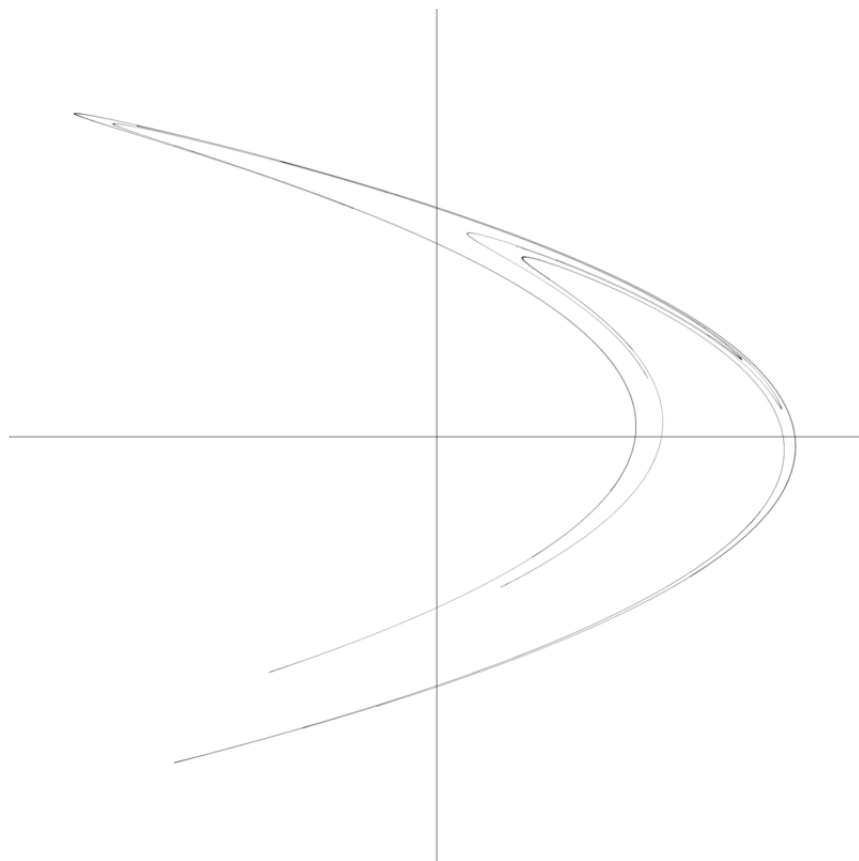
- skracanie odległości między punktami wzdłuż osi odciętych

$$\begin{aligned} x'' &= bx' \\ y'' &= y' \end{aligned} \quad (29)$$

- obrót o kąt prosty

$$\begin{aligned} x''' &= y'' \\ y''' &= -x'' \end{aligned} \quad (30)$$

Oczywiście parametrem bifurkacji jest w tym przypadku $b = 1$. Atraktor tego układu przedstawia Rysunek (10).



Rysunek 10: Atraktor Henona

Bifurkacja Rosslera Ostatnim klasycznym przykładem będzie trójwymiarowy układ Otto Rosslera. Ma on następującą postać:

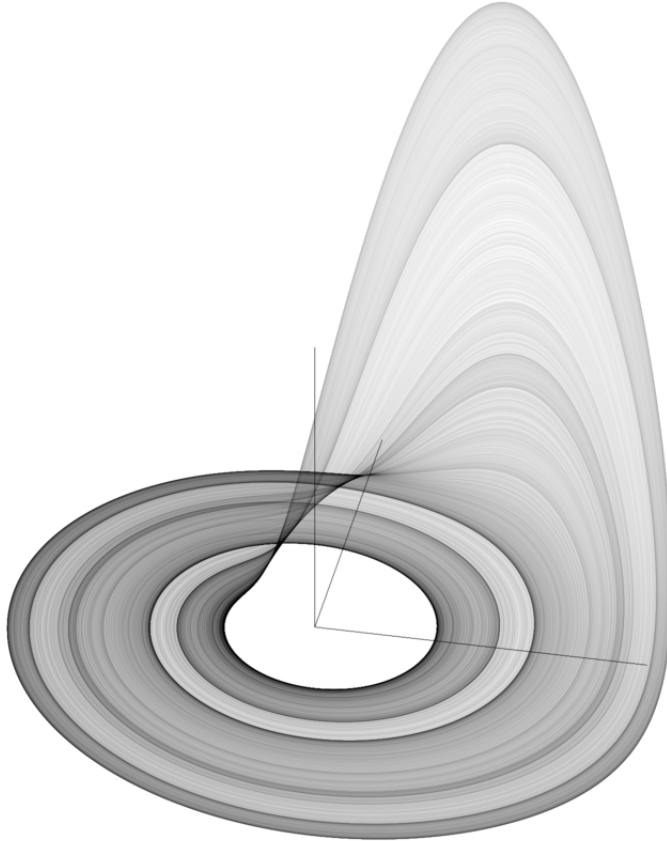
$$\begin{aligned}\frac{dx}{dt} &= -y - z \\ \frac{dy}{dt} &= x + ay \\ \frac{dz}{dt} &= bx - cz + xz\end{aligned}\quad (31)$$

Parametry $a, b, c \in \mathbb{R}_+$ warunkują zachowanie się rozwiązań.

Łatwo zauważyć, iż kłopotliwe jest ostatnie z równań ze względu na obecność członu nieliniowego. Można spróbować zrozumieć dynamikę poprzez tymczasowe usunięcie tego równania (formalnie $z = 0$). Mamy wówczas

$$\begin{aligned}\frac{dx}{dt} &= -y \\ \frac{dy}{dt} &= x + ay\end{aligned}\quad (32)$$

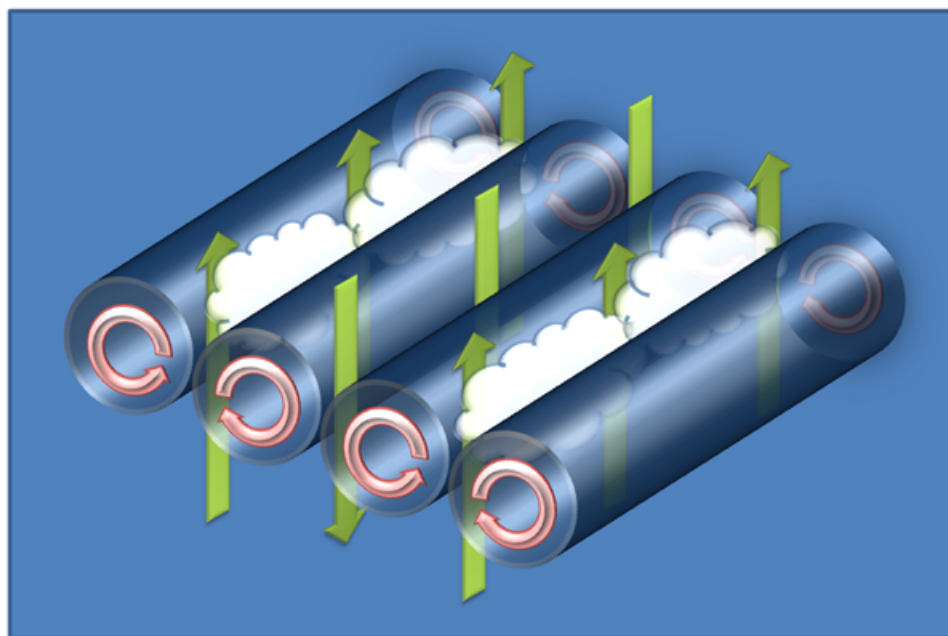
Dla $a \in (0, 2)$ otrzymujemy ponownie spiralę, ale tym razem rozwijającą się. Dodanie ostatniego równania spowoduje 'wygięcie' się owej spirali (człon nieliniowy), przy czym będzie ono stabilne o ile $x < c$. Wówczas wysokość jego osiąga $z = \frac{b}{x-c}$. Dalszy wzrost współrzędnej x jest kontrolowany przez człon $(-z)$ w pierwszym równaniu i trajektoria wygina się w drugą stronę powracając na płaszczyznę $z = 0$. Kształt atraktora ilustruje Rysunek (11).



Rysunek 11: Atraktor Rosslera

Sama fizyka również dostarcza inspiracji do tworzenia modeli o nieprzewidywalnych zachowaniach. Układem chemicznym o takiej własności jest reakcja Biełousowa-Żabotyńskiego. Hydrodynamika nie jest tu wyjątkiem - przejścia do chaosu jest w tym przypadku na tyle popularne, iż zyskało swoją nazwę: turbulencja.

Komórki Bernada Ekperyment, który pozwolił odkryć to zjawisko miał za zadanie zbadać mechanizm wymiany ciepła pomiędzy warstwami wody znajdującej się w naczyniu podgrzewanym od dołu. Parametrem kontrolnym jest w tym przypadku gęstość strumienia ciepła przekazywana cieczi. Przy niewielkiej ilości ciepła w jednostce czasu wymiana ciepła w wodzie następuje na skutek wymiany pędu między jej cząsteczkami. Jest to zatem poziom mikroskopowy - następuje stopniowe podgrzewanie kolejnych warstw. Od pewnej krytycznej wartości (odpowiednik parametru bifurkacji) następuje destabilizacja znajdującej się w naczyniu cieczi. Tworzą się struktury dynamiczne - wirujące przeciwległe rolki zwane komórkami Bernada. Mechanizm wymiany ciepła zmienia się w sposób jakościowy - następuje makroskopowy ruch płynu w objętości wywołany dużymi różnicami gęstości. Mówi się wówczas o konwekcji naturalnej. Sytuację ilustruje Rysunek (12).



Rysunek 12: Schemat komórek Bernada

Pewną analogię stanowi tu mechaniczne zjawisko zwane wyboczeniem (analogiem sił jest tutaj oczywiście różnica temperatur, zaś samo zjawisko wyboczenia reprezentowane jest przez utratę stabilności cieczi). Pewne ilościowe wnioski na ten temat Czytelnik znajdzie w klasycznej pozycji [20].

Model matematyczny takiego zjawiska opisywany jest poprzez układ równań Naviera-Stokesa sprzężony z równaniem opisującym wymianę ciepła. Jest on zatem bardzo skomplikowany i w celu jego rozwiązania stosuje się pewne uproszczenia. Najprostszym może

być przyjęcie nieskończonej długości naczynia, co implikuje brak wymiany ciepła wzdłuż rolek konwekcyjnych - model taki stanowi podstawę innego, bardzo znanego układu, którego autorem jest Edward Lorenz. Model ten jest sztandarowym przykładem układu o skomplikowanym zachowaniu, zaś jego atraktor stał się symbolem teorii chaosu (Rysunek). Znacznie więcej informacji na temat układu Lorenza wraz z analizą wrażliwości funkcji zdefiniowanej dla tego systemu zawiera Rozdział 5.

Powyższe rozważania przekonują, iż do listy cech charakteryzujących każdy układ dynamiczny można dopisać dodatkowe dwie pozycje:

- nierozkładalność przestrzeni fazowej na dwie, dynamicznie rozłączne, niepuste podprzestrzenie,
- gęstość punktów okresowych przestrzeni fazowej.

Istota chaosu Poprzez wymienione przykłady stworzyliśmy pewien ogląd na układy dynamiczne. Zauważmy, że we wszystkich układach (modelach matematycznych lub fizycznych) istnieje generyczny schemat otrzymania chaosu. Polega on na parametryzowaniu układu w celu wzmocnienia wpływu konkretnych członów (najczęściej nieliniowych) na zachowanie rozwiązań. Jest to dość ogólne spostrzeżenie. Można wyróżnić wiele podtypów dróg do chaosu i nie istnieje na ten moment ogólna teoria ich istnienia. Najczęściej wyróżnia się trzy scenariusze (bifurkacyjne):

- poprzez podwajanie okresu, tj. według schematu Feigenbauma opisanego wcześniej w tym podrozdziale,
- poprzez tzw. *intermitencję*, czyli powstawanie podokresów chaotycznego zachowania - im większa wartość parametru kontrolnego tym (statystycznie) mniejsza długość okresów ruchu uporządkowanego,
- schemat Ruelle'a-Takens-Newhouse'a, który w istocie opiera się o omawianą w tym podrozdziale bifurkację Hopfa, jednakże w jej trójwymiarowej wersji (w przypadku dwuwymiarowej przestrzeni fazowej możliwe są tylko punkty i cykle graniczne), dla której cyklem granicznym jest powierzchnia torusa.

Przedstawione schematy powstawania chaosu ilustrują czym samo zjawisko się objawia. Również na gruncie intuicji jest ono zrozumiałe, charakteryzuje się ono bowiem skomplikowanym zachowaniem rozwiązań oraz ich wrażliwością na warunki początkowe (lub parametru modelu) w ramach deterministycznych układów nieliniowych.

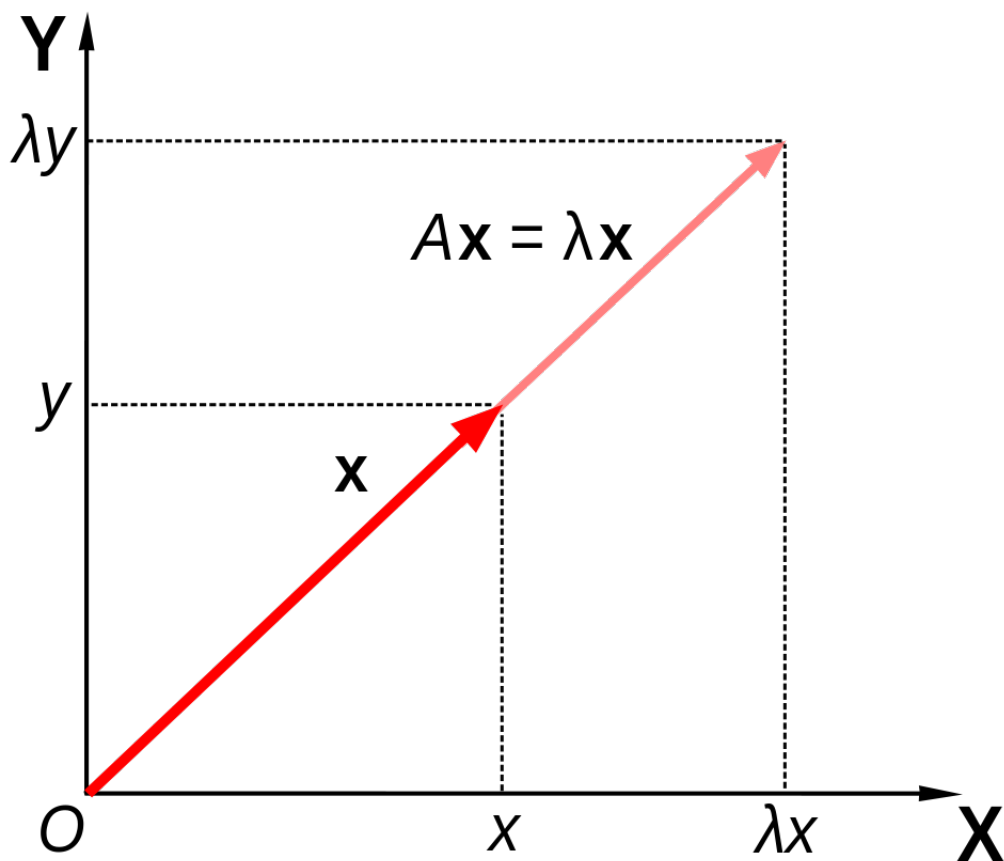
Pozostało jeszcze scharakteryzować klasę takich systemów poprzez pewną jakościową miarę - taka charakterystyka pozwala stwierdzić jak 'bardzo chaotyczny' jest układ lub też rozstrzygnąć który układ jest 'bardziej chaotyczny'. Aby wygenerować wspomnianą miarę i tym samym dobrze zrozumieć jej znaczenie zaczniemy od najprostrzej klasy odwzorowań - funkcji liniowych określonych na przestrzeniach skończonego wymiaru. W tym przypadku układ (1) sprowadza się do postaci

$$\mathbf{u}_{n+1} = \mathbf{A}\mathbf{u}_n, \quad (33)$$

gdzie $A \in M^{N \times N}$ jest macierzą przekształcenia (o stałych współczynnikach), zaś $N \in \mathbb{N}$ oznacza wymiar przestrzeni fazowej. Tutaj ponownie będziemy zainteresowani niezmiennikami - w tym przypadku interesujące są wektory, których kierunek jest niezmienniczy ze względu na transformację A , tj wektory niezerowe $\mathbf{v} \in \mathbb{R}^N$ takie, że

$$A\mathbf{v} = \lambda\mathbf{v}. \quad (34)$$

Przyjmując nomenklaturę znaną z algebry liniowej możemy mówić o takich wektorach $\mathbf{v} \in \mathbb{R}^N$ jako w wektorach własnych macierzy A (o ile takie wektory istnieją). Odpowiadające im skalary $\lambda \in \mathbb{R}$ nazywają się wartościami własnymi, których wartość modułu większa od jedności oznacza, iż punkty leżące na kierunku wektora własnego oddalają się od siebie, zaś wartości te mniejsze od jedności wskazują na zbliżanie się takich punktów względem siebie. Sytuację dla przypadku dwuwymiarowego (tj. $v = (x, y) \in \mathbb{R}^2$, gdzie $\lambda > 1$) ilustruje Rysunek (13).



Rysunek 13: Ilustracja zachowania wektorów własnych

Oczywiście różne wektory własne mogą różnie się deformować pod wpływem przekształcenia. Mówi się zatem, że danej wartości własnej odpowiada wektor własny, przy czym takie wektory są zawsze liniowo niezależne, a zatem tworzą podprzestrzeń liniową przestrzeni stanowiącej dziedzinę rozważanego przekształcenia. W tym miejscu należy

podkreślić, że wektory te w ogólności nie stanowią bazy całej przestrzeni (wymiar generowanej przez nie podprzestrzeni jest istotnie mniejszy od wymiaru całej przestrzeni) oraz wektory te nie są w ogólności ortogonalne (taka sytuacja zachodzi, gdy przekształcenie jest samosprężone, a więc w naszym przypadku, gdy macierz A jest symetryczna).

Informacja na temat tych charakterystycznych obiektów ma ogromną wartość. Mówi ona bowiem jak punkty znajdujące się w sąsiedztwie rozpatrywanego punktu trajektorii (wzdłuż wektorów własnych) będą się względe niego zachowywały. Jest to zatem ilościowa informacja na temat dynamiki układu, której poszukujemy! Wystarczy teraz tę sytuację uogólnić na przypadek dowolnego przekształcenia $f: M_{ph} \rightarrow M_{ph}$. Zakładając na chwilę, że $\dim(M_{ph}) < \infty$ i pamiętając o regularności f przekonujemy się, że odpowiednikiem macierzy A jest macierz gradientu (jacobiego) funkcji f oznaczanej dalej jako Df . Jest to zrozumiałe, gdyż funkcja będąca dyfeomorfizmem jest dobrze aproksymowana funkcją liniową, przy czym cecha ta ma charakter lokalny, co oczywiście stanowi przeszkodę z punktu widzenia numerycznego wyznaczania takich macierzy, gdy tych wymiarów jest dużo. Zauważmy, że i w tym przypadku otrzymujemy wektory i wartości własne takich macierzy, przy czym mają one dokładnie tą samą interpretację. Możemy zatem przyjąć następującą definicję

Definicja 13. Niech $\lambda_i \in \mathbb{R}_+$ stanowią wartości własne macierzy gradientu Df . Wówczas $\ln(\lambda_i) \in \mathbb{R}$ nazywamy wykładnikami Lyapunova.

Wobec takiej definicji znamy już podstawowe własności wykładników Lyapunova. Możemy już teraz zdefiniować formalnie układ przejawiający zachowania chaotyczne.

Definicja 14. Układ dynamiczny postaci (1) nazywamy chaotycznym, gdy posiada przynajmniej jeden dodatni wykładnik Lyapunova.

Definicja taka nie jest przyjęta za jedyną słuszną. Wiele innych można znaleźć w bogatej encyklopedii poświęconej zjawisku chaosu [43]. Niemniej dla celów pragmatycznych jest ona najbardziej dogodna.

Można by sądzić, że odpowiadające omawianym wartościom własnym wektory własne nazywamy wektorami (kowariantnymi) Lyapunova, jednak nie jest prawda. Wyznaczenie wektorów własnych macierzy gradientu wymaga jedynie informacji o lokalnej dynamice, podczas gdy wspomniane wektory Lyapunova zależą od macierzy gradientu we wszystkich chwilach.

W następnym podrozdziale przedstawione zostanie narzędzie, które wyróżni klasę układów strukturalnie stabilnych.

2.4 Lemat cienia

Na koniec tego rozdziału wprowadzimy lemat, który w pewnym sensie gwarantuje, iż przedstawiona w dalszej części pracy metoda jest deterministyczna ⁵. Wprowadźmy najpierw kilka pojęć pomocniczych. Skupimy się przy nich na układach dyskretnych, co

⁵Mówiąc nieco ściślej - trajektoria cienia zawsze istnieje przy przyjmowanych założeniach, ale nie zawsze jest ona reprezentatywna, o czym będzie mowa w późniejszych rozdziałach

Definicja 15. Trajektorię (orbitę) dynamicznego układu (24) z czasem dyskretnym nazywamy zbiór

$$O(u, f) = \{f^k(u) : k \in \mathbb{Z}\}$$

Zauważmy, że $u \in M_{ph}$ oznacza warunek początkowy, a trajektoria określona jest dla dodatnich i ujemnych wskaźników iteracyjnych, zgodnie z poprzednimi uwagami. Można wobec tego wprowadzić dodatnią i ujemną półtrajektorię.

$$\begin{aligned} O^+(u, f) &= \{f^k(u) : k \in \mathbb{N}\} \\ O^-(u, f) &= \{f^k(u) : k \in \mathbb{Z} - \mathbb{N}\} \end{aligned} \quad (35)$$

Oczywiście nic nie stoi na przeszkodzie, aby wprowadzić trajektorię dla zbiorów:

$$O(U, f) = \{f^k(U) : k \in \mathbb{Z}\}, \quad (36)$$

gdzie $U \subset M_{ph}$. Pozwala to na badanie lokalnej dynamiki.

Kluczowym pojęciem tego rozdziału jest pseudotrajektoria.

Definicja 16. Niech $\delta \in \mathbb{R}_+$ δ -pseudotrajektorią (δ -pseudoorbitą) dynamicznego układu (24) z czasem dyskretnym nazywamy zbiór

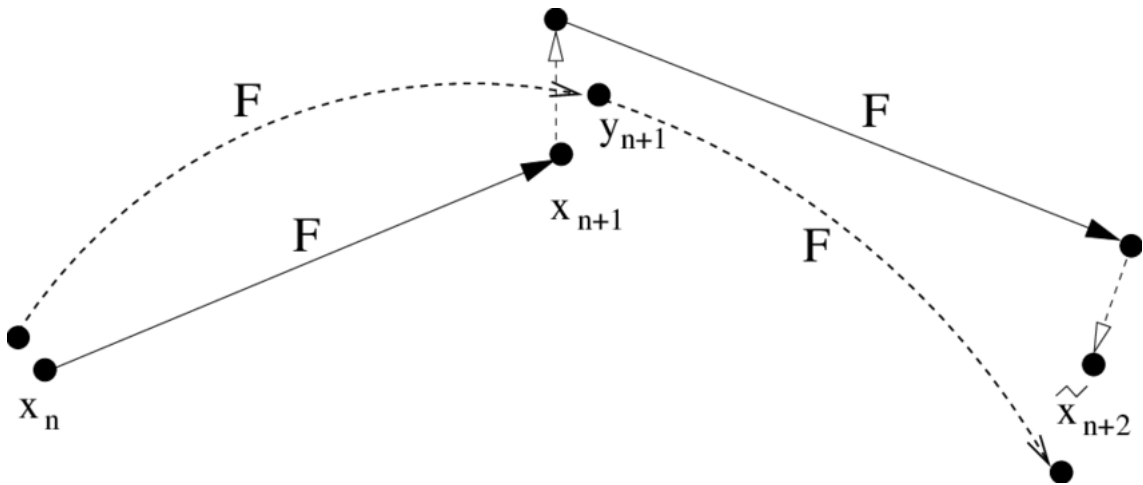
$$\xi = \{u_k \in M_{ph} : k \in \mathbb{Z}\},$$

o ile

$$d(u_{k+1}, f(u_k)) < \delta, \quad k \in \mathbb{Z},$$

gdzie d oznacza metrykę przestrzeni metrycznej M_{ph} .

Widzimy (Rysunek (14)), że punkty δ -pseudotrajektorii są blisko punktami prawdziwej trajektorii jedynie lokalnie, tzn. przechodząc od zadanego punktu przestrzeni fazowej do jego obrazu (zgodnie z dynamiką układu), znajdziemy się δ -otoczeniu punktu δ -pseudoorbity.



Rysunek 14: Ilustracja pseudotrajektorii dla układu punktów w płaszczyźnie.

Istotną cechą układów dynamicznych jest globalny odpowiednik tej własności, która w literaturze nosi nazwę *własności cienia* (ang. *shadowing property*).

Definicja 17. Powiemy, że układ postaci (1) ma własność cienia o ile dla dowolnego $\epsilon \in \mathbb{R}_+$ istnieje $\delta \in \mathbb{R}_+$ taka, że dla dowolnej δ -pseudotrajektorii układu (1) istnieje punkt $p \in M_{ph}$ taki, że

$$d(u_k, f^k(p)) < \epsilon, k \in \mathbb{Z}$$

Istnieje wiele wariacji tej definicji. Na przykład lipszycowska własność cienia określa wspólne ograniczenie na odległości występujące w definicji szacując je stałą nazywaną stałą lipszyca. W związku z czym tą bazową własność określa się mianem standardowej - w dalszej części pracy mówiąc o własności cienia mamy na myśli jej standardową wersję. Przyjrzyjmy się chwilę co oznacza iż rozważany układ ma wspomnianą własność. W poprzednich podrozdziałach wielokrotnie wspomniane zostało, iż jedną z cech charakterystycznych układów chaotycznych jest ich wrażliwość na warunki początkowe. Można by przypuszczać, iż nawet niewielkie perturbacje warunków początkowych (rzędu błędów numerycznych) spowodują całkowitą rozbieżność wyników. Własność cienia daje podstawy, aby sądzić, iż nie stanie się tak koniecznie. Oczywiście sama definicja nie specyfikuje oczekiwanej wartości przy której taka rozbieżność nie zachodzi, jednak daje to pewną (lokalną) regularność zachowania się rozwiązań. Własność taka jest zatem niezwykle przydatna. Szerokłą klasę układów, których ona dotyczy specyfikuje tytułowy

Lemat Cienia 1. Niech dany będzie układ dynamiczny postaci (4) i niech będą spełnione postulowane dotychczas założenia. Niech Λ oznacza atraktor tego układu. Istnieje wówczas otoczenie $U \supset \Lambda$, który ma następującą własność: dla dowolnej liczby $\delta \in \mathbb{R}_+$ istnieje $\epsilon \in \mathbb{R}_+$ taki, że dla dowolnej ϵ -pseudoorbity $\{u_n\}_{n \in \mathbb{N}} \subset U$ istnieje orbita tegoż układu $\{u_n^r\}_{n \in \mathbb{N}}$, taka że

$$\forall n \in \mathbb{N} \quad u_n \in U_\delta(u_n^r). \quad (37)$$

Dysponując już odpowiednią nomenklaturą oraz szerokim arsenalem technik możemy przyjrzeć się metodom poświęconym analizie czułości oraz dotrzeć do sedna postawionego problemu.

3 Wybrane metody optymalizacji lokalnej stosowane w technice

Rezultatem analizy wrażliwości jest oszacowanie gradientu funkcji celu w wybranym punkcie przestrzeni parametrów. Zdecydowana większość metod optymalizacji lokalnej w istocie bazuje na tak wyznaczonym gradiencie, wobec czego dokładność jego szacowania jest krytycznym momentem procesu optymalizacyjnego. W dalszej części tego rozdziału zaprezentowane zostały trzy klasyczne metody, z których pierwsza, metoda różnic skończonych, jest dogodna do wyznaczania pochodnych względnie prostych funkcji oraz szacowania błędów dwóch następnych: metody Tangent oraz metody operatora sprzężonego, które znajdują liczne zastosowania w technice.

3.1 Metoda różnic skończonych

Metoda ta jest rozwinięciem intuicji stojących za wyznaczaniem pochodnej funkcji. Granicę ilorazu różnicowego przybliża się poprzez tzw. *schematy różnicowe*, które wymagają wyznaczenia wybranej funkcji w kilku punktach. Schematy te mają przykładowe postaci (do końca pracy pochodna poszukiwanej funkcji oznaczana będzie symbolem J , zaś jej parametr symbolem s):

$$\frac{dJ}{ds}(s_0) \approx \frac{J(s_0 + \Delta s) - J(s_0)}{\Delta s}, \quad (38)$$

$$\frac{dJ}{ds}(s_0) \approx \frac{J(s_0) - J(s_0 - \Delta s)}{\Delta s}, \quad (39)$$

$$\frac{dJ}{ds}(s_0) \approx \frac{J(s_0 + \Delta s) - J(s_0 - \Delta s)}{\Delta s}, \quad (40)$$

którym odpowiadają odpowiednio: schemat różnicowy przedni, wsteczny oraz centralny. Metoda ta jest bardzo prosta w implementacji i dość łatwo pozwala kontrolować dokładność wyznaczanej pochodnej poprzez dobór odpowiednio małego przyrostu Δs . Należy zauważyć, że dla odpowiednio małej wartości takiego przyrostu błąd arytmetyki zmienoprzecinkowej istotnie rzutuje na dokładności wyniku. Nietrywialnym rozszerzeniem tej metody jest zastosowanie przyrostu jako liczby czysto urojonej. Dalsze postępowanie jest analogiczne do wyżej wspomnianych, przy czym operacje wykonywane są zgodnie z arytmetyką liczb zespolonych. Ostateczny wzór przyjmuje postać:

$$\frac{dJ}{ds}(s_0) \approx \frac{\text{Im}[J(s_0 + i\epsilon)]}{\epsilon}, \quad (41)$$

gdzie Im oznacza część zespoloną, zaś ϵ jest liczbą rzeczywistą. Zabieg ten pozwala na separację badanego punktu przestrzeni parametrów oraz zaburzenia na dwa niezależne obiekty numeryczne, dzięki czemu problem błędów zaokrągleń zostaje wyeliminowany. Dokładność szacowanej pochodnej pozostaje ograniczona jedynie precyzją maszynową.

3.2 Metoda Tangent

Idea tej metody polega na śledzeniu wrażliwości rozwiązania układu (1) w zależności od wektora parametrów projektowych. Zakładając odpowiednią regularność rozwiązania i funkcji f można formalnie zróżniczkować to równanie względem wektora s . W wyniku otrzymujemy nowe równanie:

$$\frac{d\mathbf{v}}{dt} = \frac{\partial \mathbf{f}}{\partial \mathbf{u}} \mathbf{v} + \frac{\partial \mathbf{f}}{\partial \mathbf{s}}, \quad (42)$$

gdzie $v = \frac{\partial \mathbf{u}}{\partial s}$. Człon ten reprezentuje wrażliwość rozwiązania na parametr projektowy, a zatem równanie (42) to opis dynamiki tej wielkości. Innymi słowy jest to propagacja początkowego zaburzenia w czasie, która dyktowana jest wrażliwością dynamiki układu na parametr projektowy. Referując teraz do chwilowej wartości funkcji postaci $J(s) = J(\mathbf{u}(t, \mathbf{s}), s)$ i różniczkując względem $s \in \mathbb{R}$ dostajemy:

$$\frac{dJ}{ds} = \frac{\partial J}{\partial \mathbf{u}} \mathbf{v} + \frac{\partial J}{\partial s} \quad (43)$$

Ostatecznie:

$$\frac{d\bar{J}}{ds} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \left(\frac{\partial J}{\partial \mathbf{u}} v + \frac{\partial J}{\partial s} \right) dt, \quad (44)$$

przy czym milcząco założono, iż prawdziwa jest równość

$$\frac{d\bar{J}}{ds} = \frac{dJ}{ds}. \quad (45)$$

W następnym Rozdziale okaże się, iż stosowalność takiego założenia jest mocno ograniczona. Jest to główna przyczyna załamania się klasycznych metod optymalizacji lokalnej dla układów o nieprzewidywalnym zachowaniu rozwiązań.

Postać wzoru (44) sugeruje potencjalne źródło problemów - rozwiązanie równania (42) jest numerycznie czasochłonne. Człon $v = \frac{\partial \mathbf{u}}{\partial s}$ to człon, którego wyznaczenie w przypadku wielowymiarowego wektora $s \in \mathbb{R}^n$ jest niezwykle czasochłonne (por. Metoda różnic skończonych). Wobec tego metoda ta może się sprawdzić w praktyce, gdy wektor parametrów projektowych ma niski wymiar.

3.3 Metoda Operatora Sprzężonego

Naszym celem jest znalezienie gradientu funkcji postaci (6). Zauważamy, że funkcja podcałkowa zależy od rozwiązania układu (1), które z kolei zależy od wektora parametrów projektowych. W praktyce oprócz takiej zagnieżdżonej zależności funkcja celu zależy *explcite* od wektora zmiennych decyzyjnych

$$J = J(\mathbf{u}(t, \mathbf{s}), \mathbf{s}), \quad (46)$$

gdzie \mathbf{u} jest rozwiązaniem układu (1) z odpowiednim warunkiem początkowym.

Należy przy tym zwrócić uwagę, że funkcja, której ekstremów szukamy nie zależy od czasu. Jawna zależność funkcji celu od zmiennych decyzyjnych może referować do składnika odpowiadającego za stan końcowy, tj stan w chwili $t = T$, a za niejawną zależność odpowiada

składnik dla chwil $t < T$.

Poszukiwana pochodna wyraża się następującym wzorem

$$\frac{d\bar{J}}{ds} = \frac{1}{T} \int_0^T \left(\frac{\partial J}{\partial \mathbf{u}} v + \frac{\partial J}{\partial \mathbf{s}} \right) dt, \quad (47)$$

gdzie $v = \frac{\partial \mathbf{u}}{\partial s}$. Funkcje celu leżące w kręgu naszych zainteresowań charakteryzują się dużą dysproporcją liczby wyjść w stosunku do liczby wejść układu (zmiennych niezależnych). W praktyce wartościami takich funkcji są skalary lub wektory o stosunkowo niewielkiej liczbie współrzędnych (najczęściej stanowiące statystyki wielkości fizycznych), wejścia natomiast charakteryzują parametry modelu (na przykład współrzędne punktów rozpinających projektowaną geometrię) i ich liczba może być rzędu dziesiątek tysięcy. Wyznaczenie gradientu ze wzoru (47) jest kosztowne numerycznie, ponieważ każda perturbacja wektora \mathbf{s} wymaga wyznaczenia v na nowo, co narzuca konieczność ponownego rozwiązania układu (1). Dla tak dużej liczby zmiennych niezależnych powoduje to duży koszt czasowy procesu optymalizacji.

Wobec powyższych rozważań chcemy przekształcić to wyrażenie, tak aby móc wyznaczyć wrażliwość rozważanej funkcji względem całego wektora zmiennych projektowych symultanicznie. Aby usunąć kłopotliwe wyrazy posłużymy się techniką tzw. *mnożników Lagrange'a* $\mathbf{w}(t)$. Zauważmy, że równość (6) możemy zapisać w postaci

$$\langle \mathbf{J} \rangle = \frac{1}{T} \int_0^T \mathbf{J}(\mathbf{u}(t, \mathbf{s})) dt + \int_0^T (\mathbf{w}(t) \left(\frac{d\mathbf{u}}{dt} - \mathbf{f}(\mathbf{u}, \mathbf{s}) \right)) dt. \quad (48)$$

Ostatni składnik po prawej stronie ma zerową wartość na mocy (1). Własność ta jest prawdziwa dla dowolnego mnożnika Lagrange'a, w związku z czym otrzymujemy otrzymujemy. Niech zatem

$$-\frac{\partial \mathbf{w}}{\partial t} = \left(\frac{\partial \mathbf{f}}{\partial \mathbf{u}} \right)^T \mathbf{w} + \left(\frac{\partial J}{\partial \mathbf{u}} \right)^T \quad (49)$$

W wyniku różniczkowania wyrażenia (48) wraz z narzuceniem warunku (49) na funkcje $\mathbf{w}(t)$ otrzymujemy wyrażenie

$$\frac{d\bar{J}}{ds} = \frac{1}{T} \int_0^T \mathbf{w}^T \left(\frac{\partial \mathbf{f}}{\partial \mathbf{s}} + \frac{\partial J}{\partial \mathbf{s}} \right) dt \quad (50)$$

Należy w tym miejscu poczynić uwagę techniczną: iloczyn \mathbf{w} i \mathbf{f} pod całką w ogólności jest dobrze zdefiniowany w sensie iloczynu skalarnego nad ustaloną przestrzenią Hilberta. W celu zachowania przejrzystości nie odróżniamy tutaj tej operacji od iloczynu liczb.

W reprezentacji dyskretnej wyrażenia ma to postać:

$$\frac{d\bar{J}}{ds} = \sum_{i=0}^{N-1} \mathbf{w}_{i+1}^T \Delta t \frac{\partial \mathbf{f}}{\partial \mathbf{s}}, \quad (51)$$

gdzie N oznacza liczbę kroków dyskretyzacji czasowej, zaś Δt krok czasowy.

Oczywiście pochodna $\frac{\partial \mathbf{f}}{\partial \mathbf{s}}$ nie stanowi dużego problemu. Podkreślamy tutaj, że chwilowa wartość funkcji zadającej dynamikę ma postać $\mathbf{f} = \mathbf{f}(\mathbf{u}, \mathbf{s})$, w związku z czym pochodna cząstkowa wzięta jest względem drugiej zmiennej traktowanej jako niezależna. Mówiąc nieco kolokwialnie: pochodna ta 'nie widzi' zależności zagnieżdżonej $\mathbf{u} = \mathbf{u}(t, \mathbf{s})$.

4 Metoda Trajektorii Cienia

4.1 Wstęp

Z praktycznego punktu widzenia obiektem zainteresowania pozostają wartości średnie wielkości chwilowych, w których okres uśredniania jest stosunkowo długi, tj. funkcje postaci (6). W rozdziale tym będziemy poszukiwali czułości tak zdefiniowanych funkcji względem parametru $s \in \mathbb{R}$ (zakładamy przy tym, że parametr projektowy jest skalar, aby nie komplikować istoty problemu niepotrzebnymi rachunkami - rozszerzenie na przypadek wektorowy nie jest trudne). Konkretnie: poszukujemy wartości wyrażenia:

$$\frac{d\bar{J}}{ds} = \lim_{\delta s \rightarrow 0} \frac{1}{\delta s} \left(\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T J(\mathbf{u}(t, s + \delta s)) dt - \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T J(\mathbf{u}(t, s)) dt \right). \quad (52)$$

Fizyczne przykłady takich wielkości to między innymi:

- średni opór aerodynamiczny skrzydła samolotu w trakcie lotu,
- średni współczynnik przejmowania ciepła na powierzchni wymiennika,
- średnia globalna temperatura powietrza,
- średnia siła nośna działająca na łopatkę turbiny.

Istnienie takich długookresowych średnich całkowych związane jest z założeniem ergodyczności (podobnie jak ich niezależność od warunków początkowych), jednakże ich różniczkowalność nie jest oczywista. Głębokie rezultaty poczynione przez D. Ruelle'a przekonują ([31], [32]), że dla klasy układów zwanych 'quasi-hyperbolicznymi' takie wartości są nie tylko dobrze określone, ale również wyrażają się formułą:

Do systemów o takiej charakterystyce zalicza się z pewnością układ Lorenza. Istnieją ponadto przesłanki, że duża część istotnych z inżynierskiego punktu widzenia przypadków przepływów również wykazuje cechę 'quasi-hyperboliczności' [39] (również heurystyczne obserwacje bazujące na eksperymentach oraz symulacjach numerycznych pokazują, że takie wielkości statystyczne zmieniają się stosunkowo niewiele wraz z perturbacją dynamiki), wobec czego z teoretycznego punktu widzenia nie ma przeszkód w wyznaczaniu czułości tak zdefiniowanych funkcji. Jednakże zlinearyzowane równania używane do tego celu (niezależnie od wybranej metody, których przegląd zawiera Rozdział) skutkują w wynikach o bardzo niskiej jakości. Przyczyną załamania się klasycznych metod analizy wrażliwości w przypadku układów chaotycznych jest ich duża wrażliwość na niewielkie zaburzenia dynamiki układu, która przejawia się w rozbieżności rozwiązań wraz z upływem czasu (tempo tej rozbieżności determinowane jest przez wykładniki Lyapunowa, zaś jej kierunek przez wektory kowariantne Lyapunowa - patrz Rozdział). Innymi słowy:

$$\frac{d\bar{J}}{ds} \neq \lim_{T \rightarrow \infty} \frac{d}{ds} \frac{1}{T} \int_0^T J(s) dt. \quad (53)$$

Główną przeszkodę stanowi fakt, iż czas numerycznej symulacji jest zawsze skończony. Zmienność wartości oczekiwanej aproksymowanej średnią na skończonej trajektorii pozostaje nieograniczona (w normie) i rozważane przejścia graniczne nie komutują, tj.:

$$\lim_{s' \rightarrow s} \lim_{T \rightarrow \infty} \left(\frac{1}{T} \int_0^T \frac{J(\mathbf{u}(t, s')) - J(\mathbf{u}(t, s))}{s' - s} dt \right) \neq \lim_{T \rightarrow \infty} \lim_{s' \rightarrow s} \left(\frac{1}{T} \int_0^T \frac{J(\mathbf{u}(t, s')) - J(\mathbf{u}(t, s))}{s' - s} dt \right). \quad (54)$$

Najważniejszą ideą tytułowej Metody Trajektorii Cienia jest 'wymuszenie' jednostajnej ograniczoności odległości między trajektoriami w przestrzeni fazowej w dowolnej chwili. Innymi słowy, chcemy aby :

$$d(\mathbf{u}(t, s'), \mathbf{u}(t, s)) < M \quad \forall t \in (0, \infty), \quad (55)$$

dla pewnej stałej $M \in \mathbb{R}_+$. Wówczas rozważane przejście graniczne (54) byłoby w istocie prawdziwe. W konsekwencji prawa strona nie 'wybuchalaby' wraz z rozwojem układu, jak to ma zazwyczaj miejsce w przypadku dowolnego wyboru trajektorii zaburzonej. Jak już zostało wspomniane, jako pierwsi zjawisko to zauważyli autorzy w pracy [22] przy badaniu układu Lorenza (1963) i zaproponowali metodę regularyzacji bazującą na uśrednianiu dostatecznej liczby trajektorii (jednocześnie stanowi ona pierwszy przykład z rodziny metod wykorzystujących zespół trajektorii) [23]. Formalne uzasadnienie poprawności tego podejścia stanowią prace [31]. Metody takie są jednak numerycznie kosztowne, co pokazują prace [9]. Należy przy tym wspomnieć, że wymiar zagadnienia może być zredukowany koncentrując się jedynie na niestabilnych wektorach kowariantnych Lyapunowa (wkład do średniej pochodzący od modów stabilnych jest wyznaczany poprzez klasyczne metody analizy wrażliwości). Jest to tak zwany algorytm S3 (ang. Space Split Sensitivity Analysis algorithm) i jest opisany wraz z analizą kosztowności tego podejścia w stosunkowo nowych pracach [10],[11]. Niniejsza praca koncentruje się na metodach, które wykorzystują jedną (generyczną) trajektorię, co od razu wyznacza kres górny wymiaru zagadnienia.

Podstawową ideą Metody Trajektorii Cienia jest regularyzacja poprzez odpowiedni wybór trajektorii zaburzonej przy innym warunku początkowym, tak aby zachodziło (55). Postępowanie takie jest usprawiedliwione na mocy założenia ergodyczności (Rozdział). Głównym narzędziem jest przy tym operator cienia, którego własności zostaną przedstawione w następnych sekcjach. Skutkiem jego działania będzie uzyskanie trajektorii bliskiej wyjściowej dla dowolnego czasu, co pozwoli uzyskać równość w (53).

Do końca pracy przyjmujemy założenia z Rozdziału. W szczególności te dotyczące ergodyczności oraz hiperboliczności układu (1), jak i regularności funkcji $\mathbf{f}: M_{ph} \rightarrow M_{ph}$.

4.2 Metoda Operatora Cienia

Przypuśćmy, że dana jest trajektoria pewnego układu dynamicznego postaci (1), którą zapiszemy w postaci

$$\mathbf{u}(t) = \mathbf{u}_t, \quad (56)$$

gdzie $\forall t \in \mathbb{R}_+ \quad \mathbf{u}_t \in M_{ph}$. Wprowadźmy zaburzenie w postaci rodziny funkcji gładkich i wspólnie ograniczonych $\{\delta \mathbf{u}_t \mid t \in \mathbb{R}_+\}$, takiej że $\forall t \in \mathbb{R}_+ \quad (\mathbf{u}_t + \epsilon \delta \mathbf{u}_t) \in M_{ph}$, gdzie $\epsilon \in \mathbb{R}_+$

jest pewną stałą. Mamy wówczas

$$\mathbf{u}'(t) = \mathbf{u}'_t = \mathbf{u}_t + \epsilon \delta \mathbf{u}_t. \quad (57)$$

Zauważmy, że z przyjętych na początku tego Rozdziału założeń wynika, że tak zdefiniowana trajektoria odpowiada układowi postaci

$$\frac{d}{dt} \mathbf{u}'(t) = \mathbf{f}(u) + \epsilon \frac{\partial \delta \mathbf{u}}{\partial \mathbf{u}} \mathbf{f}(u) = \mathbf{f}(u') - \epsilon \frac{\partial \mathbf{u}}{\partial \mathbf{u}} \delta \mathbf{u}(\mathbf{u}') + \epsilon \frac{\partial \delta \mathbf{u}}{\partial \mathbf{u}} \mathbf{f}(u') + O(\epsilon^2), \quad (58)$$

gdzie ostatnia równość wynika z przyjętych założeń na temat zaburzenia $\delta u(u)$ [39]. Wobec tego można to równanie zapisać w postaci

$$\frac{d}{dt} \mathbf{u}'(t) = \mathbf{f}(\mathbf{u}') + \epsilon \delta \mathbf{f}(\mathbf{u}') + O(\epsilon^2), \quad (59)$$

gdzie perturbacja dyfeomorfizmu $\delta \mathbf{f}: M \rightarrow M$ wyraża się formułą

$$\delta \mathbf{f}(\mathbf{u}) = -\frac{\partial \mathbf{f}}{\partial \mathbf{u}} \delta \mathbf{u}(\mathbf{u}) + \frac{d \delta \mathbf{u}}{dt} = S_f \delta \mathbf{u}(\mathbf{u}). \quad (60)$$

Wprowadzony powyżej liniowy operator $S_f: \delta \mathbf{u} \mapsto \delta \mathbf{f}$ nazywamy *operatorem cienia*. Tak zdefiniowany operator jest dobrze określony ([39]), tj. w sposób jednoznaczny definiuje gładkie zaburzenie $\delta \mathbf{f}: M \rightarrow M$ odpowiadające gładkiej perturbacji wyjściowej trajektorii (56). Innymi słowy: zaburzona trajektoria (57) spełnia jednoznacznie zaburzony układ (59). Pewną komplikację stanowi fakt, że z punktu widzenia wyznaczania wrażliwości (52) istotna jest operacja odwrotna. Perturbacja parametru projektowego s powoduje zmianę dynamiki $\delta \mathbf{f}(s)$, dla której poszukiwana jest perturbacja wyjściowej trajektorii $\delta \mathbf{u}(\delta \mathbf{f}(s)) = S^{-1}(\delta \mathbf{f}(s))$ (pamiętając przy tym, że musi być ona 'dostatecznie mała'). Mając opracowaną procedurę wyznaczania odwrotności tego operatora mamy:

$$\lim_{T \rightarrow \infty} \lim_{s' \rightarrow s} \left(\frac{1}{T} \int_0^T \frac{J(\mathbf{u}(t, s')) - J(\mathbf{u}(t, s))}{s' - s} dt \right) = \lim_{T \rightarrow \infty} \int_0^T \frac{\partial J}{\partial \mathbf{u}} \delta \mathbf{u}(\delta \mathbf{f}(s)) dt. \quad (61)$$

Przypomnijmy, że wyrażenie po lewej stronie (61) wyznacza poszukiwaną wrażliwość tylko w przypadku, gdy zaburzona trajektoria pozostaje blisko w sensie metryki przestrzeni fazowej trajektorii generycznej.

Problem inwersji operatora cienia (przy założeniu, że takie odwzorowanie odwrotne istnieje) można rozwiązać numerycznie wykorzystując spektrum Lyapunowa. Konkretnie założmy, że dla każdej chwili $t \in (0, T]$ dany jest układ wektorów Lyapunowa:

$$\phi_1(\mathbf{u}(t, s)), \phi_2(\mathbf{u}(t, s)), \dots, \phi_n(\mathbf{u}(t, s)). \quad (62)$$

Założmy ponadto, że rozpinają one całą przestrzeń styczną w punkcie $\mathbf{u}(t, s) \in M_{ph}$, tj.

$$\text{span}(\phi_1(\mathbf{u}(t, s)), \phi_2(\mathbf{u}(t, s)), \dots, \phi_n(\mathbf{u}(t, s))) = T_u M_{ph}. \quad (63)$$

Pamiętając, że $\delta \mathbf{f}, \delta \mathbf{u} \in T_u M_{ph}$ otrzymujemy

$$\begin{aligned} \delta \mathbf{f}(\mathbf{u}) &= \sum_{i=1}^n a_i^f \phi_i(\mathbf{u}) \\ \delta \mathbf{u}(\mathbf{u}) &= \sum_{i=1}^n a_i^u \phi_i(\mathbf{u}) \end{aligned}, \quad (64)$$

gdzie przez a_i^f, a_i^u $i = 1, \dots, n$ oznaczono pola skalarne (jednoznacznie wyznaczone) odpowiadające δf oraz δu , odpowiednio. Następnie działamy operatorem cienia na jeden ze składników sumy w powyższych wzorach i korzystając z (59) dostajemy:

$$S_f(a_i^u \phi_i(\mathbf{u})) = -a_i^u \frac{\partial f}{\partial \mathbf{u}} \phi_i(\mathbf{u}) + \frac{da_i^u}{dt} \phi_i(\mathbf{u}) + a_i^u \frac{d\phi_i(\mathbf{u})}{dt}. \quad (65)$$

Zauważmy ponadto, że na mocy (60) możemy zapisać dynamikę rozwoju wektorów Lyapunova jako:

$$\frac{d\phi_i(\mathbf{u})}{dt} = \frac{\partial f}{\partial \mathbf{u}} \phi_i(\mathbf{u}) - \lambda_i \phi_i(\mathbf{u}). \quad (66)$$

Aplikując (66) do (65) otrzymujemy:

$$S_f(a_i^u \phi_i(\mathbf{u})) = \left(\frac{da_i^u}{dt} - \lambda_i a_i^u \right) \phi_i(\mathbf{u}). \quad (67)$$

Korzystając z liniowości S_f ostatecznie dostajemy:

$$\delta f = \sum_{i=1}^n S_f(a_i^u \phi_i(\mathbf{u})) = \sum_{i=1}^n a_i^f \phi_i(\mathbf{u}), \quad (68)$$

gdzie

$$a_i^f = \frac{da_i^u}{dt} - \lambda_i a_i^u \quad (69)$$

Powyższe formuły sugerują, że konsekwencją działania operatora cienia jest transformacja skalarne pola $a_i^u \rightarrow a_i^f$, co z liniowości tego operatora daje pełną informację o przekształceniu $\delta u \rightarrow \delta f$. Zauważmy również, iż dysponując a_i^f jesteśmy w stanie wyznaczyć a_i^u poprzez równanie (69), a zatem i odwrócić działanie operatora S_f , pod warunkiem, że jest on odwracalny. Może się jednak okazać, że w dla pewnych δf operator ten jest nieodwracalny i uzyskane wyniki nie mają sensu. Sytuacja taka nie jest abstrakcyjna, bowiem każdy ciągły układ dynamiczny o nietrywialnym atraktorze posiada zerowy wykładnik Lyapunova ([39]). W sytuacji takiej (69) redukuje się do równości

$$a_{n_0}^f = \frac{da_{n_0}^u}{dt}, \quad (70)$$

gdzie przez $n_0 \in \mathbb{N}$ oznaczono indeks zerowego wykładnika Lyapunova. Biorąc średnią całkową i przechodząc do granicy w równości (70) dostajemy

$$\lim_{T \rightarrow \infty} \int_0^T a_{n_0}^f(u) dt = \lim_{T \rightarrow \infty} \frac{a_{n_0}^u(u(T)) - a_{n_0}^u(u(0))}{T} = 0. \quad (71)$$

Widzimy zatem, że przy nieodpowiednim zaburzeniu dynamiki (na przykład $\delta f = f$) operator cienia jest nieodwracalny, tj. $S_f^{-1} \delta f = \emptyset$.

Aby pokonać tę trudność raz jeszcze możemy wykorzystać założenie ergodyczności. Pozwala ono bowiem wprowadzić parametr, którego wartość możemy modyfikować nie zmieniając przy tym interesujących nas średnich. Zauważmy, że dla $\epsilon \in \mathbb{R}_+$ średnia zdefiniowana jak w (6) dla zmodyfikowanego równania postaci

$$\frac{d\mathbf{u}}{dt} = (1 + \epsilon \eta) f(\mathbf{u}) \quad (72)$$

ma tę samą wartość na mocy wspomnianego założenia. Oczywiście w ogólności może zachodzić $\eta = \eta(t)$, a więc dysponujemy wektorem zmiennych, których wartości możemy dobrać dowolnie, a więc tak aby dla dowolnego zaburzenia δf (pamiętając, że musi być ono gładkie) istniał przeciwobraz S_f^{-1} , tj. operator cienia był odwracalny. Istotnie bowiem równanie charakteryzujące zerowy wykładnik Lyapunova (71) zamienia się w równość

$$\lim_{T \rightarrow \infty} \int_0^T a_{n_0}^f(u) dt + \eta(t) = 0, \quad (73)$$

co daje związek na $\eta(t)$. Wówczas możemy wyznaczyć

$$\delta u = S_f^{-1}(\delta f + \eta f). \quad (74)$$

Pozostaje się zastanowić czy dla dowolnej trajektorii rozważanego układu istnieje gładkie zaburzenie o którym mowa w początkowym założeniu tego rozdziału (tzn. takie, które pozostaje jednostajnie blisko (55)). Twierdzącą odpowiedź uzyskujemy powołując się na Lemat Cienia, co w konsekwencji daje deterministyczny algorytm obliczania wrażliwości funkcji postaci (6).

4.3 Metoda Trajektorii Cienia

Sposób wyznaczania gradientu funkcji w układach dynamicznych oparty o odwracanie operatora cienia jest numerycznie kosztowny ze względu na konieczność obliczenia spektrum Lyapunova.

W sytuacjach podobnych do zaistniałej dobrym pomysłem jest próba przeformułowania problemu. Inspirację mogą stanowić techniki rozwiązywania równań różniczkowych częstkowych. Wykorzystuję się w nich podejście wariacyjne, zarówno w ujęciu numerycznym (na tym oparta jest Metoda Elementów Skończonych) jak i analitycznym (dowodzi się istnienia słabych rozwiązań, które to spełniają pewne wariacyjne równości całkowe, a następnie podnosi się ich regularność wykorzystując topologię przestrzeni w jakiej są one zdefiniowane). Wspomniane już zostało, że aby w (54) zachodziła równość musimy zapewnić spełnienie warunku (55). Chcemy zatem znaleźć trajektorię pozostającą blisko trajektorii wyjściowej w dowolnej chwili. Formułując to spostrzeżenie w języku optymalizacji - chcemy zminimalizować normę odległości pomiędzy dwiema trajektoriami (genericzną oraz tą, której poszukujemy) w przestrzeni fazowej. Poszukujemy zatem funkcji $\mathbf{u}_s : [0, T] \rightarrow M_{ph}$ takiej, że

$$L'(\mathbf{u}_{shadow}) = \inf_{\mathbf{u}} \frac{1}{T} \int_0^T d^2(\mathbf{u}(t), \mathbf{u}_r(t)) dt, \quad (75)$$

gdzie przez $d(\cdot, \cdot)$ oznaczono metrykę w przestrzeni fazowej M_{ph} , zaś \mathbf{u}_r jest rozwiązaniem równania (1), czyli stanowi trajektorię referencyjną. Winni jesteśmy ponadto narzucić warunki na (75) w celu dookreślenia postawionego problemu. Wiemy, iż poszukiwana funkcja $\mathbf{u}_s : [0, T] \rightarrow M_{ph}$ powinna spełniać równanie (1) dla odpowiednio sperturbowanego parametru $s \in \mathbb{R}^n$. Raz jeszcze zaznaczamy, że nie ma konieczności, aby warunek początkowy był taki sam jak dla trajektorii referencyjnej.

Problem w takim sformułowaniu jest oczywiście dobrze postawiony (powołując się na ponownie na Lemat Cienia wnioskujemy, że zbiór trajektorii, dla których wartość funkcjonału (75) jest skończona jest niepusty). Pozostaje zauważyć, że rozumowanie dotyczące dodatkowej funkcji spełniającej rolę swobodnego parametru $\eta : [0, T] \rightarrow \mathbb{R}$ pozostaje słuszne dokładnie z tych samych powodów co poprzednio. Aby możliwe było jej zaimplementowanie do wariacyjnego sformułowania konieczne jest zrozumienie wpływu tej funkcji na dynamikę układu. W poprzednim rozdziale została ona wprowadzona jako techniczny parametr odpowiednio rozszerzający przeciwobraz operatora cienia tak aby dla dowolnego gładkiego zaburzenia dynamiki możliwe było znalezienie zaburzenia trajektorii. Samej dynamiki oczywiście nie możemy zmieniać, ale jej tempo już tak. Pomysł polega na tym, aby manipulować chwilami występowania punktów w przestrzeni fazowej zachowując ich kolejność. W konsekwencji obraz trajektorii nie zmienia się, jednak funkcja będąca rozwiązaniem będzie już inna. Można to zrobić poprzez transformację dziedziny. W tym celu wprowadźmy funkcję $\tau \in C^1([0, T])$, tj. funkcję różniczkowalną w sposób ciągły transformującą przedział $[0, T]$ w siebie w sposób wzajemnie jednoznaczny. Otrzymane złożenie $\mathbf{u}(\tau(t))$ spełnia:

$$\frac{d\mathbf{u}(\tau(t))}{dt} = \mathbf{f}(\mathbf{u}(\tau(t)), s). \quad (76)$$

Taka konstrukcja pozwala na kontrolę tempa występowania konkretnych punktów w przestrzeni fazowej co w konsekwencji pozwala zachować odpowiednio małą odległość od trajektorii referencyjnej w odpowiadających sobie chwilach.

Funkcja taka odpowiedni 'przyspiesza' lub 'spowalnia' rozwiązanie w wybranych chwilach wzdłuż trajektorii. Oczywiście w początkowej chwili musi zachodzić $\tau(t=0) = 0$. Interesująca jest zatem jej pochodna. W szczególnym przypadku braku jej wpływu, tj. gdy $\tau(t) = id$ mamy $\frac{d\tau(t)}{dt} = 1$ dla każdego $t \in [0, T]$. Wprowadźmy oznaczenie $\eta(t) = \frac{\tau(t)}{dt} - 1$. Poprzednie spostrzeżenia wskazują, że 'przyspieszanie' rozwiązania następuje przy $\eta(t) > 0$, zaś 'hamowanie' dla $\eta(t) < 0$, w związku z czym funkcję $\eta(t)$ nazywamy *dylatacją czasu*. Wyjątkowo prosta jej interpretacja zachodzi dla układów z czasem dyskretnym, a więc i dla realizacji numerycznych równań zwyczajnych. Sprowadza się ona bowiem do powszechnie stosowanej techniki adaptacyjnego kroku czasowego. Wprowadzając taką funkcję modyfikujemy postać funkcjonału (75). Otrzymujemy

$$L(\mathbf{u}_{shadow}) = \inf_{\mathbf{u}, \tau} \frac{1}{T} \int_0^T (d^2(\mathbf{u}(\tau(t)), \mathbf{u}_r(t)) + \alpha^2 \eta^2(t)) dt, \quad (77)$$

gdzie poszukiwana funkcja $\mathbf{u}(\tau(t))$ spełnia (76). Minimalizacja tej formy całkowej stanowi istotę Metody Trajektorii Cienia.

Wartość funkcji celu można wyznaczyć ze wzoru przybliżonego:

$$\bar{J}^\infty(s) \approx \bar{J}_{shadow}^T(s) = \frac{1}{\tau(T) - \tau(0)} \int_{\tau(0)}^{\tau(T)} J(\mathbf{u}_{shadow}^T(t, s), s) dt. \quad (78)$$

Oczywiście skończony okres uśredniania T powinien być odpowiednio długi, aby oddać asymptotyczne zachowanie rozważanego układu.

Dysponując taką trajektorią oraz funkcją celu dla niej określoną jesteśmy w stanie wyznaczyć poszukiwaną wrażliwość. Można to zrobić za pomocą dowolnej z metod opisanych

w Rozdziale. Jest to możliwe, gdyż znaleziona w opisany sposób trajektoria pozostaje w odpowiednio bliskim sąsiedztwie trajektorii referencyjnej przez cały okres uśredniania i w konsekwencji otrzymujemy równość w (54).

Zdyskretyzowane równania Metody Trajektorii Cienia Za pomocą równanie (42) otrzymujemy formułę wyrażającą poszukiwaną wrażliwość:

$$\frac{d\langle \mathbf{J} \rangle}{ds} \approx \frac{\int_0^T (\frac{\partial \mathbf{J}}{\partial u} v + \frac{\partial \mathbf{J}}{\partial s} + \eta(\mathbf{J} - \bar{\mathbf{J}})) dt}{T}, \quad (79)$$

gdzie

$$\bar{\mathbf{J}} = \frac{1}{T} \int_0^T (\mathbf{J}) dt. \quad (80)$$

Dyskretna postać zagadnienia (78) wyraża się wzorem:

$$\min_{v_i, \eta_i} \sum_{i=0}^{m-1} \frac{\|v_{i+\frac{1}{2}}\|}{2} + \alpha^2 \sum_{i=1}^{m-1} \frac{\eta_i^2}{2}, \quad (81)$$

przy czym zachodzi:

$$E_i v_{i-\frac{1}{2}} + f_i \eta_i + G_i v_{i+\frac{1}{2}} = b_i \quad 1 \leq i < m, \quad (82)$$

gdzie

$$\begin{aligned} E_i &= -\frac{1}{\Delta t} I - \frac{\partial f}{\partial u}(u_{i-\frac{1}{2}}, s), \\ f_i &= \frac{u_{i+\frac{1}{2}} - u_{i-\frac{1}{2}}}{\Delta t}, \\ G_i &= \frac{1}{\Delta t} I - \frac{\partial f}{\partial u}(u_{i+\frac{1}{2}}, s), \\ b_i &= \frac{1}{2} \left(\frac{\partial f(u_{i-\frac{1}{2}}, s)}{\partial s} + \frac{\partial f(u_{i+\frac{1}{2}}, s)}{\partial s} \right) \end{aligned} \quad (83)$$

Ostatecznie dochodzimy do układu równań liniowych (układu Karusha-Kuhna-Tuckera), który w przypadku metody Tangent ma postać:

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{V} \\ \mathbf{W} \end{bmatrix} = \begin{bmatrix} \mathbf{F} \\ \mathbf{H} \end{bmatrix}, \quad (84)$$

gdzie

$$\mathbf{A} = \begin{bmatrix} I & & & & & \\ & \alpha^2 & & & & \\ & & I & & & \\ & & & \alpha^2 & & \\ & & & & I & \\ & & & & & \ddots \\ & & & & & & \alpha^2 \\ & & & & & & & I \end{bmatrix}, \mathbf{B} = \begin{bmatrix} E_1^T & & & & & \\ f_1^T & & & & & \\ G_1^T & E_2^T & & & & \\ & f_2^T & & & & \\ & G_2^T & & & & \\ & & \ddots & & & \\ & & & E_m^T & & \\ & & & f_m^T & & \\ & & & G_m^T & & \end{bmatrix}, \quad (85)$$

oraz

$$\mathbf{C} = \begin{bmatrix} E_1 & f_1 & G_1 & & & \\ & E_2 & f_2 & G_2 & & \\ & & & & \ddots & \\ & & & & & \ddots \\ & & & & & E_m & f_m & G_m \end{bmatrix}, \mathbf{V} = \begin{bmatrix} v_{\frac{1}{2}} \\ \eta_1 \\ v_{1+\frac{1}{2}} \\ \eta_2 \\ v_{3+\frac{1}{2}} \\ \vdots \\ v_{m-\frac{1}{2}} \\ \eta_{m-1} \end{bmatrix}, \mathbf{H} = \begin{bmatrix} -b_1 \\ -b_2 \\ \vdots \\ -b_3 \end{bmatrix}, \quad (86)$$

zaś \mathbf{D} oraz \mathbf{F} są macierzami zerowymi.

Układ taki można uprościć stosując klasyczną metodę eliminacji Gaussa uzyskując postać nazywaną w literaturze dopełnieniem Shura [40]:

$$\mathbf{B}\mathbf{B}^T \mathbf{w} = \mathbf{b}, \quad (87)$$

gdzie

$$\mathbf{B} = \begin{bmatrix} E_1 & \frac{f_1}{\alpha} & G_1 & & & \\ & E_2 & \frac{f_2}{\alpha} & G_2 & & \\ & & & & \ddots & \\ & & & & & \ddots \\ & & & & & E_m & \frac{f_m}{\alpha} & G_m \end{bmatrix}, \mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_3 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_3 \end{bmatrix}. \quad (88)$$

Zauważmy, że macierz główna układu $\mathbf{B}\mathbf{B}^T$ jest symetryczna oraz dodatnio określona. Jest ona także blokowo-trójdzielna. Ma ona następującą strukturę:

$$\mathbf{B}\mathbf{B}^T(i, i) = E_i E_i^T + f_i f_i^T + G_i G_i^T \quad (89)$$

oraz

$$\mathbf{B}\mathbf{B}^T(i, i+1) = \mathbf{B}\mathbf{B}^T(i+1, i) = G_i E_{i+1}^T, \quad (90)$$

gdzie para indeksów oznacza macierz kwadratową, której wymiar jest równy wymiarowi punktu przestrzeni fazowej (w kontekście numerycznej mechaniki płynów jest on proporcjonalny do wymiarów siatki obliczeniowej). Pozostałe bloki to macierze zerowe. Wspomnijmy, iż używając bezpośrednich metod rozwiązywania takiego układu, koszt numeryczny jest rzędu $O(mn^3)$ (przypominamy, iż przez m oznaczono liczbę kroków czasowych, zaś n to wymiar punktu przestrzeni fazowej). W związku z tym, dla odpowiednio gęstej dyskretyzacji obszaru obliczeniowego lepszym wyborem są metody iteracyjne. Powróćmy do tematu wydajności różnych metod numerycznych w rozważanej klasie przypadków w Rozdziale.

Następnie z (83) dostajemy

$$\begin{aligned} v_{i+\frac{1}{2}} &= -G_i^T w_i - E_{i+1}^T w_{i+1} \quad 1 \leq i < m; \\ \eta_i &= -\frac{f_i^T w_i}{\alpha^2} \quad 0 < i < n \end{aligned} \quad (91)$$

Ostatecznie otrzymujemy (dla warunków $w_0 = w_{m+1} = 0$):

$$\frac{d\bar{J}}{ds} \approx \frac{1}{m} \sum_{i=0}^{m-1} \left(\frac{\partial J(u_{i+\frac{1}{2}}, s)}{\partial u} v_{i+\frac{1}{2}} + \frac{\partial J(u_{i+\frac{1}{2}}, s)}{\partial s} \right) + \frac{1}{m-1} \sum_{i=1}^{m-1} \eta_i \tilde{J}_i, \quad (92)$$

gdzie

$$\tilde{J}_i = \frac{J(u_{i-\frac{1}{2}}, s) + J(u_{i+\frac{1}{2}}, s)}{2} - \frac{1}{m} \sum_{i=0}^{m-1} J(u_{i+\frac{1}{2}}, s) \quad 1 \leq i < m. \quad (93)$$

Z powyższych wzorów wnioskujemy, iż postać równania Tangent w Metodzie Trajektorii Cienia zawiera dodatkowy człon związany z dylatacją czasu w stosunku do (44). Jest to jednak jedynie techniczny parametr regularyzujący, a cała idea tkwi w odpowiednim rozwiązaniu $\{v_i\}_{i=1, \dots, m}$, który nazywamy *kierunkiem cienia*. W Rozdziale zostało wspomniane, iż w celu wyznaczenia gradientu funkcji celu posługujemy się techniką mnożników Lagrange'a. W przypadku minimalizacji funkcjonału (77) z ograniczeniem w postaci (76) odpowienia funkcja ma postać:

$$\Lambda = \int_0^T (v^T v + \alpha^2 \eta^2 + 2w^T + 2w^T (\frac{dv}{dt} - \frac{\partial f}{\partial u} v - \frac{\partial f}{\partial s} - \eta f)), \quad (94)$$

gdzie przez $w = w(t)$ oznaczono wspomniane mnożniki. Rozumowanie przedstawione w Rozdziale prowadzi do układu

$$\begin{cases} \frac{dv}{dt} - \frac{\partial f}{\partial u} v - \frac{\partial f}{\partial s} - \eta f = 0 \\ \frac{dw}{dt} + \frac{\partial f}{\partial u}^T w - v = 0 \\ w(0) = w(T) = 0 \\ \alpha^2 \eta - w^T f = 0 \end{cases}. \quad (95)$$

Dyskretna postać jest taka sama jak w przypadku metody Tangent (84), przy tych samych oznaczeniach (83), przy czym prawa strona tego układu jest inna: macierz \mathbf{H} składa się z samych zer zaś

$$\mathbf{F} = \begin{bmatrix} \frac{1}{m} \frac{\partial J(u_{\frac{1}{2}}, s)}{\partial u} \\ \frac{1}{m-1} \bar{J}_1 \\ \frac{\partial J(u_{1+\frac{1}{2}}, s)}{\partial u} \\ \frac{1}{m} \frac{\partial J(u_{2+\frac{1}{2}}, s)}{\partial u} \\ \frac{1}{m-1} \bar{J}_2 \\ \vdots \\ \frac{1}{m-1} \bar{J}_{m-1} \\ \frac{\partial J(u_{m-\frac{1}{2}}, s)}{\partial u} \\ \frac{1}{m} \end{bmatrix}. \quad (96)$$

Wobec powyższego komplementacja Shura ma postać

$$\mathbf{B}\mathbf{B}^T \mathbf{w} = \mathbf{B}\mathbf{g}, \quad (97)$$

gdzie obie strony wyrażają się wzorami (88), (89) oraz (90).

Zauważmy na koniec, iż obie metody charakteryzuje ta sama macierz główna układu.

5 Implementacja metody LSS

W niniejszym rozdziale przedstawione zostały wyniki symulacji numerycznej dla dwóch scenariuszy:

- Scenariusz I - układ Lorentza (1963)
- Scenariusz II -

5.1 Scenariusz I - konceptualne preliminarium

5.1.1 Model matematyczny

Zastosujemy teraz opisaną metodę trajektorii cienia w analizie wrażliwości za pomocą zagadnienia sprzężonego dla układu Lorentza, tj. układu chaotycznego opisywanego autonomicznym układem nieliniowych równań różniczkowych zwyczajnych postaci:

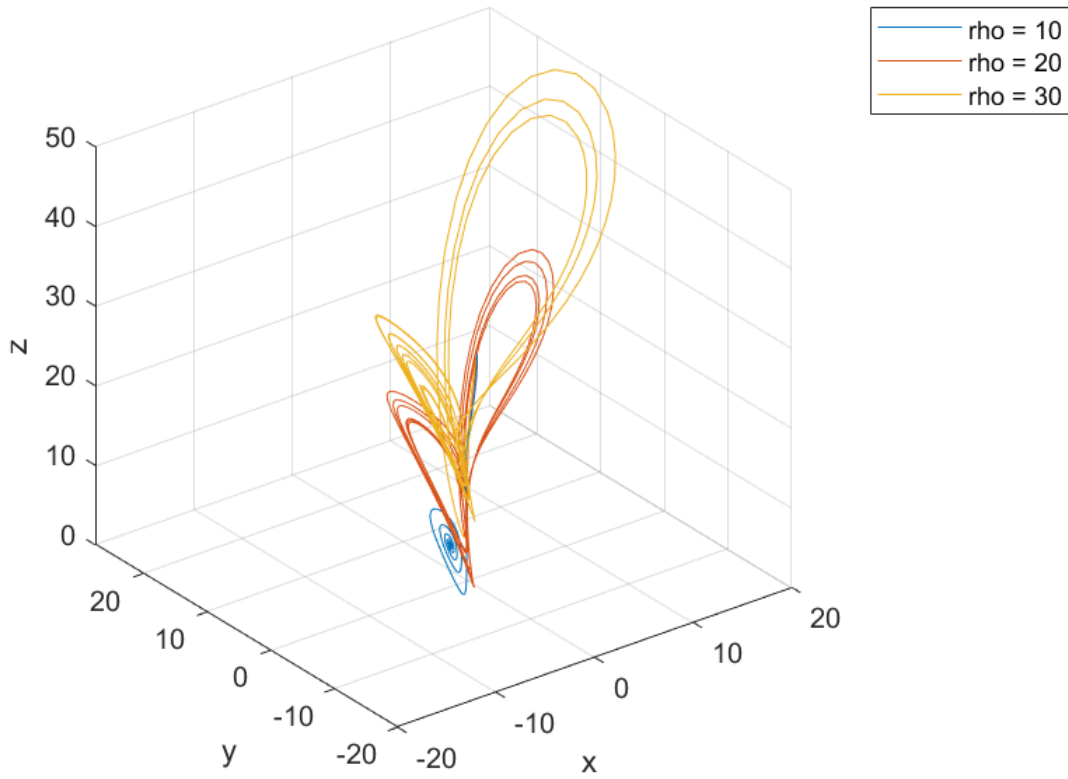
$$\begin{cases} \frac{dx}{dt} = \sigma(y - z) \\ \frac{dy}{dt} = x(\rho - z) - y \\ \frac{dz}{dt} = xy - \beta z \end{cases}, \quad (98)$$

gdzie σ, ρ, β są pewnymi stałymi. Nie zostały określone warunki początkowe, zakładając jedynie, że nie przyjmują one poniższych wartości (punkty krytyczne układu Lorentza):

- $x = y = z = 0$,
- $x = y = \sqrt{\beta(\rho - 1)}, z = \rho - 1$,
- $x = y = -\sqrt{\beta(\rho - 1)}, z = \rho - 1$,

Układ ten jest chaotyczny w ścisłym sensie i jest bardzo prostym modelem ruchu konwekcyjnego powietrza atmosferycznego (oryginalny układ Lorentza składał się z równań, rozważany jest skrajnie uproszczony - jest on minimalnym, w sensie ilości stopni swobody, układem, który może przejawiać chaotyczne zachowania), przy czym stała σ charakteryzuje lepkość ośrodka, ρ współczynnik przewodzenia ciepła, zaś β wymiar charakterystyczny.

Trajektorie dla trzech wybranych parametrów ρ w przypadku $x(0) = 1, y(0) = 0, z(0) = 28$ przedstawia Rysunek (15).



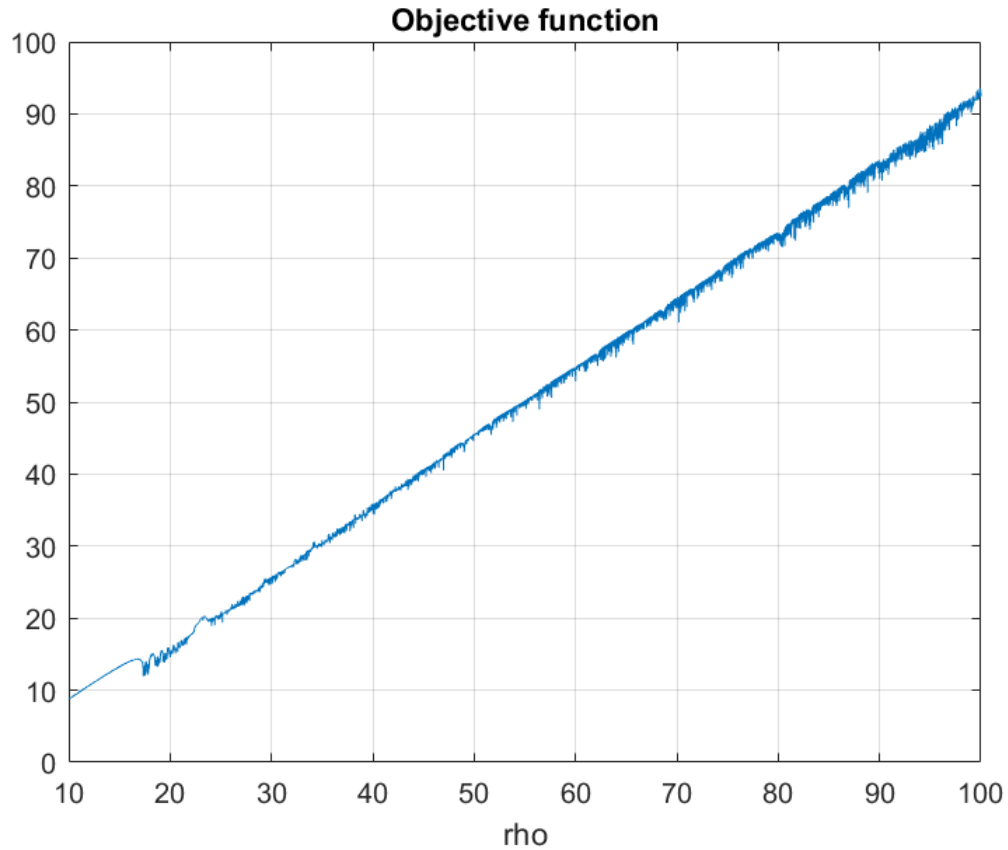
Rysunek 15: Trajektorie układu Lorentza.

5.1.2 Wyniki symulacji numerycznej

Funkcja celu została zdefiniowana jako

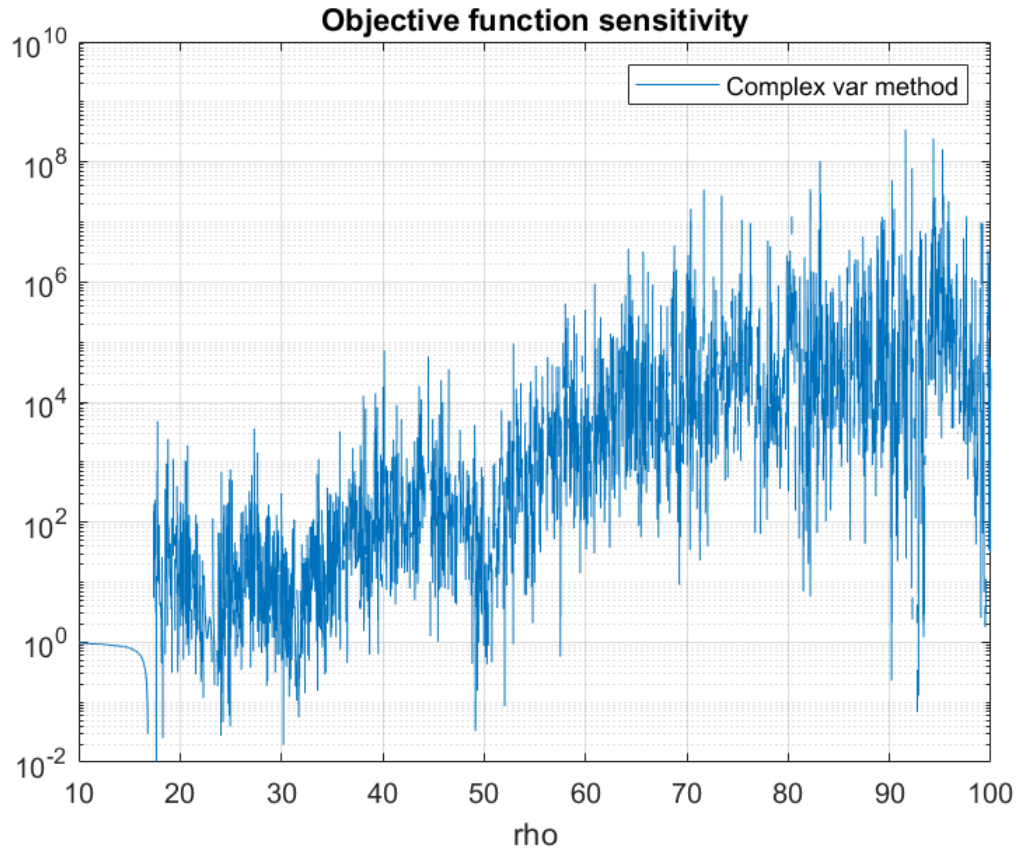
$$\langle \mathbf{J} \rangle = \lim_{T \rightarrow \infty} \int_0^T z dt \quad (99)$$

Założone zostały wartości parametrów: $\sigma = 10$, $\beta = \frac{8}{3}$, zaś ρ pozostawiono jako zmienną projektową (której zakres zmienności przebiega od 10 do 100). Rysunek (16) przedstawia zmienność funkcji celu w rozważanym zakresie zmian ρ , gdzie okres uśredniania $T = 10$.



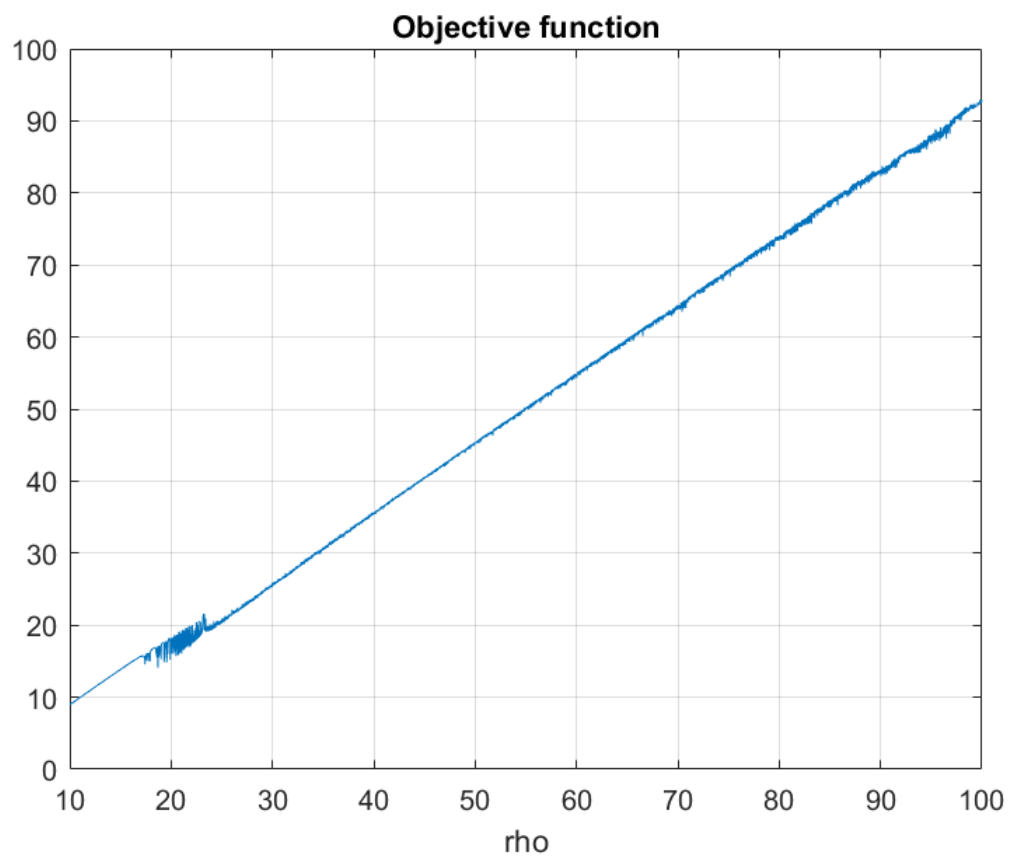
Rysunek 16: Zmienność funkcji celu w zależności od zmiennej decyzyjnej ($T = 10$)

Należy zauważyć, że chociaż ogólny trend krzywej sugeruje stałe jej nachylenie (wartość tangensa kąta tego nachylenia jest bliska jedności), to występują lokalne fluktuacje, które są przyczyną problemów w wyznaczenia wrażliwości tej funkcji na zmianę parametrów projektowych. Wyznaczona została dla tego przypadku pochodna względem parametru ρ , której wykres ilustruje Rysunek (17) (oś rzędnych jest przedstawiona w skali logarytmicznej). Wykorzystano do tego celu metodę zaburzenia zaburzenia części urojonej zmiennej projektowej traktowanej jako liczbę zespoloną. Technika ta pozwala uzyskać pochodną z dokładnością maszynową (patrz Rozdział 3.1).

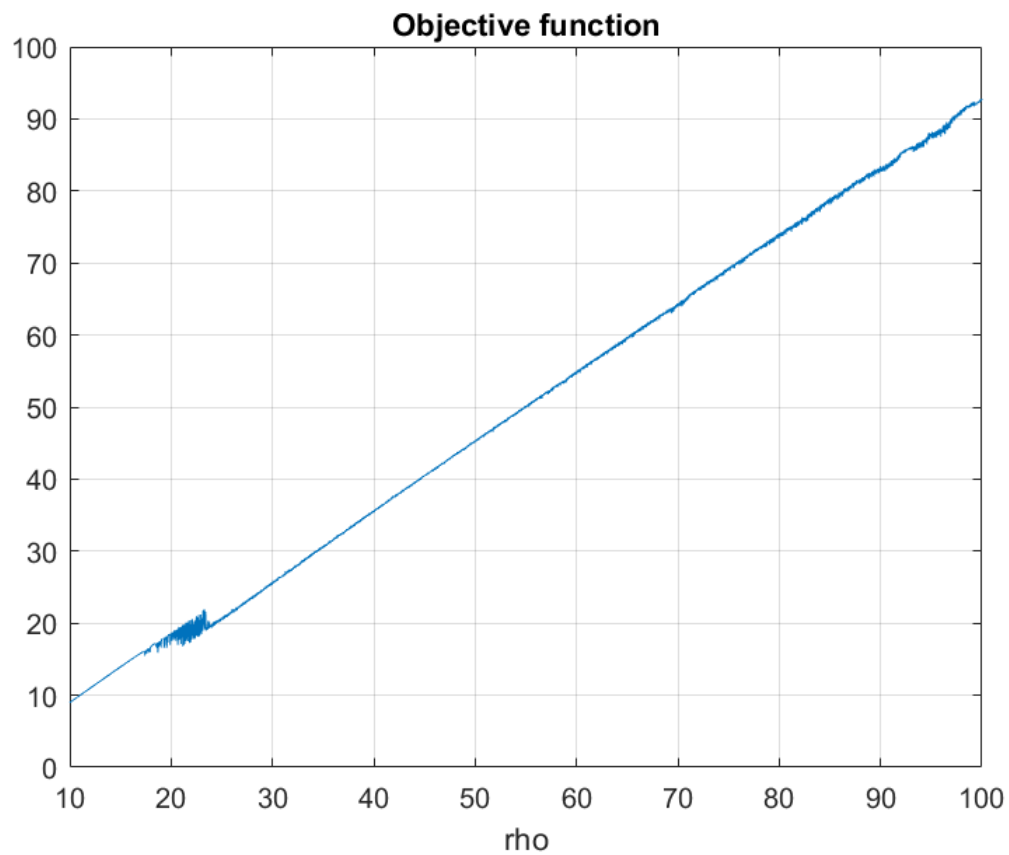


Rysunek 17: Wrażliwość funkcji celu wyznaczona metodą zaburzenia zespolonego

Tak wyznaczona czułość jest praktycznie nieprzydatna. Jednym ze sposobów 'wygładzenia' trendu zmian funkcji celu w zależności od zmiennej decyzyjnej jest zwiększenie okresu uśredniania (słuszność tego podejścia sugeruje założenie ergodyczności). Na Rysunkach (18) oraz (19) przedstawiono rozpatrywaną funkcję w przypadkach $T = 50$ oraz $T = 100$.

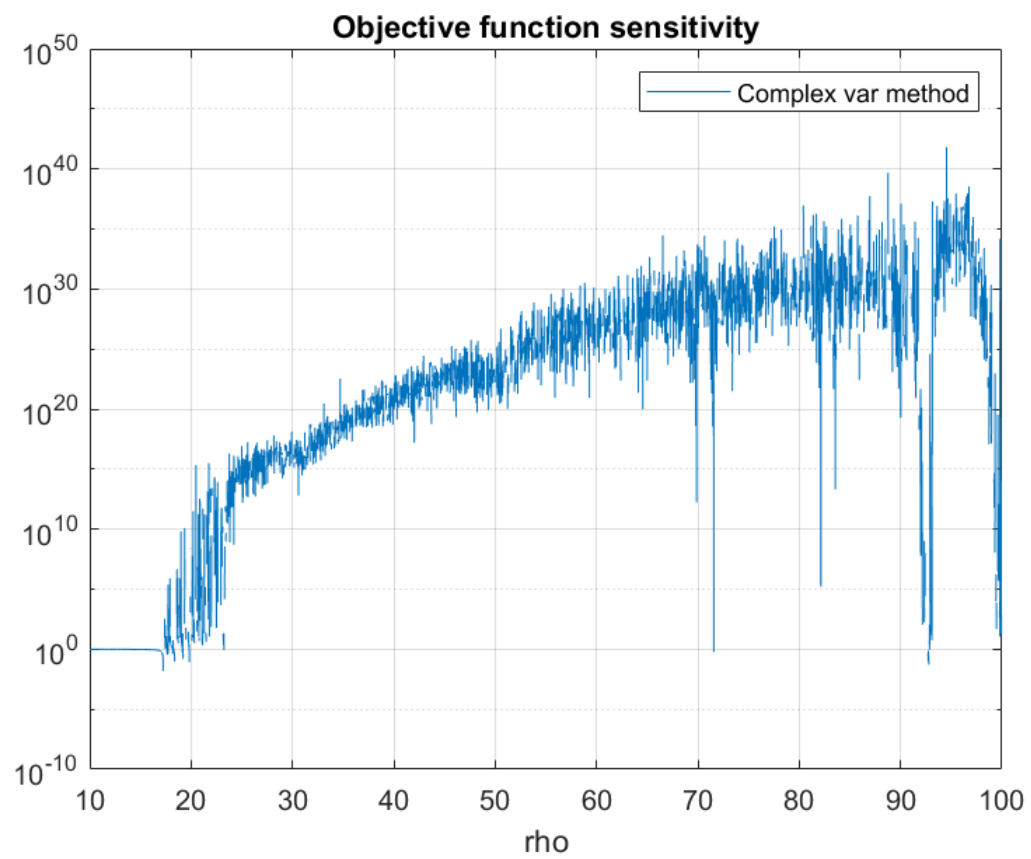


Rysunek 18: Zmienność funkcji celu w zależności od zmiennej decyzyjnej ($T = 50$)

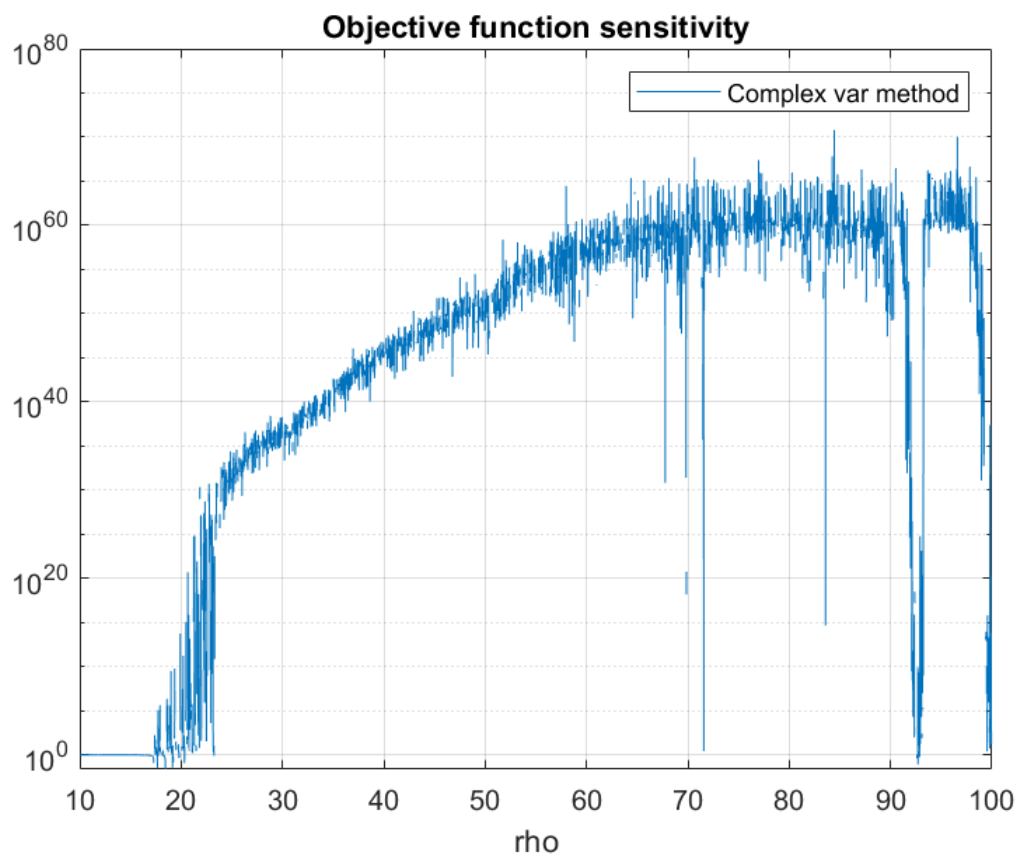


Rysunek 19: Zmienność funkcji celu w zależności od zmiennej decyzyjnej ($T = 100$)

Zgodnie z oczekiwaniami amplituda fluktuacji maleje, jednak ich obecność nie została wyeliminowana, co w istotny sposób rzutuje na wyznaczanie poszukiwanego gradientu (Rysunki 20 i 21).

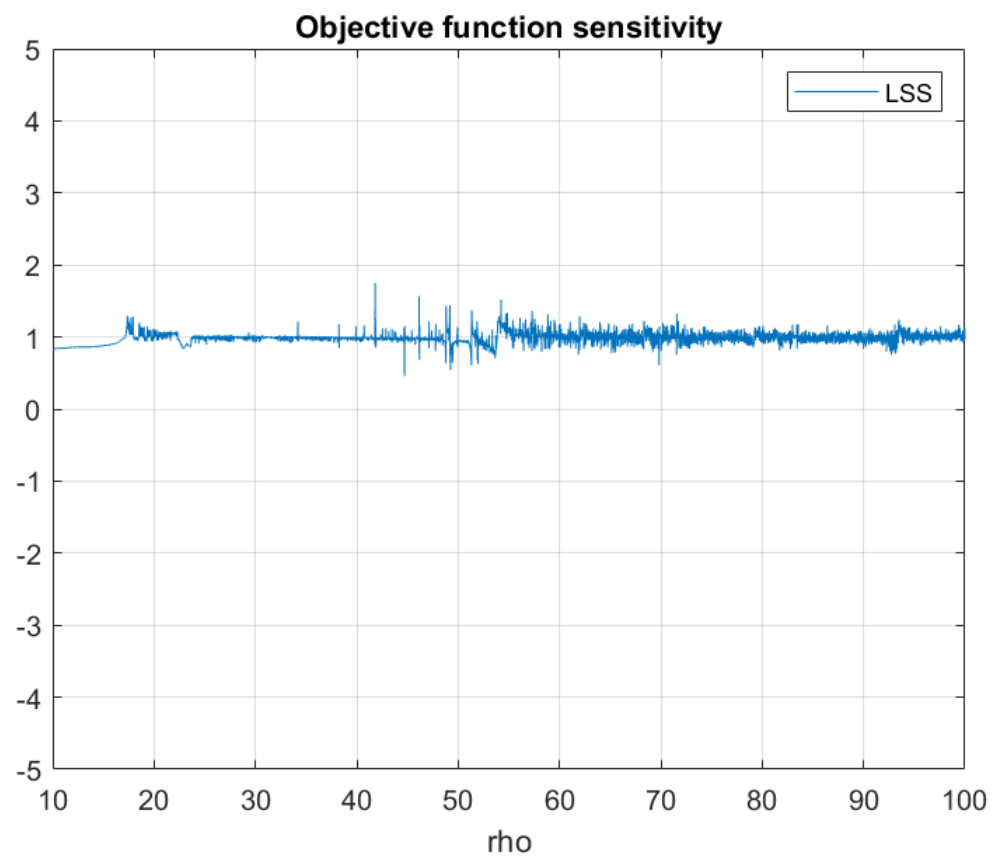


Rysunek 20: Zmienność funkcji celu w zależności od zmiennej decyzyjnej ($T = 50$)

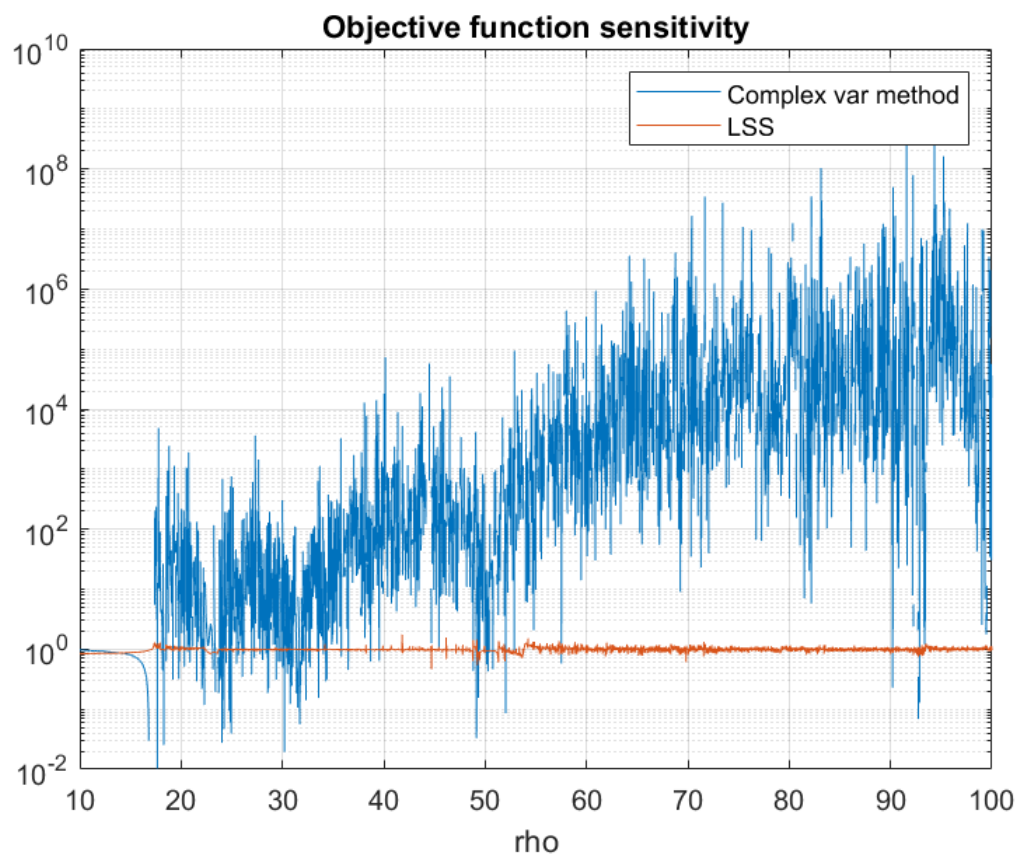


Rysunek 21: Zmienność funkcji celu w zależności od zmiennej decyzyjnej ($T = 100$)

Metoda trajektorii cienia została wykorzystana do wyznaczenia wrażliwości funkcji celu (przy użyciu operatora sprzężonego) dla pierwszego przypadku ($T = 10$). Wynik ilustruje Rysunek (22) , zaś Rysunek (23) pokazuje porównanie obu metod.

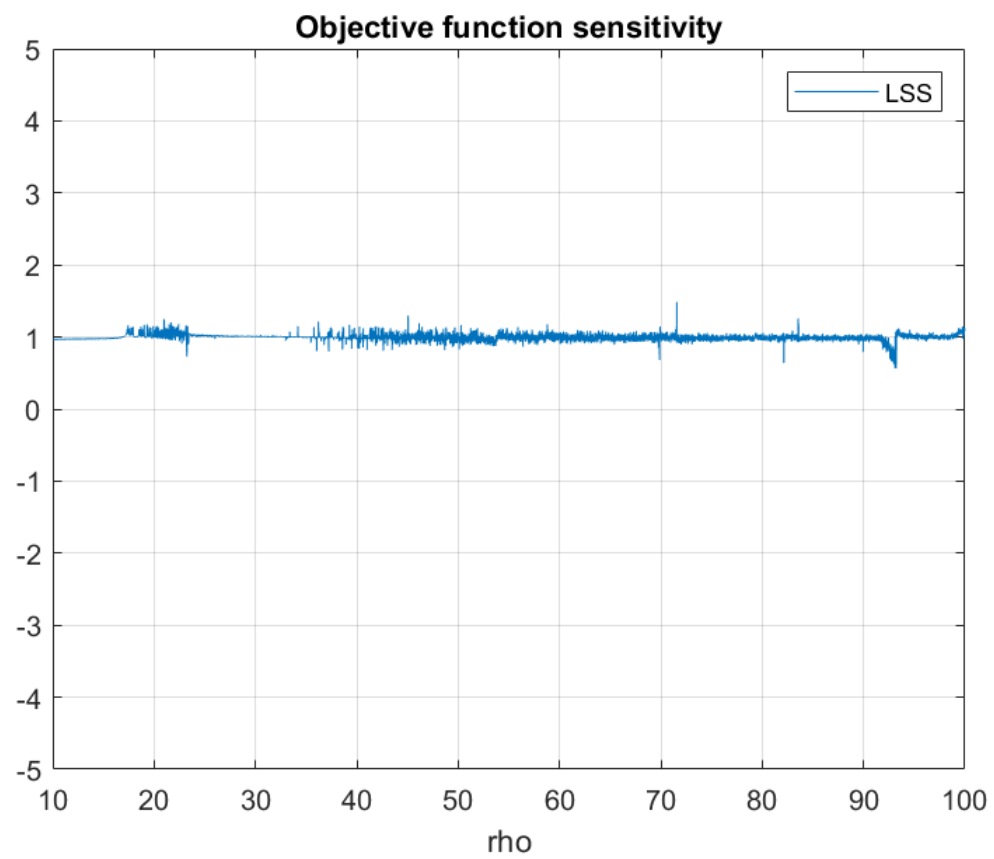


Rysunek 22: Czułość funkcji celu wyznaczona metodą trajektorii cienia.

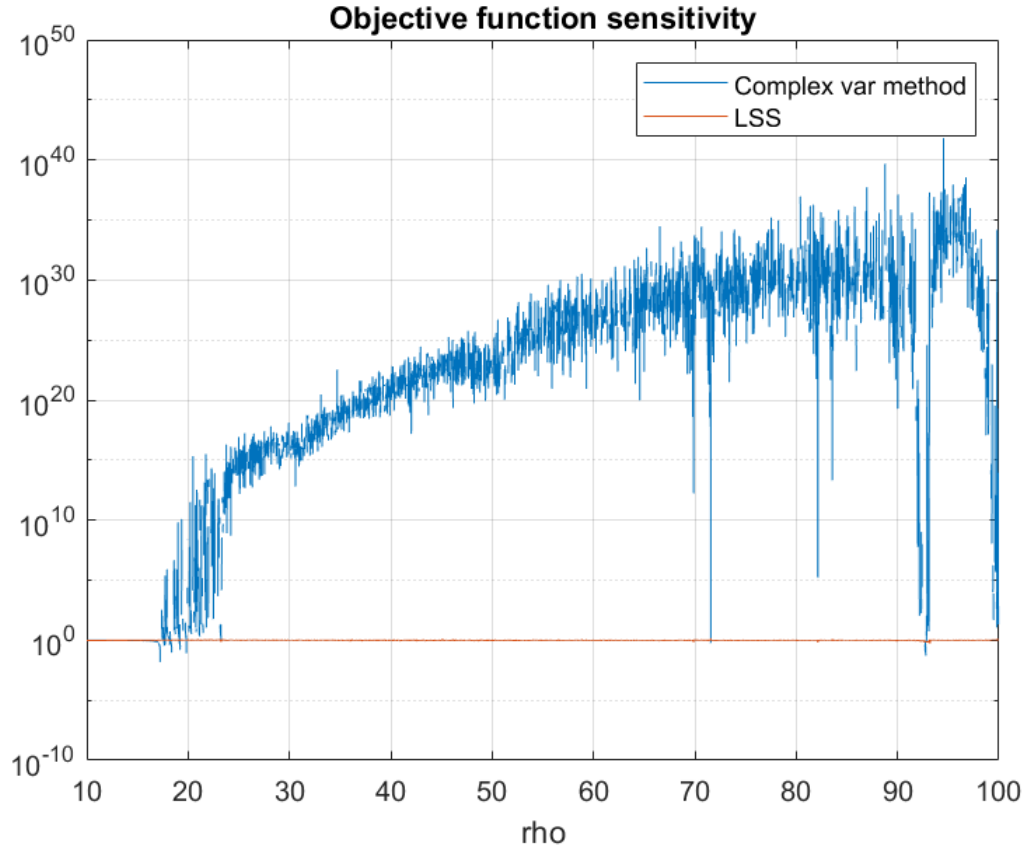


Rysunek 23: Porównanie obu metod

Podobna analiza została wykonana dla drugiego przypadku ($T = 50$), a jej wyniki przedstawiają Rysunki (24) oraz (25).



Rysunek 24: Czulość funkcji celu wyznaczona metodą trajektorii cienia.



Rysunek 25: Porównanie obu metod

Kod zaimplementowany w środowisku skryptowym MATLAB jest dostępny na publiczny repozytorium autora [25].

5.2 Scenariusz II - oblicze chaosu

5.2.1 Słowo wstępne

Jak zostało zaznaczone na wstępie, głównym celem niniejszej pracy jest analiza wrażliwości wybranej funkcji celu dla zagadnienia przepływowego modelowanego równaniem Boltzmanna. Konkretniej: chcemy zastosować Metodę Trajektorii Cienia do zagadnienia rozwiązywanego Metodą Siatek Boltzmanna (ang. *Lattice Boltzmann Method*). Istnieje cały szereg zalet nautry numerycznej, które pokrótce zostaną przedstawione w następnym podrozdziale. Jednak już w tym momencie należy wmienić obszary praktycznych zastosowań, które wykorzystują wspomnianą metodę. Są to między innymi ([13]) obliczenia związane z

- ośrodkami porowatymi (na przykład filtracja gleby)
- ogniwami paliwowymi

W poprzednim Rozdziale została przedstawiona metoda LSS w sposób generyczny. Obecnie należy dopasować jej schemat w celu regularyzacji równań opisujących dynamikę rozważanego przepływu. W momencie pisania tej pracy autor nie spotkał się z żadną istniejącą implementacją tej metody dla zdyskretyzowanych równań Boltzmanna - znane są jedynie aplikacje dla solverów bazujących na równaniach Naviera-Stokesa. Należy przy tym dodatkowo zaznaczyć, że wszystkie one oparte są o przechowywanie powstałej macierzy układu równań w pamięci. W dalszej części pracy ukazane zostaną wady takiego podejścia (w szczególności, zajętość pamięci operacyjnej w takich przypadkach całkowicie niweluje, zdaniem autora, jakiegokolwiek szanse na ich użycie w praktycznych zastosowaniach - nawet dla układów dwuwymiarowych korzystających z modeli turbulencji klasy RANS). Praca owa skupia się na podejściu, w którym macierz układu trajkuje się w sposób 'abstrakcyjny' - jako funktor, którego argumentami są wektory wejściowe, zaś wyjście stanowi iloczyn macierzy i wektorów wejściowych. Do opisu implementacji Metody Trajektorii Cienia konieczna jest znajomość równań wykorzystywanych w Metodzie Siatek Boltzmanna, dlatego też w dalszej części pracy skrótowo opisane zostały podstawowe jej założenia i płynące z nich wnioski w postaci gotowych formuł. Następnie zaprezentowane zostaną wyniki obliczeń dla przykładowego zagadnienia, w którym następuje oderwanie strugi na skutek skokowej zmiany geometrii. Zbadana zostanie przy tym czułość średniego spadku ciśnienia w długim horyzoncie czasowym w stosunku do modułu prędkości zadanej na wlocie do obszaru obliczeniowego.

5.2.2 Metoda Siatek Boltzmanna

Jak już zostało nieformalnie wspomniane (przy okazji omawiania pojęcia ergodyczności) w mechanice statystycznej definiuje się rozkład prawdopodobieństwa związany z nazwiskami Boltzmanna oraz Maxwella. Ma on następującą postać

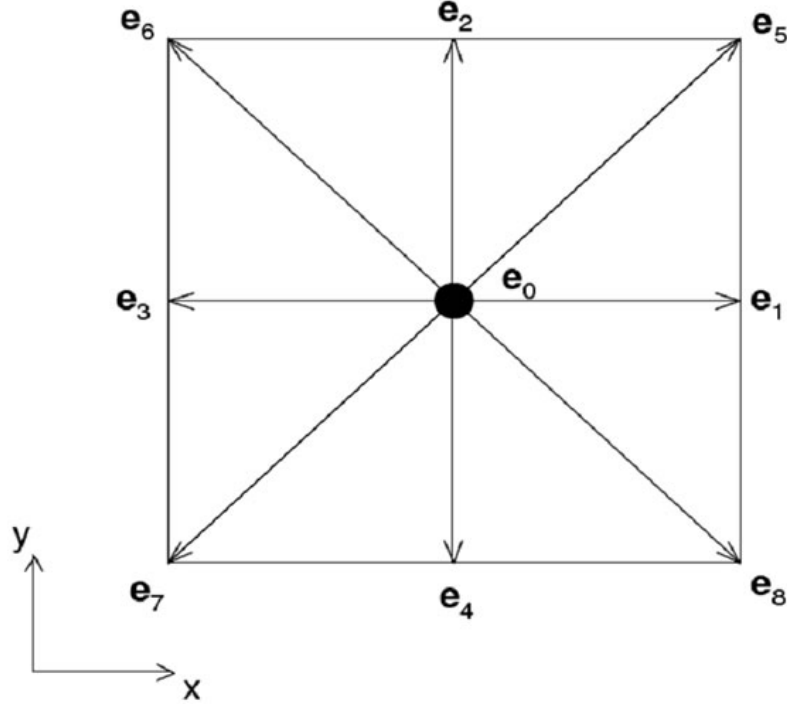
$$f(v) = 4\pi \left(\frac{m}{2\pi kT}\right)^{\frac{3}{2}} v^2 \exp\left(-\frac{\frac{1}{2}mv^2}{kT}\right), \quad (100)$$

gdzie przez v oznaczono prędkość cząstki, m jej masę, T temperaturę bezwzględną, oraz przez k stałą Boltzmana. Podstawowym założeniem prowadzącym do takiej dystrybucji prędkości wśród cząstek jest przyjęcie modelu gazu doskonałego oraz założenie równowagi termodynamicznej między układem jego otoczeniem. Należy zwrócić uwagę na znaczenie tego rozkładu - jest to gęstość prawdopodobieństwa stanowiąca miarę liczności cząstek w danym zakresie prędkości. Jak każda funkcja gęstości prawdopodobieństwa jest ona równa prawie wszędzie pochodnej dystrybuanty związanej z tym rozkładem. Wobec tego, na poziomie intuicji, można taką funkcję utożsamiać ze stosunkiem liczby cząstek mających ograniczony zakres przyjmowanych prędkości w stosunku do całkowitej ich liczby. Zauważmy, że podejście to nie opisuje każdej cząstki jako oddzielnego składnika złożonego systemu (podejście mikroskopowe). Jednakże jest ono również odmienne od dobrze znanego z obecnie najpopularniejszych metod modelowania komputerowego przepływów uśredniania w relatywnie niewielkiej (w sensie rozciągłości przestrzennej) objętości parametrów statystycznie istotnej liczby cząstek, z którego generowane są pola parametrów makroskopowych takich jak ciśnienie, temperatura czy prędkość (podejście to bazuje na

założeniu ciągłości badanego ośrodka i nazywane jest podejściem makroskopowym). Głównym zamysłem Metody Siatek Boltzmanna jest symulowanie ewolucji w czasie analogicznego rozkładu (niekoniecznie dla stanów równowagowych).

5.2.3 Model matematyczny

Jako schemat połączeń węzłów siatki przyjęty został typ d2Q9 (Rysunek (26))



Rysunek 26: Schemat połączeń węzłów siatki

W przyjętym modelu metody siatek Boltzmanna dyskretne równanie zadające dynamikę ma następującą postać:

$$f_i(x + \Delta x, t + \Delta t) = f_i(x, t) - \frac{1}{\tau_f} [f_i(x, t) - f_i^{eq}(x, t)]. \quad (101)$$

Przyjmujemy, że $f_i(x, t)$ oznacza funkcję rozkładu cząstek dla i -tej prędkości w punkcie x w chwili t . Parametry siatki oznaczamy przez Δx oraz Δt , zaś τ_f jest charakterystycznym czasem relaksacji. Jeśli pozostawimy układ bez wpływu czynników zewnętrznych, będzie on dążył do stanu równowagi, dla której równowagowa funkcja gęstości rozkładu f^{eq} , która ma postać:

$$f_i^{eq} = \rho w_i \left[1 + \frac{e_i \cdot u}{c_s^2} + \frac{(e_i \cdot u)^2 - (c_s |u|)^2}{2c_s^2} \right], \quad (102)$$

gdzie $w_0 = \frac{4}{9}$, $w_{1,2,3,4} = \frac{1}{9}$, $w_{5,6,7,8} = \frac{1}{36}$, $c_s = \sqrt{3}c$, gdzie $c = \frac{\Delta x}{\Delta t}$, przy czym c_s interpretujemy

jako prędkość dźwięku w płynie. Prędkość u wyraża się wzorem:

$$u = \sum_{i=0}^8 f_i e_i, \quad (103)$$

gdzie prędkości w kierunku i są postaci:

$$e_i = \begin{cases} (0, 0)c & \text{gdy } i = 0 \\ (\cos[(i-1)\frac{\pi}{2}], \sin[(i-1)\frac{\pi}{2}])c & \text{gdy } i = 1, 2, 3, 4 \\ (\sqrt{2}\cos[(2i-9)\frac{\pi}{4}], \sqrt{2}\sin[(2i-9)\frac{\pi}{4}])c & \text{gdy } i = 5, 6, 7, 8 \end{cases} \quad (104)$$

Gęstość w punkcie x w chwili t jest postaci:

$$\rho(x, t) = \sum_{i=0}^8 f_i(x, t). \quad (105)$$

Dodatkowo, gdy na cząsteczki działa zewnętrzna siła F , to makroskopowa prędkość płynu wyraża się następująco:

$$U(x, t) = u(x, t) + \frac{1}{2} \frac{F(x, t)}{\rho(x, t)}. \quad (106)$$

Dynamika w podejściu tej metody modelowana jest zatem dwustopniowo: pierwszy etap stanowi kolizja (reprezentowana operatorem kolizji), drugi zaś stanowi propagację funkcji rozkładu zgodnie z dyskretnymi kierunkami wybranych wektorów prędkości. W pracy operator kolizji jest aproksymowany dobrze rozpoznany modelem BGK. Należy zwrócić uwagę, iż rozkład równowagowy jest wielomianowym przybliżeniem rozkładu Maxwella-Boltzmann.

5.2.4 Aspekty numeryczne

Zanim przejdziemy do omówienia wyników należy chwilę poświęcić na aspekty implementacyjne, gdyż dla zagadnień wielowymiarowych odgrywają one kluczową rolę. W typowych przypadkach przepływów mogą nawet zdecydować o możliwości wykorzystania danej metody.

W Rozdziale 4.3.1 wspomniane zostało, iż zarówno Metoda Tangent jak i Metoda Operatora Sprzężonego charakteryzują się tą samą macierzą główną dla układu KKT. Macierz ta ma strukturę opisaną wzorami (89) oraz (90). Zauważmy, że macierze E_i, G_i są rzadkie (większość elementów ma zerową wartość), zaś f_i reprezentują wektory, które jednakże w ogólności nie są rzadkie. Ma to swoje poważne konsekwencje w odniesieniu do bloków diagonalny (bloki pod- i nad-diagonalne opisane równaniami (90) są macierzami rzadkimi jako iloczyn takich macierzy) - iloczyn $f_i f_i^T$ jest pełną macierzą kwadratową (przypominamy, że jej wymiar to wymiar zagadnienia m). Bloków takich jest tyle ile jest kroków czasowych symulacji n . W związku z tym, chcąc przechowywać macierz układu (w obydwu metodach) w pamięci, dostajemy dolne ograniczenie na ilość liczb zmiennopozycyjnych. Stosując notację *dużego* O otrzymujemy

$$N = O(m^2 n). \quad (107)$$

Zakładając, iż wymiar zagadnienia zależy potęgowo od gęstości dyskretyzacji obszaru (kwadratowo w przypadku dwuwymiarowym oraz sześciennie dla sytuacji trójwymiarowej), oznaczając przez k średnią ilość punktów dyskretyzacji wzdłuż jednego wymiaru geometrycznego oraz także biorąc $n = k$ mamy

$$N = O(k^{2d+1}), \quad (108)$$

gdzie $d = 2, 3$ oznacza ilość wymiarów geometrycznych. Dla uproszczenia przyjęliśmy, że każdy węzeł siatki obliczeniowej przechowuje jedną liczbę, co nie jest prawdą zarówno w przypadku zdyskretyzowanych równań Naviera-Stokesa jak i równań Boltzmanna.

Przyjmując dalej, iż liczba zmiennopozycyjna na typowej architekturze sprzętowej zajmuje 4 bajty oraz zakładając $k = 100$ (co w przypadku dwuwymiarowym implikuje jedyne 10000 węzłów) otrzymujemy szacunkową wartość $N \approx 4 \cdot 10^{10}$ bajtów, co oznacza, iż rozwiązując stosunkowo małe, dwuwymiarowe zagadnienie przepływowe (nieustalone) potrzebne jest ponad 36 GB pamięci. Dla symulacji trójwymiarowej sytuacja jeszcze bardziej się komplikuje (tzw. *klątwa wymiarowości*).

Powyższe szacunki pokazują, iż konieczna jest specjalna technika rozwiązywania układu, którego macierz opisywana jest równaniami (89) oraz (90). Wspomnieliśmy już, że jest to macierz symetryczna i dodatnio określona, więc jedną z pierwszych metod, która przychodzi na myśl jest Metoda Gradientów Sprzężonych (jako bardzo intuicyjna i stosunkowo prosta w implementacji). Wobec wspomnianej regularności macierzy jest ona w tym przypadku tożsama z popularną metodą *GMRES* (ang. *Generalized Minimal Residual*). Z założenia więc nie jest nam potrzebna macierz układu, a jedynie jej działanie na dowolny wektor zgodny z jej wymiarem. Potrzebujemy zatem funktora reprezentującego tę macierz (technika przynależąca do klasy metod tzw. *matrix-free*).

Podejście takie wymaga prostych reguł wyznaczania współczynników takiej macierzy. W tym momencie pojawia się główny problem: zgodnie z definicjami macierzy E_i, G_i , a także wektorów f_i zawierają one różnego rodzaju pochodne (patrz 84). Ze względu na wymiary zagadnienia, a także na dokładność wyznaczanie ich w sposób numeryczny jest niesatysfakcjonujące. Ze względu na iteracyjny charakter układów dynamicznych różniczkowanie symboliczne również nie stanowi rozwiązania.

Pomysł polega na zastosowaniu automatycznego różniczkowania, techniki stanowiącej niejako połączenie wspomnianych dwóch podejść. Bazuje ona na tzw. liczbach dualnych. Ich koncepcja, a także zastosowanie jest zbliżone do wyznaczania pochodnej za pomocą metod różnic skończonych w wariancie z zespolonym zaburzeniem (zob. Rozdział 3.1). Liczba dualna jest bowiem reprezentowana przez parę liczb rzeczywistych $(a, b) \in \mathbb{R}^2$, co możemy zapisać jako $a + b\epsilon$, gdzie odpowiednik jednostki urojonej, ϵ (tzw. nilpotent), utożsamiamy z zaburzeniem. Twór ten można sobie wyobrazić jako niezerowy obiekt, taki, że $\epsilon^2 = 0$ (a więc i $\epsilon^k = 0$ dla $k > 2$). Własność ta jest odwzorowaniem faktu, iż po podzieleniu przez nilpotent oraz przejściu granicznym $\epsilon \rightarrow 0^+$ otrzymujemy zerową wartość o ile ϵ był w potęgde większej od jedności. Widzimy tu analogię do wspomnianego wariantu metody różnic skończonych – speracja zaburzenia od punktu bieżącego na niezależne twory numeryczne, co w praktyce oznacza brak ograniczeń spowodowanych błędami zaokrągleń.

Mając tak zdefiniowane liczby dualne można bez trudu sprawdzić, iż dla wielomianu $P \in \mathbb{R}[x]$ zachodzi $P(a + b\epsilon) = P(a) + bP'(a)\epsilon$, a więc $P(a + \epsilon) = P(a) + P'(a)\epsilon$. W ogólności

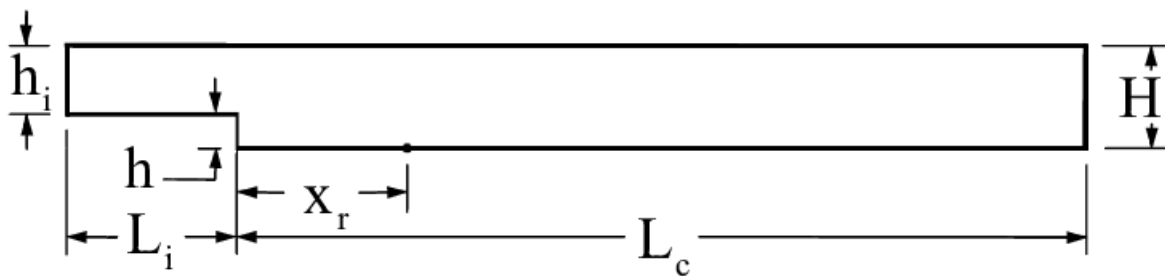
funkcje, które mają taką własność określane są mianem przestępnych. Aby więc policzyć pochodną funkcji przestępnej, wystarczy zbudować graf obliczeń (tak jak ma to miejsce w przypadku różniczkowania symbolicznego) z zaimplementowaną arytmetyką liczb dualnych, dzięki czemu ‘za jednym zamachem’ otrzymujemy wartość funkcji w rozważanym punkcie oraz jej pochodną tamże (arytmetyka symboliczna na grafie odpowiadającym funkcji jest dużo prostsza od generowania nowego grafu odpowiadającego jej pochodnej). Metoda ta posiada podobną przypadłość do Metody Tangent – jest wrażliwa na funkcje o wielowymiarowej dziedzinie. Dla każdej zmiennej niezależnej trzeba bowiem ‘przejsć’ po grafie aby wyznaczyć pochodną względem tej zmiennej (wartość samej funkcji jest redundantnie wyznaczana za każdym takim przejściem). Również tutaj istnieje podobne rozwiązanie – analog Metody Operatora Sprzężonego (w swojej istocie jest to dokładnie to samo, tj. zadziałanie operatorem na rozważane przekształcenie) i nazywa się odwrotnym różniczkowaniem automatycznym.

Konkluzje są tutaj takie same – aby wyznaczyć pochodną w danym punkcie należy wyznaczyć wartość funkcji w tym punkcie, tj. przejście po grafie obliczeniowy w jego naturalnym kierunku, a następnie przejście w drugą stronę. Interesujące jest, iż mechanizm taki jest bardzo popularny w technikach uczenia maszynowego. W szczególności głębokie sieci neuronowe bazują na optymalizacji lokalnej, w której gradient wyznaczany jest wspomnianą metodą – jest to popularny algorytm wstecznej propagacji błędów (ang. Backpropagation algorithm). Co więcej techniki automatycznego różniczkowania wstecznego są stosowane w popularnych bibliotekach związanych z wspomnianym tematem (jako dobry przykład wystarczy wspomnieć otwartą bibliotekę TensorFlow autorstwa Google Brain Team).

Wracając do postawionego problemu można uznać, iż dysponując odpowiednimi narzędziami możliwe jest zaimplementowanie macierzy w formie funkcji – dalsze kroki procesu optymalizacyjnego przebiegają standardowo, zgodnie z Metodą Gradientów Sprzężonych, która została wykorzystana w niniejszej pracy. Automatyczne różniczkowanie jest wykonywane za pomocą otwartego silnika TAPENADE stworzonego przez grupę INRIA Sophia-Antipolis (zob. Dokumentacja [18]). Możemy już teraz przejść do analizy wyników symulacji numerycznej przedstawionych w następnym podrozdziale.

5.2.5 Wyniki symulacji

Kształt obszaru obliczeniowego został zilustrowany na Rysunku (27)



Rysunek 27: Geometria obszaru obliczeniowego

W rozważanym przypadku $H = 2$, $h_i = h = L_i = 1$ oraz $L_c = 9$. Pomimo tak prostej geometrii w przepływie takim występuje szereg zjawisk o skomplikowanej naturze:

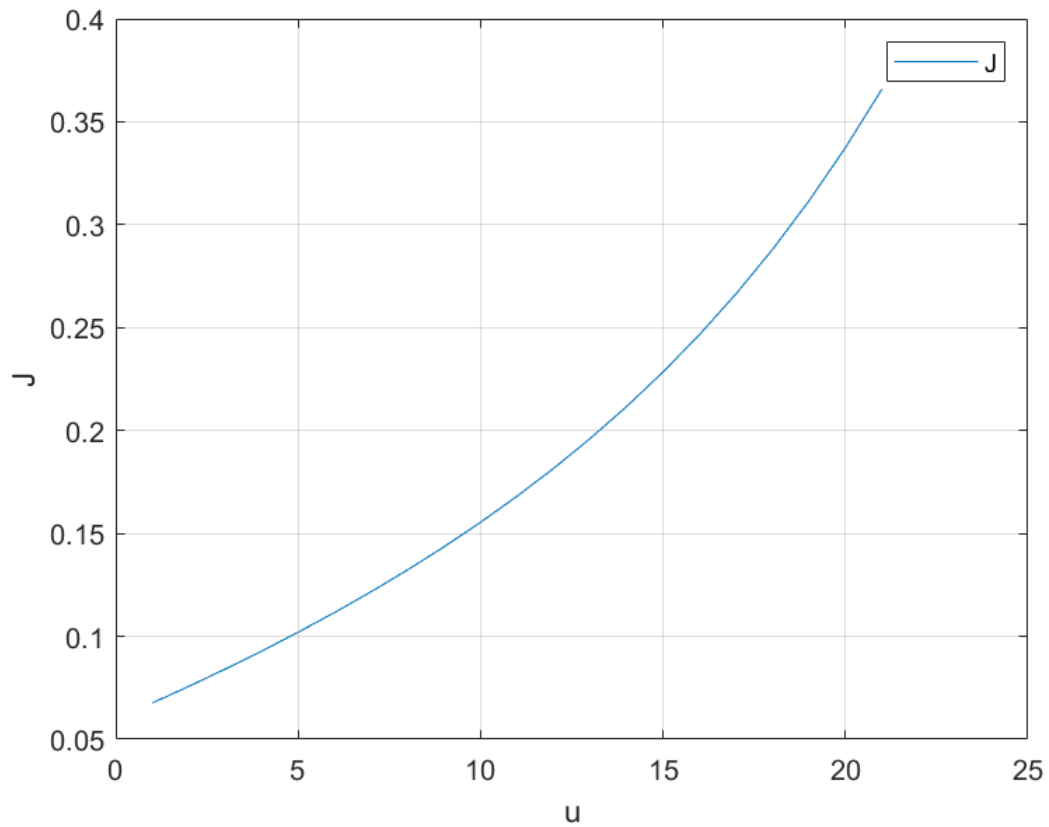
- niestabilność przepływu
- przejście laminarno-turbulentne
- separacja warstwy przyściennej
- oderwanie warstwy przyściennej
- recyrkulacja,

wobec czego przypadek ten często traktowany jest jako sprawdzian testowy dla różnych modeli turbulencji [33].

Jako funkcję celu przyjęty został średni spadek ciśnienia między wlotem a wylotem, zaś jako parametr kontrolny przyjęty został moduł prędkości wlotowej (przyjęty został tło-kowy rozkład prędkości na wlocie, zaś liczba kroków czasowych $T = 4000$):

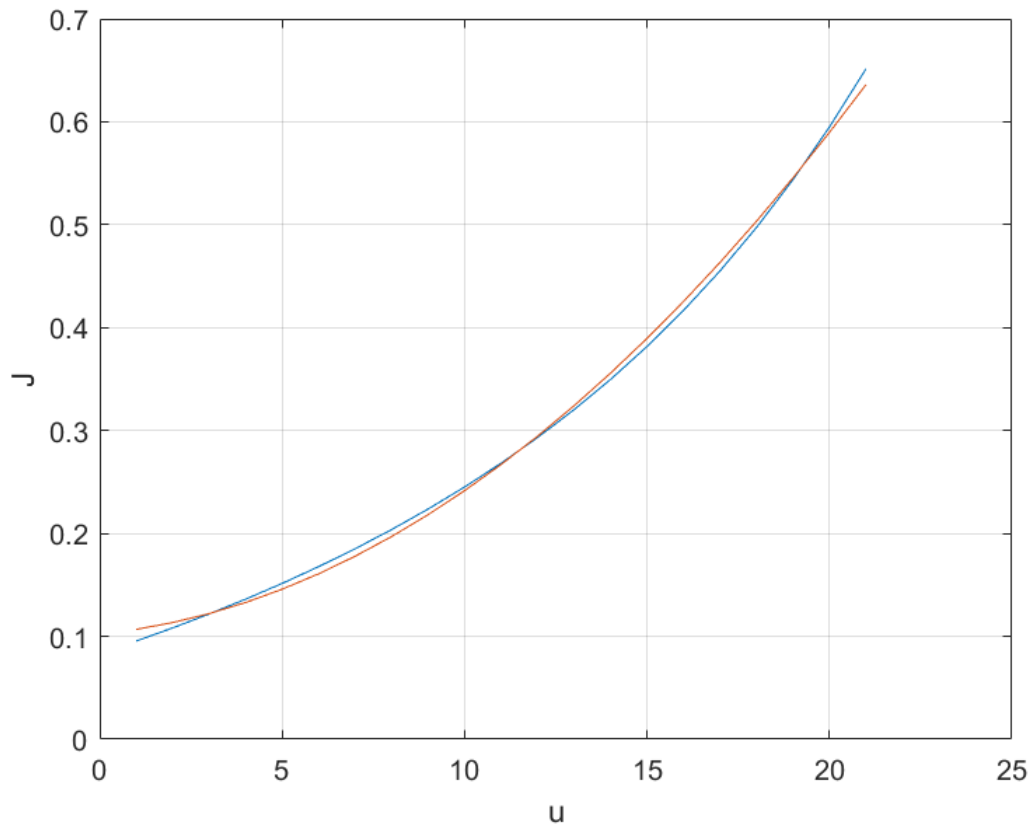
$$\bar{J} = \frac{1}{T} \int_0^T (\hat{p}_{in}(t) - \hat{p}_{out}(t)) dt, \quad (109)$$

gdzie $\hat{p}_{in}(t)$ oraz $\hat{p}_{out}(t)$ oznaczają średnie ciśnienie w chwili t na wlocie oraz wylocie, odpowiednio. Funkcja ta dla badanej geometrii w przypadku różnych prędkości $u \in [0.05; 0.15]$ na wlocie została zilustrowana na Rysunku (28).



Rysunek 28: Wartość funkcji celu w zależności od parametru kontrolnego

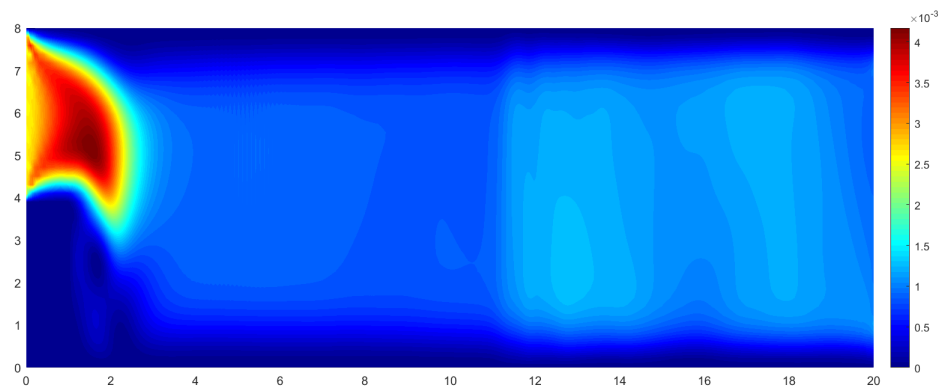
Zależność ta ma kształt paraboliczny, co jest zgodne z oczekiwaniami. Uwagę tę potwierdza Rysunek (29), który ilustruje krzywą na tle najlepiej dopasowanej doń paraboli (oznaczoną kolorem czerwonym)



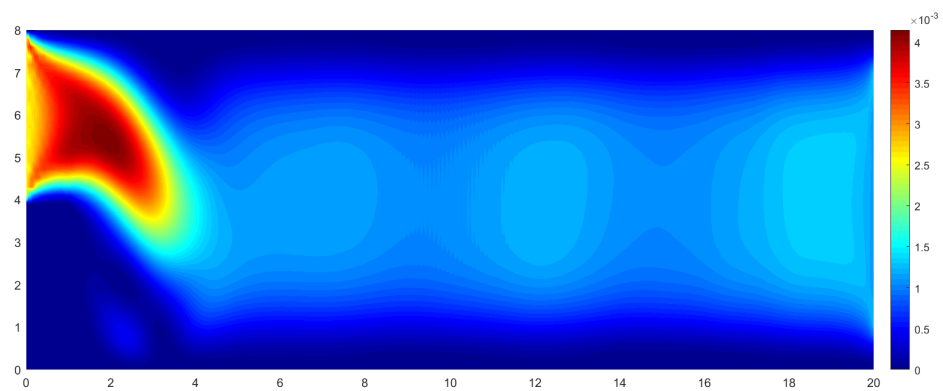
Rysunek 29: Najlepiej dopasowana parabola na tle krzywej funkcji celu

Przebadane zostało pięć przypadków różniących się liczbą Reynoldsa (modułem prędkości wlotowej). Rozmiary siatki pozostawały niezmiennie: w kierunku wlotu zastosowano podział na $m = 20$ komórek, zaś w kierunku poziomym $n = 10m$.

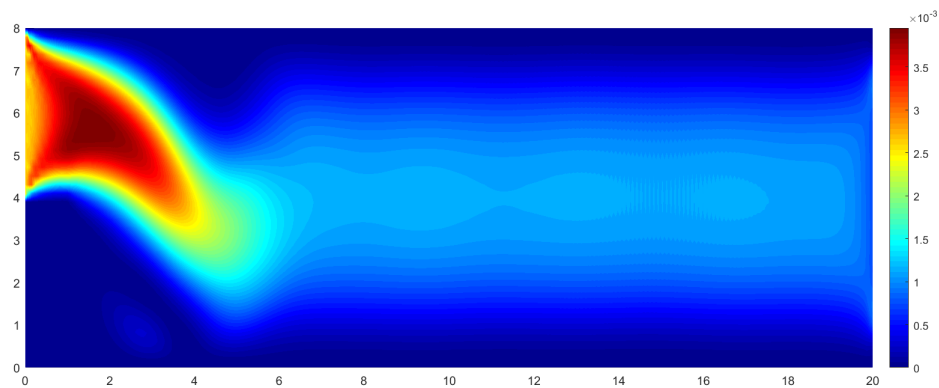
Liczba Reynoldsa: 200 Rysunki (30), (31), (32) oraz (33) przedstawiają mapy modułu prędkości po 1000, 2000, 3000 oraz 4000 kroków czasowych, odpowiednio (rysunki zostały odpowiednio przeskalowane w kierunku osi pionowej dla przejrzystości poglądu).



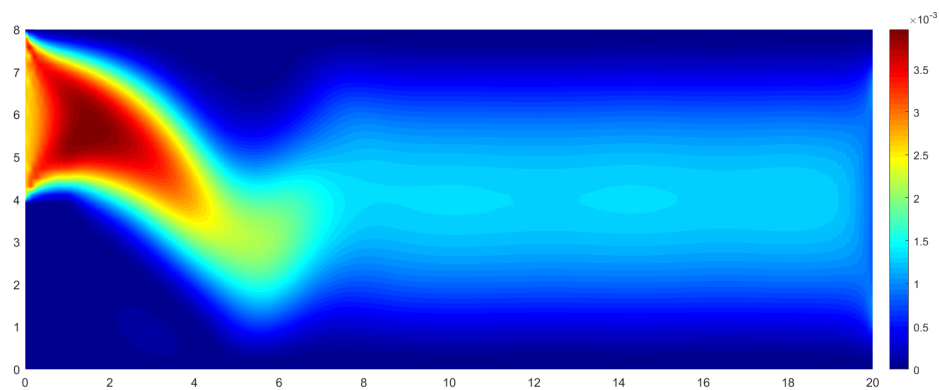
Rysunek 30: Mapa modułu prędkości po $T = 1000$ kroków czasowych



Rysunek 31: Mapa modułu prędkości po $T = 2000$ kroków czasowych

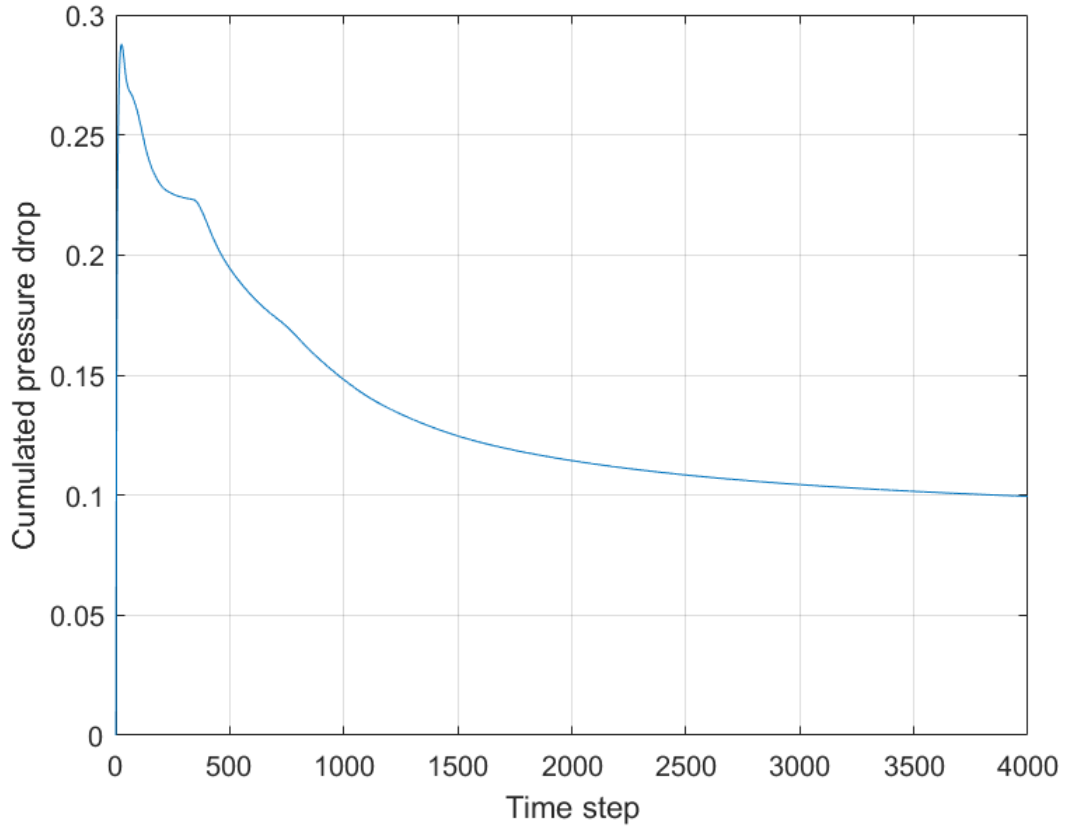


Rysunek 32: Mapa modułu prędkości po $T = 3000$ kroków czasowych



Rysunek 33: Mapa modułu prędkości po $T = 4000$ kroków czasowych

Dobłą charakterystyką jest średni skumulowany spadek ciśnienia (tj. średni spadek ciśnienia liczony od początku do danej chwili), którego wykres dla pierwszych $T = 4000$ kroków czasowych oraz dla parametru wlotowego $u = 0.05$ przedstawia Rysunek (34).



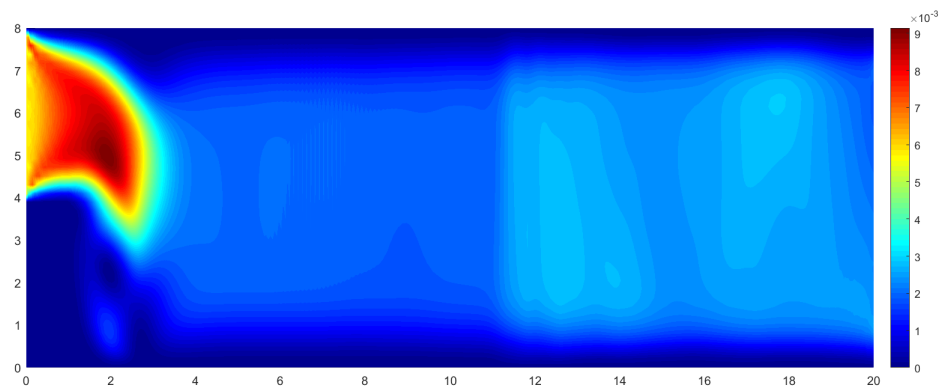
Rysunek 34: Średni skumulowany spadek ciśnienia

Wspomnijmy przy okazji, iż ergodyczność układu oznacza w tym przypadku, że taki średni skumulowany spadek ciśnienia będzie asymptotycznie dążył do tej samej wartości niezależnie od początkowego pola funkcji rozkładu, które w tej pracy inicjalizowane są poprzez rozkład równowagowy.

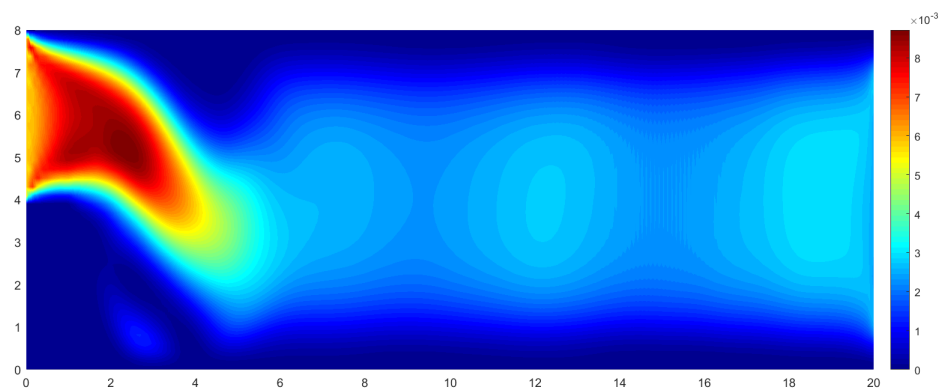
Wrażliwość wyznaczona poprzez długookresową symulację $grad\bar{J}$ oraz ta, wyznaczona z użyciem Metody Trajektorii Cienia $grad_{LSS}\bar{J}$, a także błąd względny $\epsilon = \frac{|grad\bar{J} - grad_{LSS}\bar{J}|}{grad\bar{J}}$ mają wartości

$$\begin{aligned} grad\bar{J} &= 5.72 \\ grad_{LSS}\bar{J} &= 5.23 \\ \epsilon &= 0.088 \end{aligned} \quad (110)$$

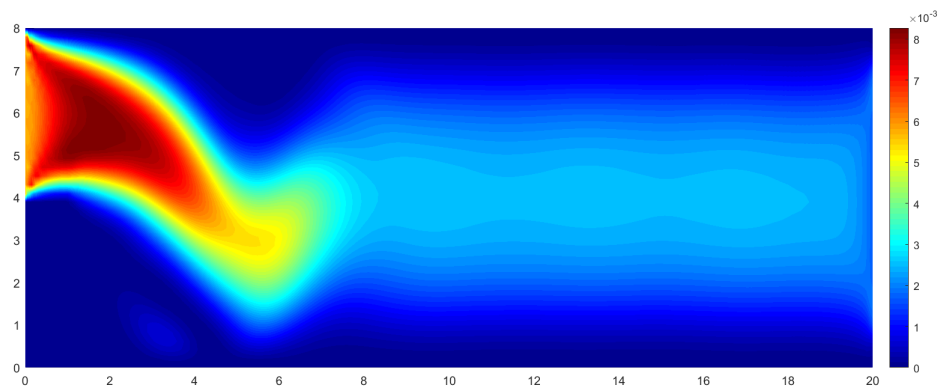
Liczba Reynoldsa: 300 Rysunki (35), (36), (37) oraz (38) przedstawiają mapy modułu prędkości po 1000, 2000, 3000 oraz 4000 kroków czasowych, odpowiednio



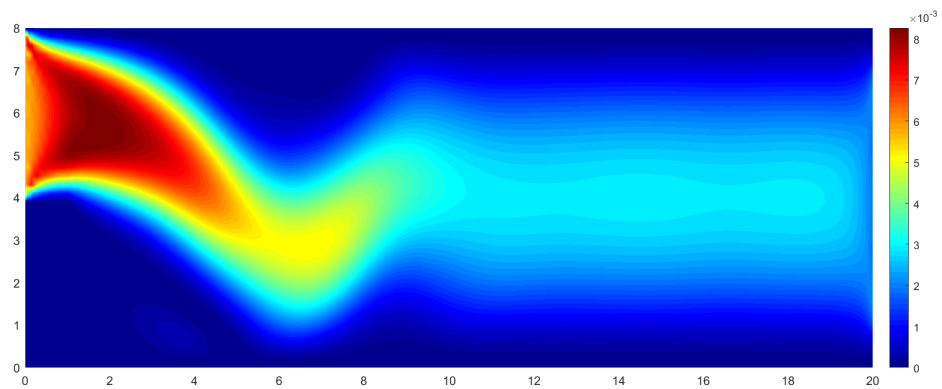
Rysunek 35: Mapa modułu prędkości po $T = 1000$ kroków czasowych



Rysunek 36: Mapa modułu prędkości po $T = 2000$ kroków czasowych

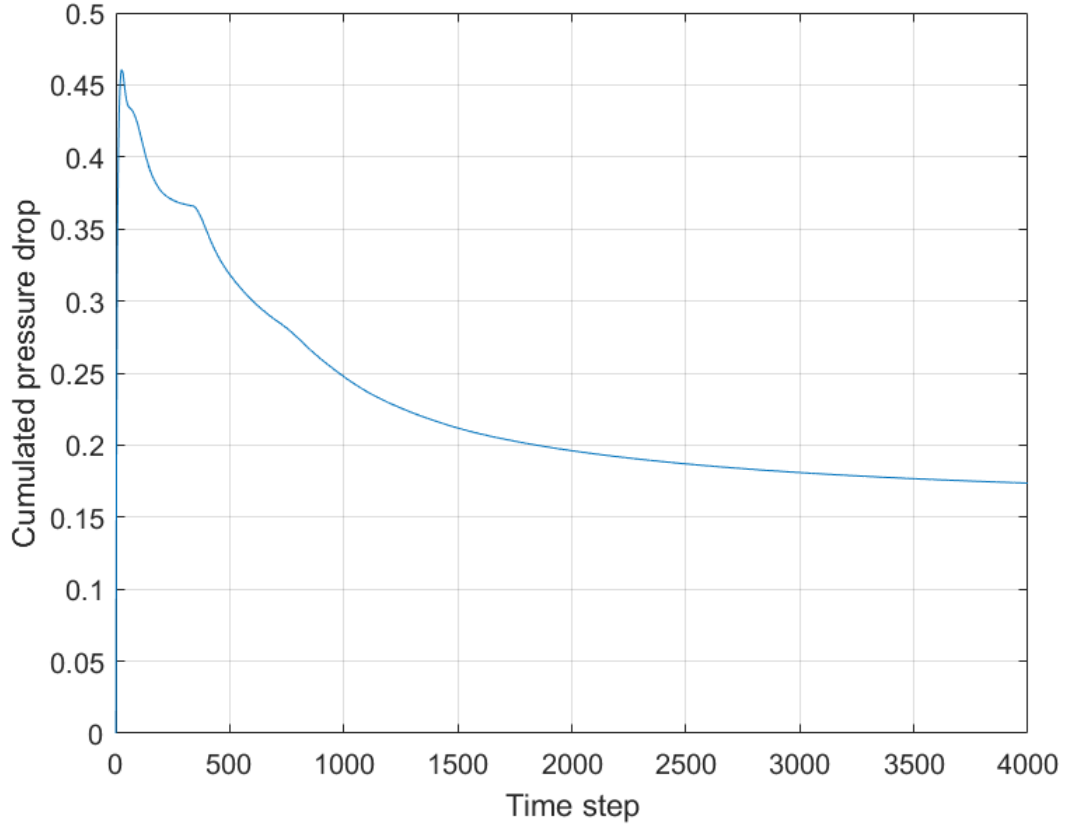


Rysunek 37: Mapa modułu prędkości po $T = 3000$ kroków czasowych



Rysunek 38: Mapa modułu prędkości po $T = 4000$ kroków czasowych

Średni skumulowany spadek ciśnienia ilustruje Rysunek (39).

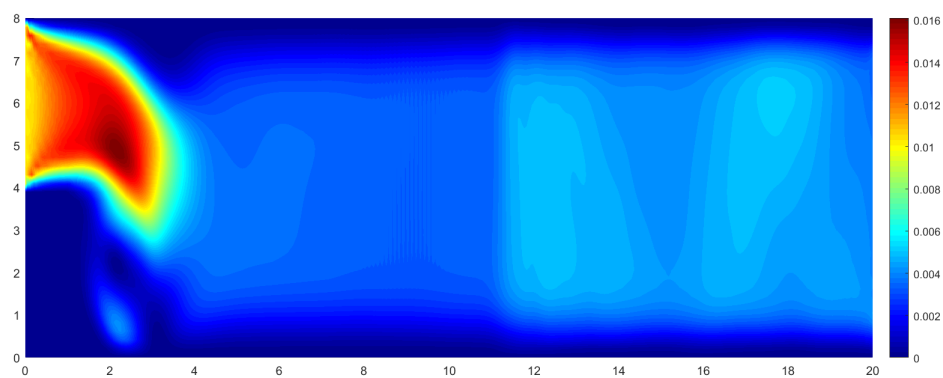


Rysunek 39: Średni skumulowany spadek ciśnienia

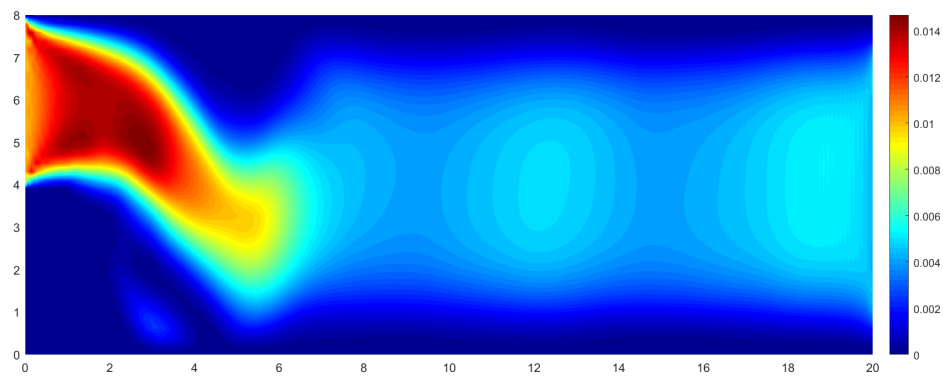
Wrażliwość wyznaczona poprzez długookresową symulację $grad\bar{J}$ oraz ta, wyznaczona z użyciem Metody Trajektorii Cienia $grad_{LSS}\bar{J}$, a także średni błąd względny ϵ mają wartości

$$\begin{aligned} grad\bar{J} &= 6.20 \\ grad_{LSS}\bar{J} &= 5.63 \\ \epsilon &= 0.0919 \end{aligned} \quad . \quad (111)$$

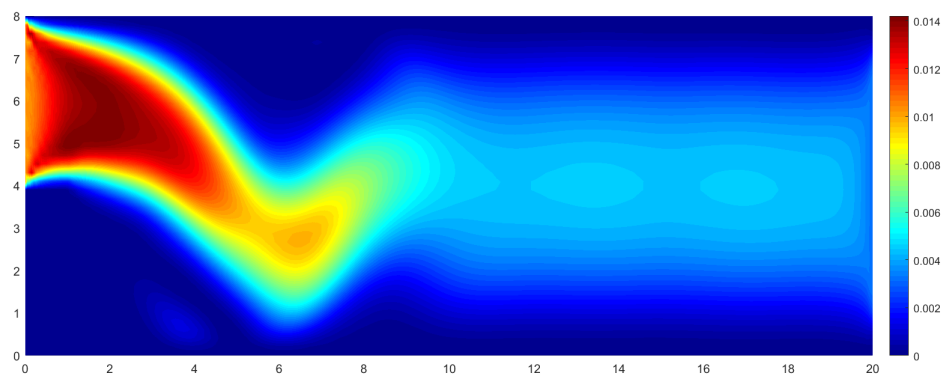
Liczba Reynoldsa: 400 Rysunki (40), (41), (42) oraz (43) przedstawiają mapy modułu prędkości po 1000, 2000, 3000 oraz 4000 kroków czasowych, odpowiednio



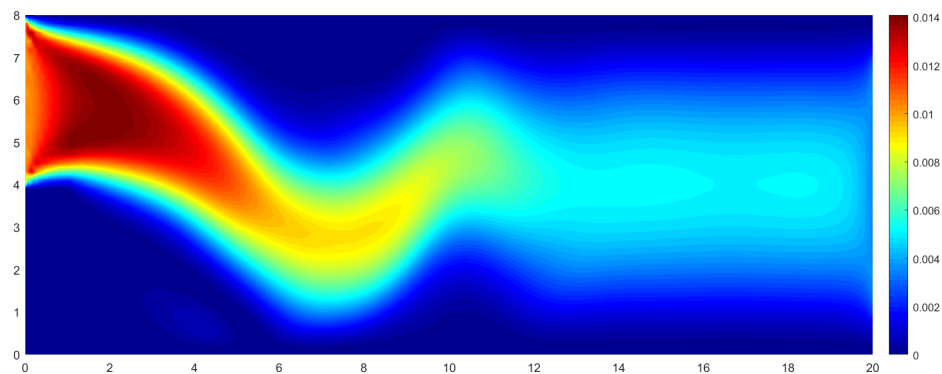
Rysunek 40: Mapa modułu prędkości po $T = 1000$ kroków czasowych



Rysunek 41: Mapa modułu prędkości po $T = 2000$ kroków czasowych

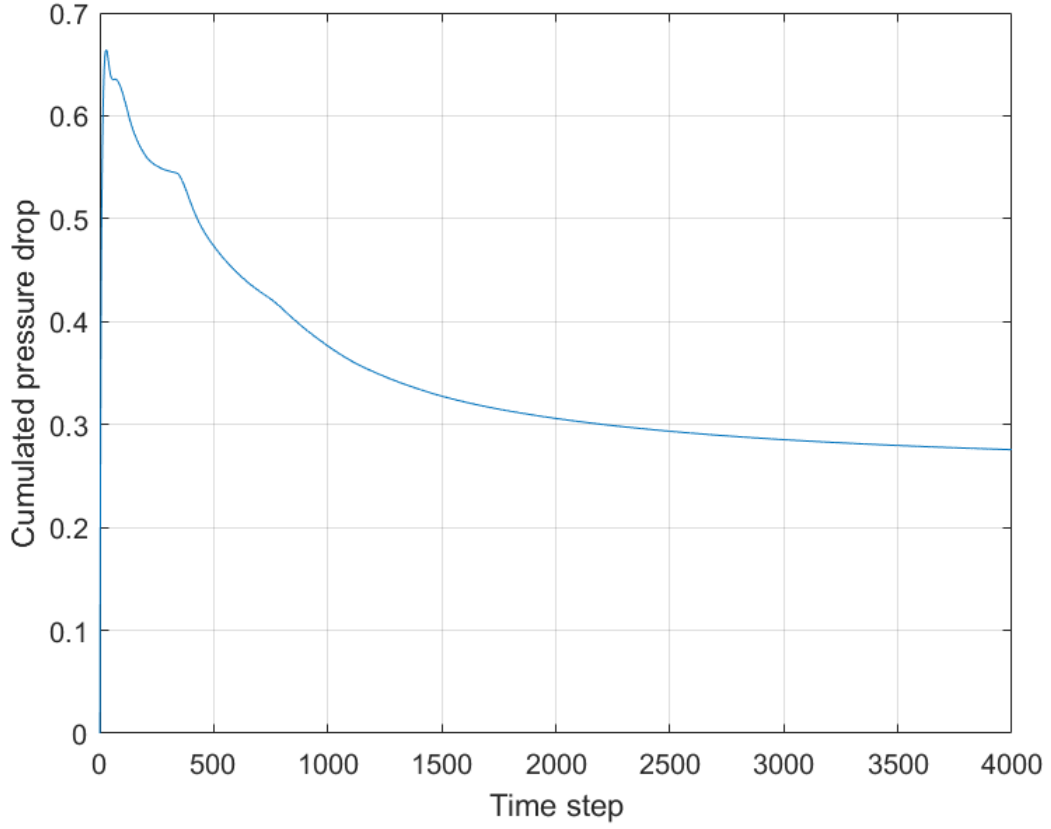


Rysunek 42: Mapa modułu prędkości po $T = 3000$ kroków czasowych



Rysunek 43: Mapa modułu prędkości po $T = 4000$ kroków czasowych

Średni skumulowany spadek ciśnienia ilustruje Rysunek (44).

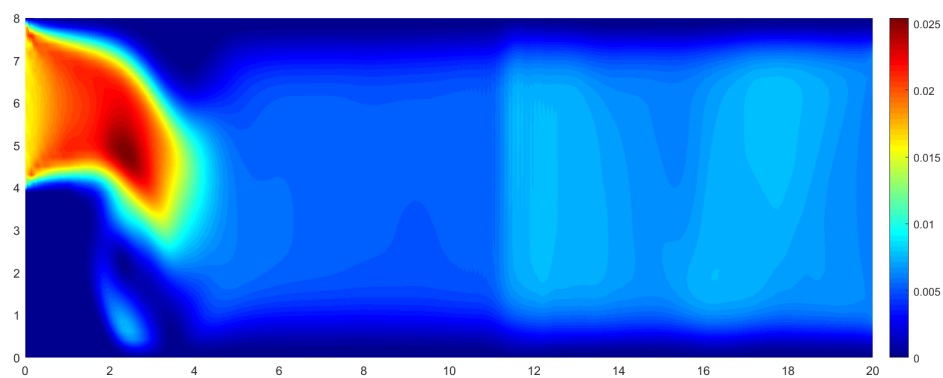


Rysunek 44: Średni skumulowany spadek ciśnienia

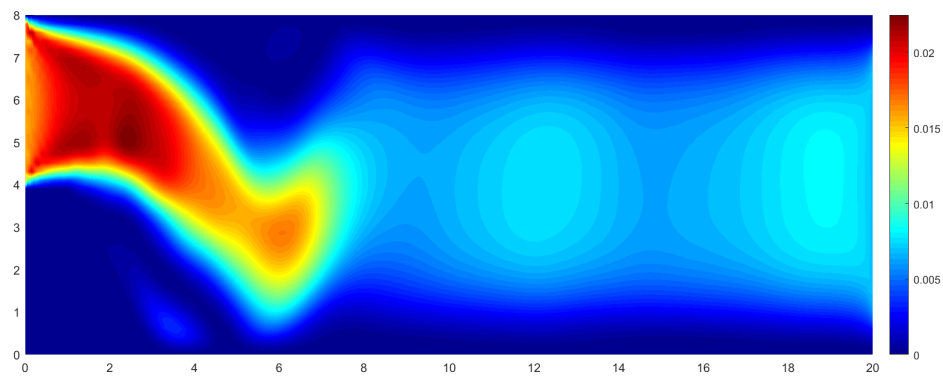
Wrażliwość wyznaczona poprzez długookresową symulację $grad\bar{J}$ oraz ta, wyznaczona z użyciem Metody Trajektorii Cienia $grad_{LSS}\bar{J}$, a także średni błąd względny ϵ mają wartości

$$\begin{aligned} grad\bar{J} &= 6.99 \\ grad_{LSS}\bar{J} &= 6.12 \\ \epsilon &= 0.1245 \end{aligned} \quad . \quad (112)$$

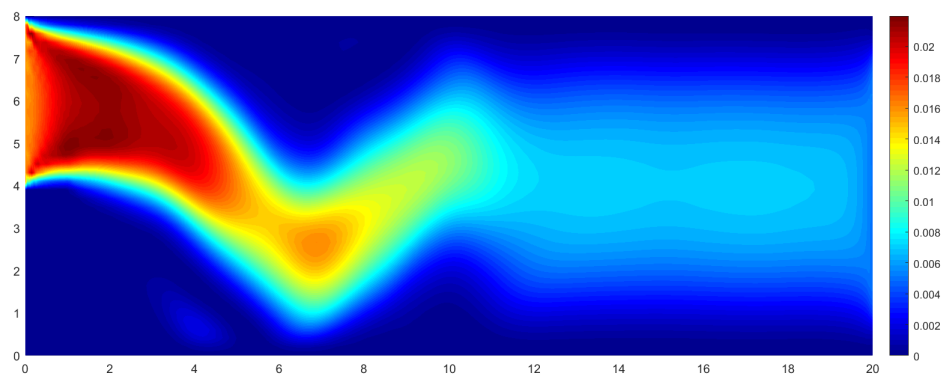
Liczba Reynoldsa: 500 Rysunki (45), (46), (47) oraz (48) przedstawiają mapy modułu prędkości po 1000, 2000, 3000 oraz 4000 kroków czasowych, odpowiednio



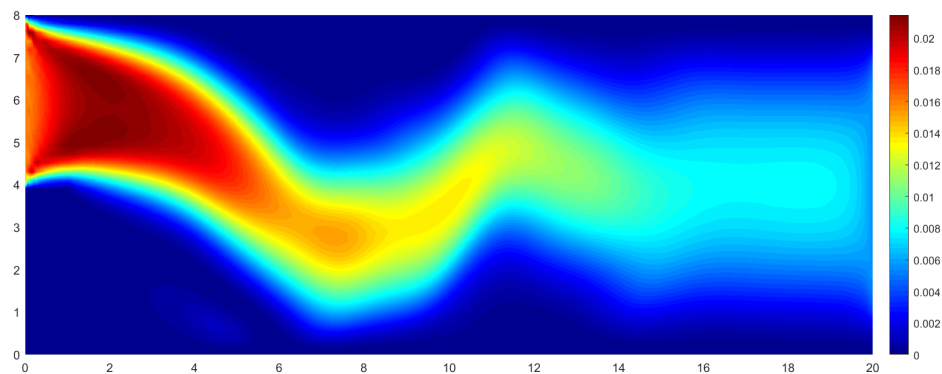
Rysunek 45: Mapa modułu prędkości po $T = 1000$ kroków czasowych



Rysunek 46: Mapa modułu prędkości po $T = 2000$ kroków czasowych

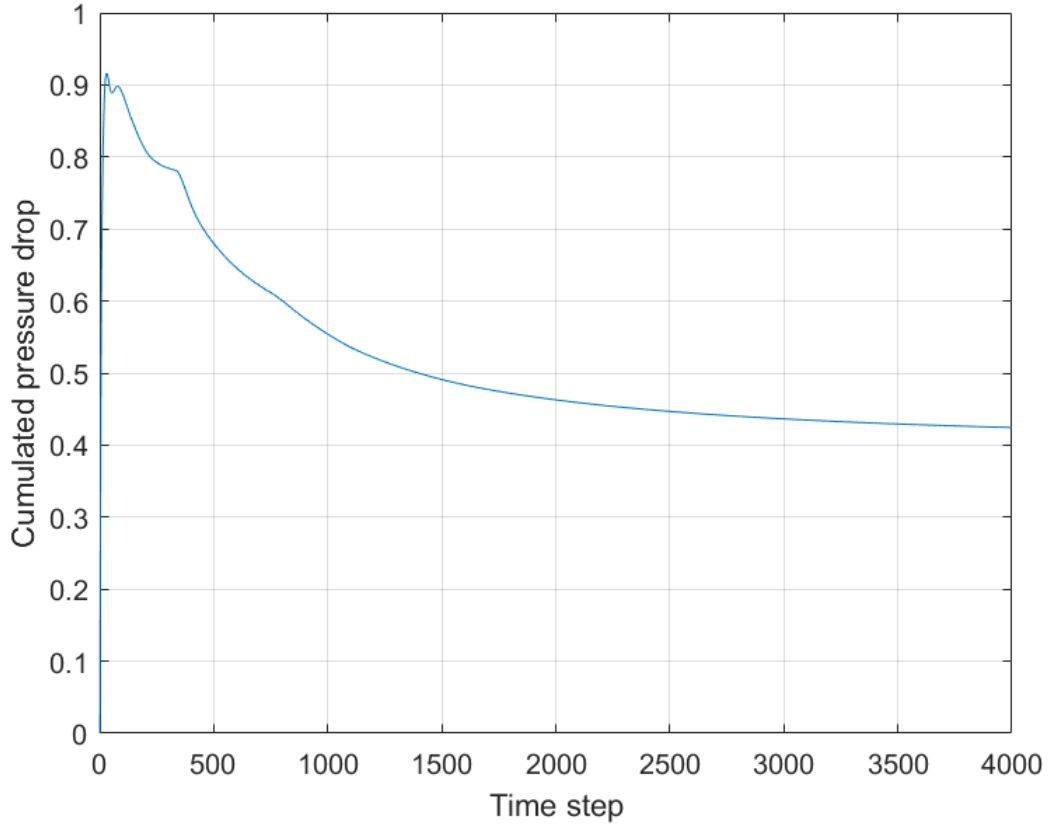


Rysunek 47: Mapa modułu prędkości po $T = 3000$ kroków czasowych



Rysunek 48: Mapa modułu prędkości po $T = 4000$ kroków czasowych

Średni skumulowany spadek ciśnienia ilustruje Rysunek (49).

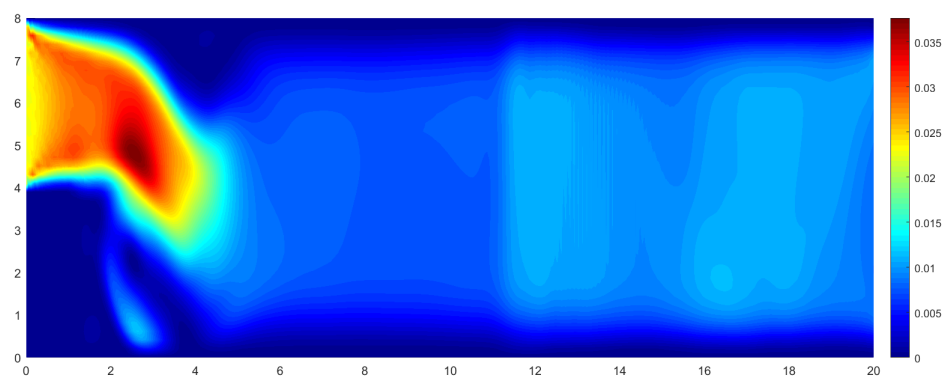


Rysunek 49: Średni skumulowany spadek ciśnienia

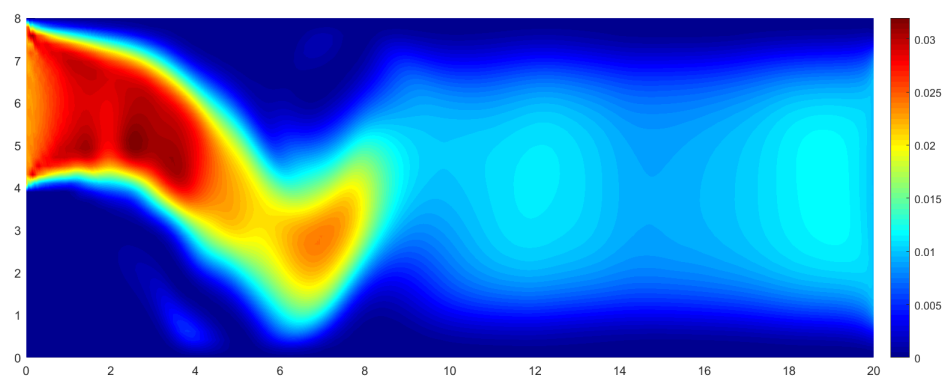
Wrażliwość wyznaczona poprzez długookresową symulację $grad\bar{J}$ oraz ta, wyznaczona z użyciem Metody Trajektorii Cienia $grad_{LSS}\bar{J}$, a także średni błąd względny ϵ mają wartości

$$\begin{aligned} grad\bar{J} &= 8.04 \\ grad_{LSS}\bar{J} &= 6.74 \\ \epsilon &= 0.1617 \end{aligned} \quad . \quad (113)$$

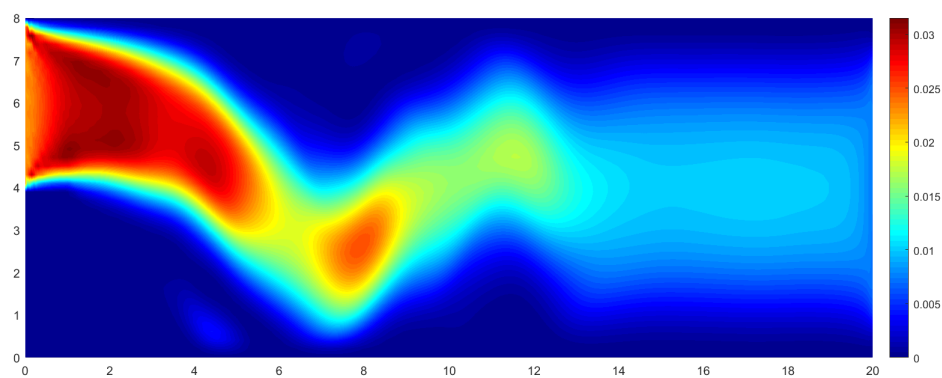
Liczba Reynoldsa: 600 Rysunki (50), (51), (52) oraz (53) przedstawiają mapy modułu prędkości po 1000, 2000, 3000 oraz 4000 kroków czasowych, odpowiednio



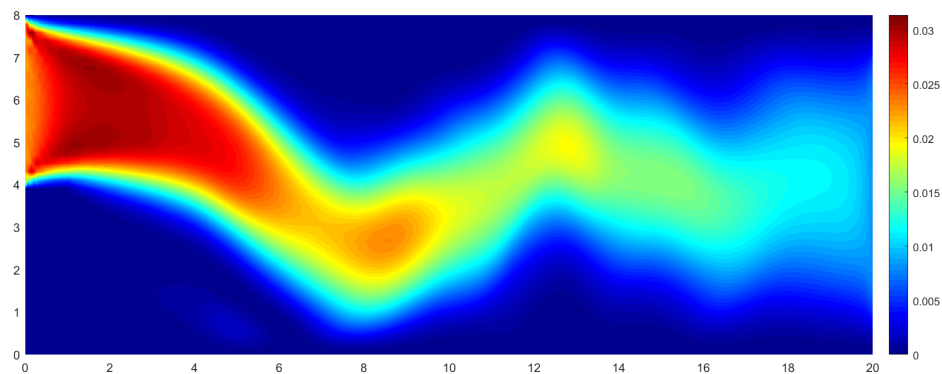
Rysunek 50: Mapa modułu prędkości po $T = 1000$ kroków czasowych



Rysunek 51: Mapa modułu prędkości po $T = 2000$ kroków czasowych

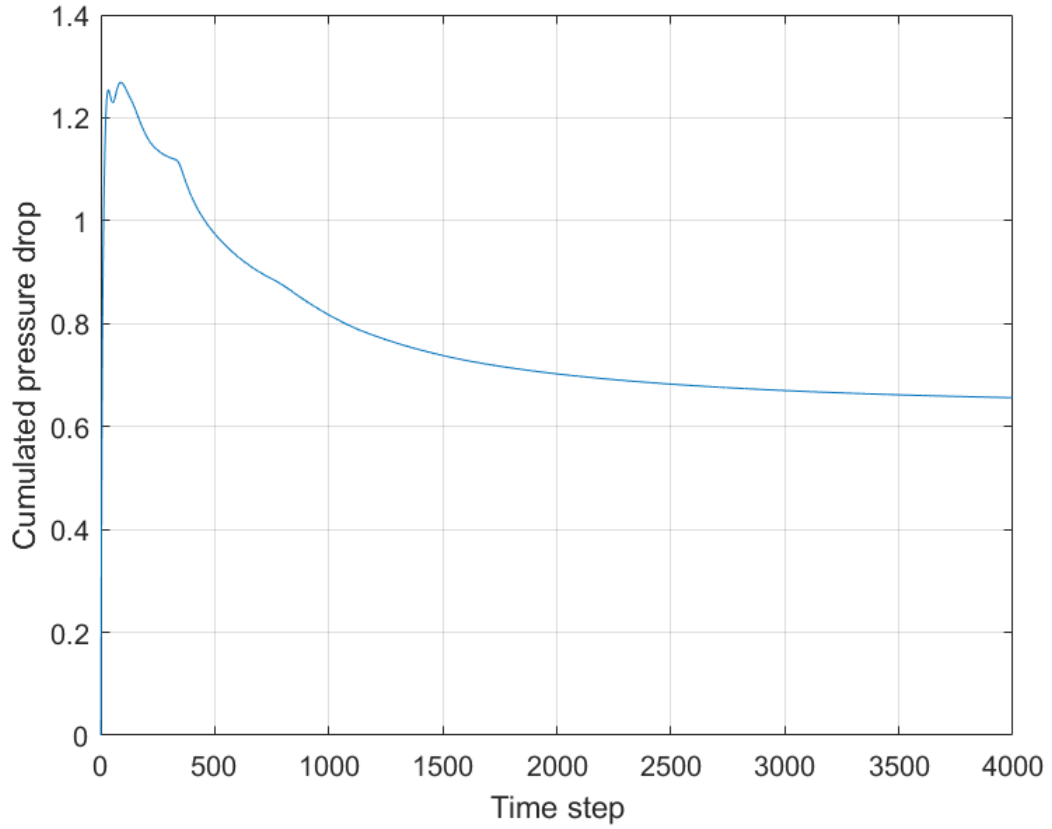


Rysunek 52: Mapa modułu prędkości po $T = 3000$ kroków czasowych



Rysunek 53: Mapa modułu prędkości po $T = 4000$ kroków czasowych

Średni skumulowany spadek ciśnienia ilustruje Rysunek (54).



Rysunek 54: Średni skumulowany spadek ciśnienia

Wrażliwość wyznaczona poprzez długookresową symulację $grad\bar{J}$ oraz ta, wyznaczona z użyciem Metody Trajektorii Cienia $grad_{LSS}\bar{J}$, a także błąd względny ϵ mają wartości

$$\begin{aligned} grad\bar{J} &= 9.52 \\ grad_{LSS}\bar{J} &= 7.57 \\ \epsilon &= 0.2048 \end{aligned} \quad . \quad (114)$$

Kod zaimplementowany w języku C++ jest dostępny na publiczny repozytorium autora [26].

6 Dyskusja wyników oraz wnioski

6.1 Scenariusz pierwszy

Wyniki symulacji dla układu Lorenza (1963) są zgodne z oczekiwaną wrażliwością funkcji celu. Zauważmy w tym miejscu, że dla przebadanych okresów symulacji (Rysunki) wyznaczona w sposób bezpośredni wrażliwość pozostaje równie bezużyteczna (Rysunki). Można stąd wysunąć wniosek, iż nawet jeżeli funkcja celu wizualnie wygląda na ‘wygładzoną’, ostatecznym weryfikatorem przydatności powinna być zawsze wyznaczona numerycznie wrażliwość.

Rysunki ilustrują, że takowa liczona względem trajektorii cienia jest bardzo stabilna. Dodatkowo zauważmy, iż przebadany został dość obszerny zakres parametrów projektowych.

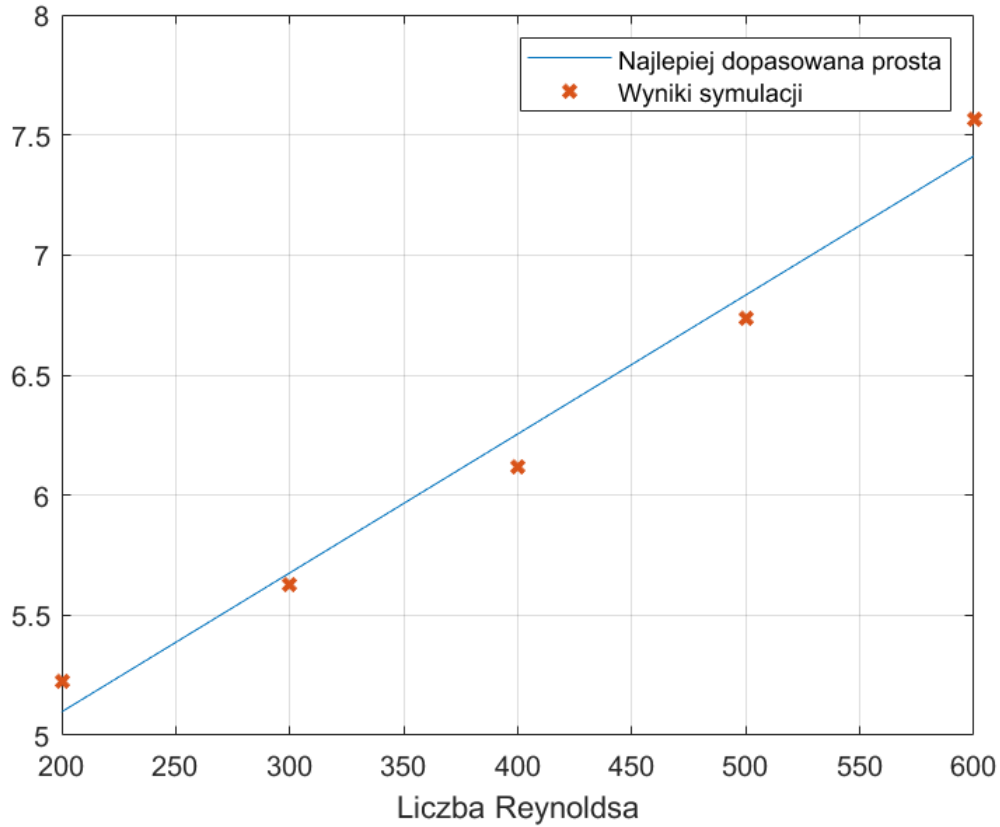
6.2 Scenariusz drugi

Na początku zauważmy, że w każdym z przebadanych przypadków przyjęte założenie ergodyczności ma uzasadnienie (dostarczone wyniki nie dają podstaw, aby to założenie odrzucić) – skumulowane spadki ciśnień dążą bowiem asymptotycznie do stałych wartości (Rysunki). Jest tak nawet pomimo różnych liczb Reynoldsa oraz wynikających z nich jakościowych zmian rozwiązań ilustrowanych poprzez mapy modułów prędkości (Rysunki). Funkcja celu (109) dla średniego w czasie spadku ciśnienia ma charakter paraboliczny, co jest zgodne z danymi literaturowymi. Stanowi to wskazówkę, iż w odpowiadających analizowanym równaniom Boltzmanna równaniach Naviera-Stokesa człon nieliniowy nie mógłby zostać zlinearyzowany bez wprowadzania istotnych błędów.

Dokładność wyników podsumowuje poniższa tabela. Błąd względny rośnie wraz ze wzrostem liczby Reynoldsa. Obserwacja taka może być uzasadniona tym, że nieliniowości zaczynają dominować dynamikę zachowania rozwiązań i w związku z tym rośnie liczba niestabilnych modów Lyapunova. Rysunek ilustruje wykres błędu względnego w funkcji liczby Reynoldsa. Charakter paraboliczny wskazuje potencjalne problemy dla bardzo dużych liczb Reynoldsa. W takich przypadkach konieczne byłoby zastosowanie dłuższych okresów symulacji (poprzedni scenariusz sugeruje, iż taki zabieg ma działanie regularyzujące).

Zauważmy też, że wrażliwości wyznaczone omawianą metodą niedoszacowują wartości referencyjnych. Taki stan rzeczy może być spowodowany skończonym czasem symulacji. Pamiętajmy, że wyznaczona trajektoria cienia pozostaje w bliskim sąsiedztwie trajektorii bazowej, jednakże nie ma gwarancji, że jest ona przez cały okres w tej samej odległości. Dysponujemy jedynie ograniczeniem górnym, które w konkretnych przypadkach może wpływać na dokładność uzyskanych wyników.

Spostrzegamy na koniec, że wyniki tworzą silnie liniowy trend. Najlepiej dopasowana prosta pokazana na tle wyników (Rysunek (55)) ilustruje tę uwagę.



Rysunek 55: Najlepiej dopasowana prosta

Wartość współczynnika korelacji Pearsona, zdefiniowanego następująco:

$$r = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y}, \quad (115)$$

która wynosi $r = 0.9899$ stanowi dodatkowe potwierdzenie (X, Y oznaczają zmienne losowe, w rozważanym przypadku są to: liczba Reynoldsa oraz wyznaczona wrażliwość, odpowiednio).

Powyższe uwagi są oczywiście efektem oczekiwanym, wobec parabolicznego charakteru funkcji celu.

6.3 Konkluzje

Metoda Trajektorii Cienia w zaimplementowanej formie stanowi bardzo obiecujące narzędzie w analizie wrażliwości funkcji opisujących układy chaotyczne. Sam algorytm, chociaż oparty na dość zaawansowanych wynikach z teorii chaosu, ma bardzo intuicyjną interpretację. Sformułowanie w formie zagadnienia minimalizacji formy całkowej z ograniczeniami w postaci prawa rządzącego dynamiką wrażliwości na parameter projektowy jest dobrze

określone w przeciwieństwie do wyjściowego problemu, który przyjmuje postać zagadnienia początkowego. Skutkuje to dobrą zbieżnością rozwiązania związanego z nim układu równań.

Należy również zwrócić uwagę, iż metoda jest w pełni deterministyczna – w żadnym kroku nie są generowane liczby losowe (pseudolosowe), a więc wyniki dla tych samych danych wejściowych pozostają takie same. Odtwarzalność taka jest ważna w kontekście praktycznych zastosowań, a także ułatwia analizę zbieżności metody. Co więcej, wyniki niniejszej pracy pokazują, że rezultaty są numerycznie stabilne – niewielkie zmiany danych wejściowych nie powodują znacznych zmian w wynikach, a sama dokładność jest na stosunkowo wysokim poziomie.

Dalsze ograniczenia błędów można uzyskać poprzez wydłużenie okresu symulacji, co przeważnie ‘wygładza’ statystyki układu. Wniosek ten jak i możliwość zastosowania opisanej metody dotyczą szerokiego zakresu układów dynamicznych. Należy tutaj jednak stanowczo zaznaczyć, iż konieczne jest, aby układ taki był ergodyczny. Brak takiego założenia skutkuje w ogólności zależnością długookresowych średnich od warunków początkowych, co jest głównym powodem złego uwarunkowania zadania wyjściowego.

Wybór między metodą optymalizacji (Tangent lub też Metoda Operatora Sprzężonego) jest względnie prosty ze względu na tę samą macierz główną układu. Wspomniane zostało już, iż pierwsza z nich lepiej nadaje się do zagadnień, w których wymiar wektora zmiennych decyzyjnych jest niski. Dla dużej liczby zmiennych projektowych Metoda Operatora Sprzężonego jest właściwszym wyborem. W przypadku pośrednim stosunkowo łatwo jest dobrać optymalną metodę poprzez ich porównanie wymagające jedynie zmiany wektora prawych stron, co z punktu widzenia istniejącej implementacji jest stosunkowo prostym zadaniem.

Dotychczasowe próby opisane w literaturze z punktu widzenia numerycznego zawsze polegały na przechowywaniu w pamięci macierzy głównej układu. W Rozdziale 5.2.4 pokazane zostało, że takie podejście jest bardzo kosztowne pod kątem pamięci operacyjnej. Eliminuje to użyteczność tej metody w większych zagadnieniach (w szczególności w przepływach trójwymiarowych). Stosowany tutaj pomysł polega na wykorzystaniu podejścia nie wymagającego przechowywania wspomnianej macierzy w pamięci (poprzez zaimplementowanie jej działania na dowolny wektor w postaci funktora) oraz technik automatycznego różniczkowania kodu (z wykorzystaniem otwartych bibliotek) do wyznaczenia współczynników takiej macierzy (narzędzia te nie generują błędów zaokrągleń, więc wyznaczane pochodne są bardzo dokładne – ograniczone jedynie precyzją maszynową).

Najistotniejszym elementem niniejszej pracy było zastosowanie Metody Trajektorii Cienia do zagadnienia przepływowego opisywanego równaniami Boltzmanna. W momencie pisania autor nie natknął się na inne tego typu próby. Niemniej praca ta daje podstawy, aby sądzić, iż jest to zadanie wykonalne, a uzyskane wyniki są obiecujące. Jest to o tyle ważne, ponieważ metoda dyskretyzacji równań ruchu płynu bazująca na podejściu Boltzmanna staje się coraz bardziej popularna ze względu na dobre własności numeryczne równań – w szczególności na możliwość masowej paralelizacji kodów obliczeniowych. Oprócz wspomnianych zalet Metoda Trajektorii Cienia (jak każda inna) posiada wady. Patrząc krytycznym okiem można wspomnieć co najmniej kilka z nich:

- Metoda jest trudna do zastosowania w gotowych kodach obliczeniowych ze względu

na konieczność zróżniczkowania kodu (w przeciwnym przypadku zajętość pamięci jest nieakceptowalnie wysoka).

- Pomimo istotnego ograniczenia wymaganej pamięci operacyjnej wciąż mogą pojawić się problemy dla naprawdę dużych zagadnień. Nadal bowiem konieczne jest przechowywanie całej trajektorii układu.
- Innym problemem wspomnianym na przykład w pracach [10], [11] jest wybór trajektorii, która jest niereprezentatywna ze względu na zaburzenie. Kłopot tkwi w tym, że wszystkie wspomniane własności (patrz Rozdział 2) zachodzą w sensie odpowiedniej miary, tj. nie zachodzą na zbiorach, których ta miara jest zerowa. Na przykład wspomniane już w Rozdziale 2. trajektorie okresowe charakteryzują się takimi własnościami. Wybierając taką trajektorię nie mamy gwarancji regularnego jej zachowania. W pracy [17] odnaleziono przykłady, w których taki zbiór (miary zero) jest gęsty w całej przestrzeni, co zwiększa szansę na wybór niereprezentatywnej trajektorii.

Wszystkie powyższe mankamenty wymagają dodatkowych badań. Obiecującym rozwiązaniem może okazać się na przykład podejście hybrydowe. Polega ono na połączeniu Metody Trajektorii Cienia z klasą metod opartych na uśrednianiu w przestrzeni całkowitej. Przykład takiego algorytmu, który w istocie bazuje na rozbiciu przestrzeni fazowej – S3 (ang. Space Split Sensitivity algorithm) zaprezentowany jest w pracach [10] oraz [11].

Literatura

- [1] Abramov R., Majda A., *Blended response algorithms for linear fluctuation-dissipation for complex nonlinear dynamical systems*, Nonlinearity, Vol. 20, No. 12, 2007, pp. 2793.
- [2] Bewley T., *Flow control: new challenges for a new Renaissance*, Progress in Aerospace Sciences, Vol. 37, No. 1, 2001, pp. 21-58.
- [3] Bewley T., Moin P., Temam R., *DNS-based predictive control of turbulence: an optimal target for feedback algorithms*, J. Fluid Mech., Vol. 447, 2001, pp. 179-225.
- [4] Blackmore D., Prykarpatsky K. A., Samoylenko V. H., *Nonlinear Dynamical Systems Of Mathematical Physics: Spectral And Symplectic Integrability Analysis*, Wydawnictwo: World Scientific, Singapore 2011.
- [5] Blonigan P. J., Wang Q., *Probability density adjoint for sensitivity analysis of the mean chaos*, Journal of Computational Physics, 235 (2014) 1-13.
- [6] Blonigan P. J., Wang Q., Nielsen E. J., Diskin B., *Least Squares Shadowing sensitivity analysis of chaotic flow around a tw-dimensional airfoil*, in: 54th AIAA Aerospace Sciences Meeting, January 2016, pp. 1-28.
- [7] Blonigan P. J., Wang Q., *Least Squares Shadowing sensitivity analysis of a modified Kuramoto-Sivashinsky equation*, Chaos Solitons Fractals, 64, 2014, 16-25.
- [8] Cavaglieri D., Beyhaghi P., Bewley T., *Low-storage IMEX Runge-Kutta schemes for the simulations of Navier-Stokes systems* 21st AIAA computational fluid dynamics conference, San Diego, CA, June 2013.
- [9] Chandramoorthy N., Fernandez P., Talnikar C., Wang Q. *An Analysis of the Ensemble Adjoint Approach to Sensitivity Analysis in Chaotic Systems* AIAA Paper 2017-3799.
- [10] Chandramoorthy N., Wang Z. N., Tucker P., *Toward computing sensitivities of average quantities in turbulent flows*, Center for Turbulence Research, 2018.
- [11] Chandramoorthy N., Wang Q., *Sensitivity computation of statistically stationary quantities in turbulent flows* Massachusetts Institute of Technology, Cambridge, MA, 02139.
- [12] Chater M., Ni A., Wang Q., *Simplified Least Squares Shadowing sensitivity analysis for chaotic ODEs and PDEs*, Journal of Computational Physics 329 (2017) 126-140.
- [13] Espinoza, Mayken *Compress effects on porosity, gas-phase tortuosity, and gas permeability in a simulated PEM gas diffusion layer*, International Journal of Energy Research. 39 (11): 1528–1536
- [14] Evans L.C., *Równania różniczkowe cząstkowe*, Wydawnictwo: PWN, Warszawa 2012.

- [15] Gleick J., *Chaos. Narodziny nowej nauki.*, Wydawnictwo: Zysk i S-ka, 2018.
- [16] Golub G. H., Loan C. V. F., *Matrix Computations*, The Johns Hopkins Univ. Press, Baltimore, 1996.
- [17] Grebogi C., Ott E., Yorke J. A., *Unstable periodic orbits and the dimensions of multifractal chaotic attractors*, Phys. Rev. A 37, 1711, 1988.
- [18] Hascoet L., Pascual V., *TAPENADE 2.1 user's guide*, HAL Id: inria-00069880
- [19] Jameson A., *Aerodynamic design via control theory*, Journal of Scientific Computing 3 (1988) 233-260.
- [20] Landau L. D., Lifszyc J. M., *Hydrodynamika*, Wydawnictwo: PWN, Warszawa 2009.
- [21] Kudrewicz J., *Fraktale i chaos*, Wydawnictwo: WNT, 1996.
- [22] Lea D., Allen M., Haine T., *Sensitivity analysis of the climate of a chaotic system*, Tellus, Vol. 52A, 2000, pp. 523-532.
- [23] Lea D., Haine T., Allen M., Hansen J., *Sensitivity analysis of the climate of a chaotic ocean circulation model*, Journal of the Royal Meteorological Society, Vol. 128, 2002, pp. 2587-2605.
- [24] Lorenz E., *Deterministic Nonperiodic Flow*, Journal of Atmospheric Sciences, Vol. 20, 1963, pp. 130-141.
- [25] Marek A., *ChaoticSystemsOptimization*, <https://github.com/BlooAM/ChaoticSystemsOptimization>, 2019.
- [26] Marek A., *LSS*, <https://github.com/BlooAM/LSS>, 2019.
- [27] Maurin K., *Analiza, część I*, Wydawnictwo: PWN, Warszawa 2010.
- [28] Maurin K., *Analiza, część II*, Wydawnictwo: PWN, Warszawa 2010.
- [29] Maurin K., *Analiza, część III*, Wydawnictwo: PWN, Warszawa 2010.
- [30] Palczewski A., *Równania różniczkowe zwyczajne. Teoria i metody numeryczne z wykorzystaniem programu rachunków symbolicznych*, Wydawnictwo: WNT, Warszawa 2017.
- [31] Ruelle D., *Differentiation of SRB states*, Communications in Mathematical Physics, Vol. 187, 1997, pp. 227-241.
- [32] Ruelle D., *A review of linear response theory for general differentiable dynamical systems*, Nonlinearity 22 (4) (2009) 855.
- [33] Salazar J., Santos L. *2d lattice-boltzmann simulations of flow over a backward-facing step at low to moderate reynolds numbers*, GRAD - Fluid Physics and Transport Phenomena Research Group, Universidade Federal de Santa Catarina

- [34] Strogatz S., *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry and Engineering*, Westview Press, Philadelphia, PA, 1994.
- [35] Tempczyk M., *Teoria chaosu dla odważnych*, Wydawnictwo: PWN, Warszawa 2002.
- [36] Teman R., *Navier-Stokes Equations and Nonlinear Functional Analysis*, Wydawnictwo: SIAM, Philadelphia 1983.
- [37] Thepaut J-N., Courtier P., *Four-dimensional variational data assimilation using the adjoint of a multilevel primitive-equation model*, Quarterly Journal of the Royal Meteorological Society 117 (502) (1991) 1225-1254.
- [38] Thuburn J., *Climate sensitivities via a Fokker-Planck adjoint approach*, Quarterly Journal of the Royal Meteorological Society, Vol. 131, No. 605, 2005, pp. 73-92.
- [39] Wang Q., *Forward and adjoint sensitivity computation of chaotic dynamical systems*, Journal of Computational Physics, 235 (2013) 1-13.
- [40] Wang Q., *Convergence of the Least Squares Shadowing Method for Computing Derivative of Ergodic Averages*, SIAM Journal of Numerical Analysis, Vol 52, No. 1, 2014, pp. 156-170.
- [41] Wang Q., *Least Squares Shadowing sensitivity analysis of chaotic limit cycle oscillations*, Journal of Computational Physics, Vol. 267, June 2014, pp. 210-224.
- [42] Venditti D., Darmofal D., *Grid adaptation for functional outputs: Application to two-dimensional inviscid flow*, Journal of Computational Physics 176(2002) 4069.
- [43] <https://plato.stanford.edu/entries/chaos/>,