
Dynamika stochastyczna

Wydanie I

Jerzy Łuczka, Łukasz Machura

18/11/2013

Spis treści

1	Dynamika deterministyczna	3
1.1	Opis i modelowanie zjawisk oraz procesów przy pomocy równań różniczkowych	3
1.2	Modelowanie z czasem dyskretnym	8
1.3	Istnienie i jednoznaczność rozwiązań	9
1.4	Układy dynamiczne z czasem ciągłym	16
1.5	Dynamiczne układy zachowawcze i dysypatywne	24
1.6	Stany stacjonarne i ich stabilność	34
1.7	Atraktory	45
1.8	Dyskretnie układy dynamiczne	50
2	Chaos w układach dynamicznych	61
2.1	Determinizm a przewidywalność	61
2.2	Model chaosu. Układ bistabilny (oscylator Duffinga)	62
2.3	Scenariusz przejścia do chaosu	80
3	Dynamika stochastyczna	97
3.1	Równania stochastyczne i ich interpretacja	97
3.2	Podstawowe procesy stochastyczne	103
3.3	Rachunki w całkach Ito - Stratonowicza	117
3.4	Modele szumów	121
3.5	Proces Ornsteina-Uhlenbecka	129
3.6	Geometryczny ruch browna	136
4	Dodatek numeryczny	141
4.1	Liczby losowe	141
4.2	Równania różniczkowe zwyczajne	153
4.3	Stochastyczne równania różniczkowe	159
4.4	Przykłady całkowania procesów stochastycznych	163
4.5	Szum dychotomiczny	168
4.6	Szum kangura	169

Autorzy Jerzy Łuczka, Łukasz Machura

Wersja 0.7 11/2013

Pobierz podręcznik (v0.7, PDF)

Dynamika deterministyczna

1.1 Opis i modelowanie zjawisk oraz procesów przy pomocy równań różniczkowych

Jednym z podstawowych praw fizyki, jakie poznajemy w szkole średniej jest II zasada dynamiki Newtona. Opisuje ona klasyczne układy mechaniczne. Układy te są idealizacją realnych układów występujących w otaczającym nas świecie. W najprostszej wersji II zasada dynamiki Newtona w odniesieniu do jednej cząstki poruszającej się tylko wzdłuż jednej osi współrzędnych, np. wzdłuż osi OX, może być sformułowana w następującej postaci:

Ruch cząstki jest zdeterminowany przez siły jakie działają na cząstkę

Z punktu widzenia matematycznego, ruch cząstki opisany jest przez równanie Newtona:

$$ma = F \quad (1.1)$$

W równaniu tym występują trzy wielkości:

m to masa cząstki a jest przyspieszeniem cząstki F jest siłą działającą na cząstkę. Ponieważ ruch zachodzi tylko wzdłuż osi OX (tak zakładamy dla prostoty), siła F działa tylko w kierunku OX oraz przyspieszenie a jest wzdłuż osi OX.

Wiemy z kursu fizyki, że przyspieszenie cząstki jest pochodną (względem czasu) pierwszego rzędu prędkości v cząstki. Z kolei prędkość cząstki v jest pochodną pierwszego rzędu położenia cząstki x .

$$a = \frac{d}{dt}v = \frac{d}{dt}\frac{d}{dt}x = \frac{d^2x}{dt^2} \quad (1.2)$$

W ogólnej postaci siła

$$F = F(x, v, t) = F(x, dx/dt, t) \quad (1.3)$$

może zależeć od położenia x cząstki, jej prędkości $v = dx/dt$ oraz czasu t .

Jeżeli przyspieszenie a oraz siłę F w takiej postaci podstawimy do równania Newtona, to jego postać jest następująca:

$$m \frac{d^2 x}{dt^2} = F \left(x, \frac{dx}{dt}, t \right) \quad (1.4)$$

W ten sposób otrzymujemy równanie różniczkowe, które opisuje jednowymiarowy ruch cząstki wzdłuż osi OX. Co możemy powiedzieć o tym równaniu:

Jest to równanie różniczkowe drugiego rzędu, ponieważ pojawia się pochodna drugiego rzędu $d^2 x/dt^2$. Jest to równanie różniczkowe zwyczajne, ponieważ nie występują pochodne cząstkowe a jedynie pochodne ze względu na jedną zmienną - w tym przypadku pochodne względem czasu t . Samo równanie Newtona nie wystarczy, aby opisać ruch cząstki. Musimy zadać warunki początkowe dla tego równania. Ponieważ jest to równanie drugiego rzędu, musimy zadać dwa warunki początkowe: początkowe położenie $x(t_0) = x_0$ oraz początkową prędkość $v(t_0) = v_0$. Warunki początkowe można zadać w dowolnej chwili czasu t_0 , ale zazwyczaj tą chwilą początkową jest umowna chwila $t_0 = 0$. Równanie (1.4) możemy zatem przedstawić w równoważnej postaci:

$$\frac{dx}{dt} = v \quad (1.5)$$

$$\frac{dv}{dt} = \frac{1}{m} F(x, v, t) \quad (1.6)$$

gdzie wprowadziliśmy nową zmienną v która ma interpretację prędkości cząstki. W ten sposób otrzymaliśmy układ 2 równań różniczkowych pierwszego rzędu. Jak później zobaczymy, taka manipulacja jest użyteczna przy wprowadzeniu pojęcia przestrzeni fazowej dla równań różniczkowych. Jeżeli siła F nie zależy w sposób jawny od czasu, to układ równań

$$\frac{dx}{dt} = v \quad (1.7)$$

$$m \frac{dv}{dt} = F(x, v) \quad (1.8)$$

nazywamy autonomicznym. Innymi słowy, jest to autonomiczny układ 2 równań różniczkowych zwyczajnych 1-rzędu. Mówimy wówczas, że jego przestrzeń fazowa jest 2-wymiarowa.

Jeżeli cząstka porusza się na płaszczyźnie (X, Y) , to równanie Newtona ma postać:

$$m \frac{d^2 x}{dt^2} = F \left(x, y, \frac{dx}{dt}, \frac{dy}{dt}, t \right) \quad (1.9)$$

$$m \frac{d^2 y}{dt^2} = G \left(x, y, \frac{dx}{dt}, \frac{dy}{dt}, t \right) \quad (1.10)$$

gdzie F i G są składowymi siły działającymi na cząstkę w kierunku x oraz y . W ogólnym przypadku siły te zależą od położenia cząstki (x, y) , jej składowych prędkości $(dx/dt, dy/dt)$ oraz czasu t .

Jeżeli składowe siły F i G nie zależą w sposób jawny od czasu, to postępując podobnie jak poprzednio otrzymamy układ:

$$\frac{dx}{dt} = v \quad (1.11)$$

$$\frac{dy}{dt} = u \quad (1.12)$$

$$m \frac{dv}{dt} = F(x, y, v, u) \quad (1.13)$$

$$m \frac{du}{dt} = G(x, y, v, u) \quad (1.14)$$

Jest to autonomiczny układ 4 równań różniczkowych zwyczajnych 1-rzędu. Mówimy wówczas, że jego przestrzeń fazowa jest 4-wymiarowa.

Dla cząstki poruszającej się w przestrzeni (X, Y, Z) , mamy 3 równania Newtona 2-rzędu. Jeżeli 3 składowe siły nie zależą w sposób jawny od czasu, to postępując podobnie jak poprzednio otrzymamy układ 6 równań różniczkowych 1-rzędu i przestrzeń fazowa jest 6-wymiarowa. W ogólności, dla N cząstek poruszających się w przestrzeni, przestrzeń fazowa ma wymiar $6N$. Analiza takich równań przekracza możliwości współczesnej matematyki w tym sensie, że mało wiemy o ogólnych własnościach konkretnych układów, które modelujemy. To powoduje, że musimy stosować numeryczne metody i komputer jest nieodzownym narzędziem analizy.

Powyżej podaliśmy jeden przykład modelowania. Bazuje on na formalizmie Newtona i równaniach ruchu Newtona. Może być stosowany do opisu dynamiki cząstek klasycznych. Czasami wygodnie jest stosować inny formalizm jak na przykład formalizm Lagrange'a lub formalizm Hamiltona. W wielu przypadkach wszystkie trzy formalizmy są równoważne. Dla tzw. układów z więzami, wygodnie jest stosować formalizm Lagrange'a lub formalizm Hamiltona.

Definiując układ równań różniczkowych jako autonomiczny, zakładaliśmy że siła nie zależy w sposób jawny od czasu. Może wydawać się, że jest to jakieś ograniczenie. Nie jest to prawdą. Układy nieautonomiczne można sprowadzić do układów autonomicznych wprowadzając dodatkową zmienną niezależną, dodatkowe "położenie". Pokażemy to na prostym przykładzie. Rozpatrzmy cząstkę poruszającą się wzdłuż osi X . Na cząstkę działa siła tarcie proporcjonalna do prędkości cząstki, $F = -\gamma v$, działa siła potencjalna $F(x) = -V'(x)$ pochodząca od energii potencjalnej $V(x)$ (nazywanej skrótowo potencjałem). Siła ta jest ujemnym gradientem potencjału (czyli pochodną $V'(x)$). Dodatkowo na cząstkę działa periodyczna w czasie siła $F(t) = A \cos(\omega t)$. Równanie Newtona ma postać

$$m\ddot{x} = -\gamma\dot{x} - V'(x) + A \cos(\omega t) \quad (1.15)$$

gdzie kropki oznaczają pochodne względem czasu, a apostrof oznacza pochodną względem x . I tak

$$\dot{x} = \frac{dx}{dt}, \quad \ddot{x} = \frac{d^2x}{dt^2}, \quad V'(x) = \frac{dV(x)}{dx} \quad (1.16)$$

Równanie to możemy przedstawić w postaci układu 3 równań różniczkowych:

$$\dot{x} = v \quad (1.17)$$

$$m\dot{v} = -\gamma v - V'(x) + A\cos(z) \quad (1.18)$$

$$\dot{z} = \omega \quad (1.19)$$

Równoważność pokazujemy w następujący sposób:

w równaniu (1.18) należy zastąpić v z równania (1.17) wyrażeniem $v = \dot{x}$ pamiętając jednocześnie że $\dot{v} = \ddot{x}$ równanie (1.19) można scałkować i otrzymamy $z = \omega t$; wstawiamy to wyrażenie do równania (1.18). W ten sposób otrzymujemy znowu równanie (1.15). Tak więc jedno równanie różniczkowe nieautonomiczne 2-rzędu jest równoważne układowi 3 równań różniczkowych 1-rzędu. Odpowiadająca temu układowi przestrzeń fazowa jest 3-wymiarowa. Z przykładu tego płynie ważna wskazówka, jak otrzymywać autonomiczny układ równań różniczkowych 1-rzędu. Liczba tych równań definiuje przestrzeń fazową układu. Wymiar tej przestrzeni jest jedną z najważniejszych charakterystyk. Proszę to zapamiętać!

Fizyka stosuje też aparat równań różniczkowych cząstkowych. Studenci kierunku fizyka i pokrewnych kierunków znają przykłady takich równań. Równanie Schrödingera, równanie falowe, równanie dyfuzji, równania Maxwella to są równania różniczkowe cząstkowe. Ich analiza jest znacznie trudniejsza. Istnieją specjalne i specyficzne metody matematyczne pozwalające otrzymać informację o własnościach układów opisywanych takimi równaniami.

W wielu dziedzinach nauki (chemia, biologia, socjologia, nauki ekonomiczne) stosuje się fenomenologiczny sposób modelowania. Aby uzmysłwić, jak go stosować podamy jeden przykład.

1.1.1 Modelowanie procesu wzrostu

Procesy wzrostu pojawiają się na wielu obszarach. Nie trzeba być bystrym obserwatorem, aby zauważyć co wokół nas może wzrastać. My rozważamy jedną z możliwych klas procesów wzrostu: wzrost populacji zajęcy czy bakterii, wzrost depozytów pieniężnych na lokatach bankowych, wzrost stężenia substancji w reakcjach chemicznych czy wzrost liczby komórek nowotworowych. Często procesom wzrostu towarzyszą procesy malenia (zaniku, śmierci, ...). My je będziemy pomijać. Rozpatrzmy konkretny przykład: wzrost pieniędzy na lokacie bankowej. Załóżmy, że w chwili czasu t jest na lokacie $x(t)$ (np. złotych polskich). Pytamy, ile pieniędzy przyrośnie po pewnym czasie h , czyli ile pieniędzy będzie w chwili $t + h$. Zaczynamy modelować ten proces. Oznaczmy, że w chwili $t + h$ jest $x(t + h)$ pieniędzy. Na tę kwotę składają się pieniądze $x(t)$ oraz przyrost δ z odsetek, czyli

$$x(t + h) = x(t) + \delta \quad (1.20)$$

Przyrost δ zależy od wielkości oprocentowania k oraz od tego jak długo (h) trzymamy pieniądze na lokacie, czyli

$$\delta \propto x(t), \quad \delta \propto k, \quad \delta \propto h \quad (1.21)$$

Możemy to skomasować pisząc:

$$\delta = kx(t)h \quad (1.22)$$

Dlatego też

$$x(t+h) = x(t) + kx(t)h \quad (1.23)$$

Relacje tę możemy przepisać w postaci

$$\frac{x(t+h) - x(t)}{h} = kx(t) \quad (1.24)$$

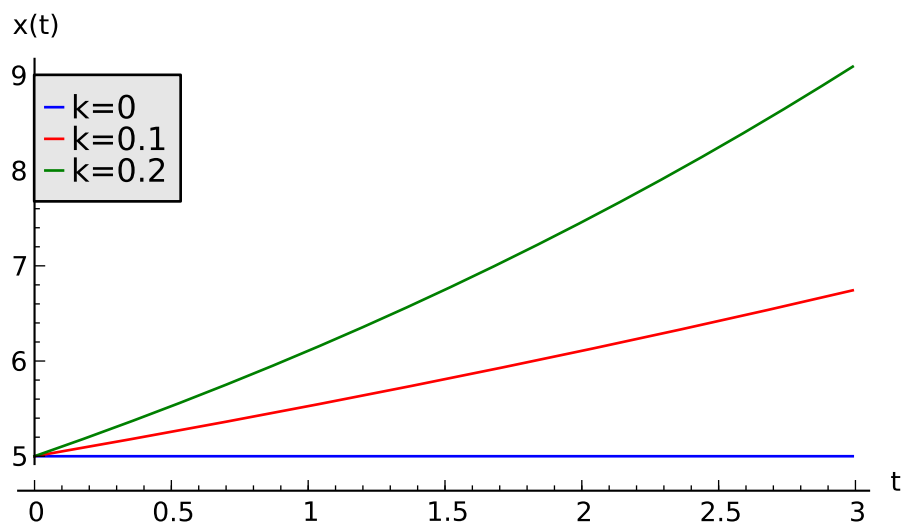
W granicy małych przyrostów czasu $h \rightarrow 0$, lewa strona jest definicją pochodnej

$$\frac{dx(t)}{dt} = kx(t), \quad x(0) = x_0 \quad (1.25)$$

gdzie x_0 jest wartością początkową naszej lokaty. W ten oto sposób otrzymaliśmy równanie opisujące dynamikę wzrostu pieniędzy na naszej lokacie bankowej. Jest to równanie różniczkowe zwyczajne, 1-go rzędu, autonomiczne. Jego przestrzeń fazowa jest 1-wymiarowa.

Poniżej pokazujemy rozwiązania tego równania dla 3 różnych wartości k .

```
var('N1,N2,N3')
T = srange(0,3,0.01)
# rozwiązania dla różnych wartości k=0, 0.1, 0.2
de = vector([0,0.1*N2,0.2*N3])
sol=desolve_odeint(de,[5,5,5],T,[N1,N2,N3])
```



Rysunek 1.1: Rozwiązania równania (1.25).

Inne procesy wzrostu także można modelować tym równaniem. Równanie to jest też punktem wyjściowym do modyfikacji, uogólnień, rozszerzeń, itp. Proste rozszerzenie polega na uzależnieniu współczynnika tempa wzrostu k od dodatkowych czynników. Na przykład przy modelowaniu wzrostu populacji zajęcy, możemy uzależnić tempo wzrostu k od liczby zajęcy w populacji: duża ilość zajęcy powoduje dużą konsumpcję pożywienia, a to z kolei skutkuje

zmaleniem ilości pożywienia i utrudnieniami w zdobywaniu pożywienia. W efekcie zmniejsza się tempo wzrostu k . Innymi słowy, k powinno być malejącą funkcją $x(t)$ liczebników w populacji. Istnieje nieskończenie wiele takich funkcji. Na przykład

$$k \rightarrow k(x) = \exp(-bx), \quad b > 0 \quad (1.26)$$

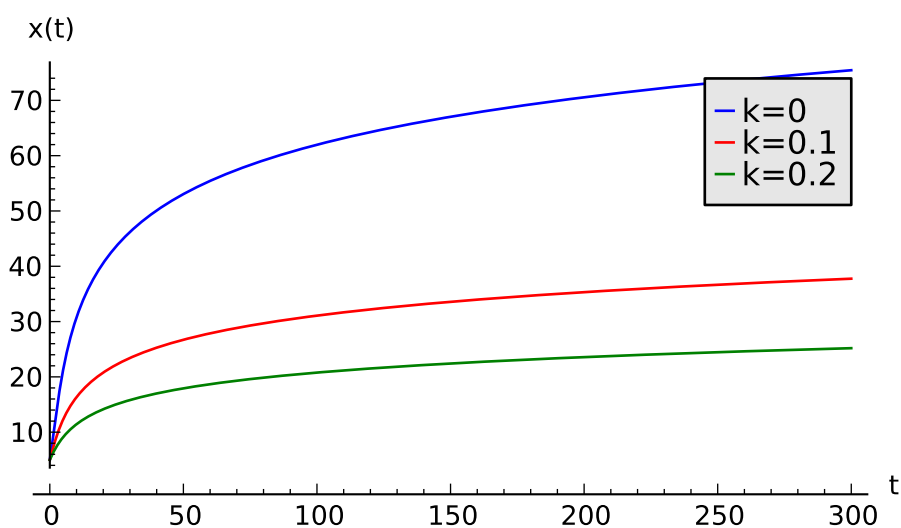
jest malejącą funkcją x . Teraz równanie różniczkowe ma postać

$$\frac{dx}{dt} = x e^{-bx}, \quad x = x(t), \quad x(0) = x_0 \quad (1.27)$$

Jakie są skutki takiej zmiany? Pokazujemy to na poniższym rysunku. Zauważamy, że tempo wzrostu populacji zmniejsza się w porównaniu z poprzednim przypadkiem.

Model można rozszerzyć uwzględniając procesy śmierci: te naturalne i te wskutek istnienia drapieżników, które zjadają osobników populacji. Prosty model ofiara-drapieżca jest 2-wymiarowy: opisuje zmiany w populacji ofiar i zmiany w populacji drapieżników. Jest to autonomiczny układ 2 równań różniczkowych zwyczajnych.

```
var('x,y,z')
U = srange(0, 300, 0.01)
de = vector([x*exp(-0.1*x), y*exp(-0.2*y), z*exp(-0.3*z)])
sol=desolve_odeint(de, [5,5,5], U, [x,y,z])
```



Rysunek 1.2: Rozwiązania równania (1.27).

1.2 Modelowanie z czasem dyskretnym

Powyżej otrzymaliśmy takie oto wyrażenie na przyrost:

$$x(t+h) = x(t) + khx(t) \quad (1.28)$$

Jeżeli zmiany następowałyby nie w sposób ciągły lecz dyskretny (np. co 1 dzień, co jedną godzinę) wówczas krok czasowy h jest dyskretny. Można wprowadzić oznaczenia

$$x_n = x(t), \quad x_{n+1} = x(t+h) \quad (1.29)$$

i wówczas równanie dla przyrostu ma postać

$$x_{n+1} = x_n + \alpha x_n, \quad \alpha = kh \quad (1.30)$$

W ten sposób otrzymujemy równanie z czasem dyskretnym. Ogólna postać tego typu równania to

$$x_{n+1} = f(x_n) \quad (1.31)$$

które mówi nam, jaką wartość przyjmuje dana wielkość w następnym kroku $n + 1$ jeżeli znana jest wartość tej wielkości w kroku n . Równanie to nazywa się też równaniem rekurencyjnym. W zależności od postaci funkcji $f(x)$ otrzymujemy różne modele dynamiki układów.

Układ 2 równań z czasem dyskretnym ma postać

$$x_{n+1} = f(x_n, y_n) \quad (1.32)$$

$$y_{n+1} = g(x_n, y_n) \quad (1.33)$$

Analiza jakościowa takiego układu jest bardzo trudna. Czasami nieumiejętne stosowanie numerycznej analizy może skutkować tym, że umkną nam istotne cechy takiego układu, zwłaszcza gdy w układzie występują dodatkowo parametry których zmiana może powodować coś, co nazywa się bifurkacjami. Ale o tym w dalszej części książki.

1.3 Istnienie i jednoznaczność rozwiązań

Do opisu realnych zjawisk przy pomocy równań różniczkowych zwyczajnych z warunkami początkowymi zadanymi w chwili czasu $t = 0$, potrzebne nam są rozwiązania dla czasów $t > 0$ (ewolucja czasowa). Można też rozpatrywać przypadek $t < 0$ ale to zaliczyłbym do ćwiczeń matematycznych. Ważnym zagadnieniem jest istnienie rozwiązań równań różniczkowych. Możemy zapytać, czy zawsze rozwiązania równań różniczkowych istnieją i jeżeli istnieją, to czy to są jedyne rozwiązania z warunkiem początkowym. Oczywiście dla różnych warunków początkowych układ może różnie ewoluować, ale gdy startuje zawsze z tego samego stanu (warunku) początkowego to czy ewolucja jest taka sama? Na tym polega problem jednoznaczności rozwiązań. Jeżeli dla danego warunku początkowego istnieją np. 3 rozwiązania, to jak ewoluuje układ: istnieją 3 możliwości i którą możliwość wybiera układ? Gdyby tak było dla realnych układów to nie moglibyśmy przewidywać ewolucji układu, nie moglibyśmy sterować układami, brak byłoby determinizmu. W rozwoju nauk ścisłych to właśnie determinizm stał się kołem napędowym rozwoju cywilizacyjnego ludzkości. To determinizm pozwala budować urządzenia, które działają tak jak my sobie tego życzymy: telewizor odbiera wybrany przeze mnie program, używam telefonu do komunikacji z moją rodziną, wystrzelona rakietą ma taką trajektorię jaką zaplanowałem, itd. Zbadamy 3 przykłady, które przybliżą nam powyższą problematykę. Źródło tych przykładów jest w książce: J. Hale, H. Kocak, “Dynamics and Bifurcations”. Książka jest znakomita.

1.3.1 Przykład 1

Równanie

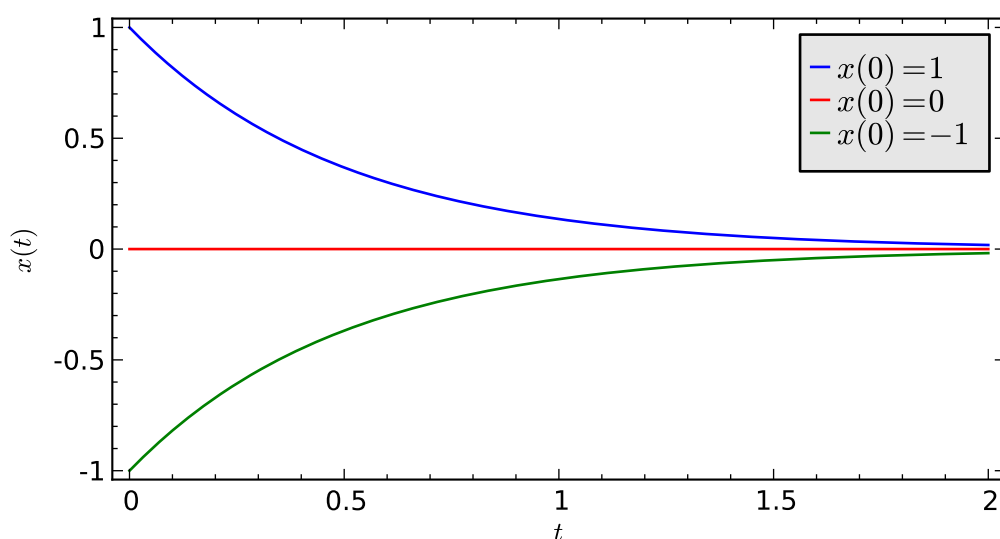
$$\frac{dx}{dt} = -2x, \quad x(0) = x_0 \quad (1.34)$$

jest równaniem różniczkowym liniowym. Jest to jedno z najprostszych równań różniczkowych. Można łatwo sprawdzić, że funkcja

$$x(t) = x_0 e^{-2t} \quad (1.35)$$

jest rozwiązaniem i spełnia warunek początkowy $x(0) = x_0$. Funkcja ta jest dobrze określona dla wszystkich skończonych wartości czasu $t \in (-\infty, \infty)$. Nie ma tu większych ograniczeń. Jest to jedyne rozwiązanie. Poniższy rysunek daje wyobrażenie o rozwiązaniach $x(t)$ dla 3 różnych warunków początkowych. Przy okazji zauważmy, że wszystkie trzy rozwiązania dążą do tego samego stanu $x = 0$ dla długich czasów $t \rightarrow \infty$.

```
var('t')
g(t,a)=a*exp(-2*t)
p=plot(g(t,a=1), (t,0,2), legend_label=r"$x(0)=1$", color='blue')
p+=plot(g(t,a=0), (t,0,2), legend_label=r"$x(0)=0$", color='red')
p+=plot(g(t,a=-1), (t,0,2), legend_label=r"$x(0)=-1$", color='green')
show(p)
```



Rysunek 1.3: Rozwiązania (1.34) liniowego równania różniczkowego (1.34).

1.3.2 Przykład 2

Równanie

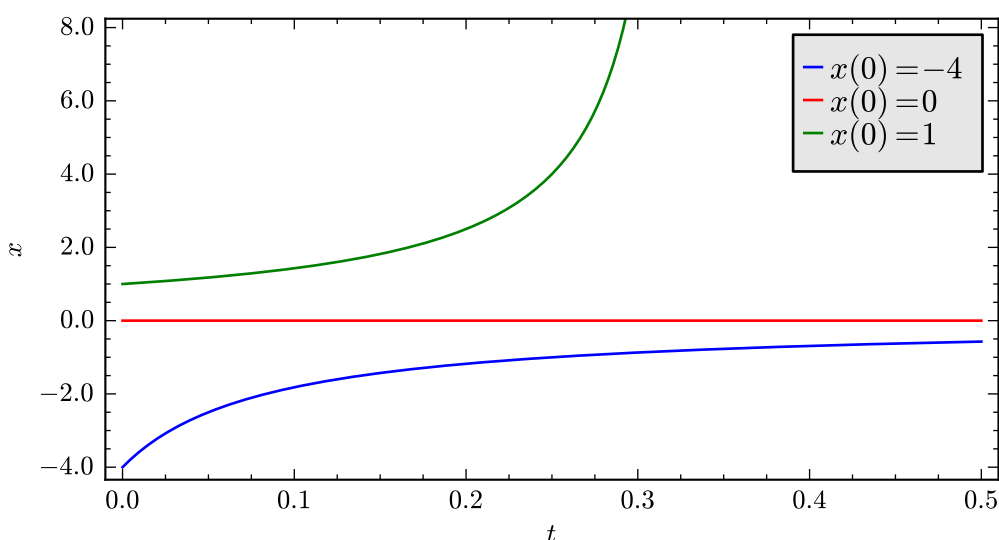
$$\frac{dx}{dt} = 3x^2, \quad x(0) = x_0 \quad (1.36)$$

jest równaniem różniczkowym nieliniowym. Prawa strona tego równania jest określona dla wszystkich wartości x . Podobnie jak poprzednie równanie, można je rozwiązać metodą separacji zmiennych. Otrzymamy funkcję

$$x(t) = \frac{x_0}{1 - 3x_0 t} \quad (1.37)$$

która jest rozwiązaniem i spełnia warunek początkowy. Funkcja ta nie jest określona dla wszystkich skończonych wartości czasu $t \in (-\infty, \infty)$. Istnieją ograniczenia dla wartości czasu t . Ale jest to jedyne rozwiązanie.

```
var('t')
g=plot(-4.0/(1+12*t), (t,0,0.5), legend_label=r'$x(0)=-4$', color='blue')
g+=plot(lambda t: 0.0, (t,0,0.5), legend_label=r'$x(0)=0$', color='red')
g+=plot(1.0/(1-3*t), (t,0,1/3), legend_label=r'$x(0)=1$', color='green')
g.show()
```



Rysunek 1.4: Rozwiązania (1.37) równania różniczkowego (1.36).

Wszystkie rozwiązania z ujemnym warunkiem początkowym $x(0) < 0$ są dobrze zdefiniowane dla wszystkich czasów $t > 0$ (krzywa niebieska). Podobnie jest z rozwiązaniem $x(t) = 0$ dla warunku początkowego $x(0) = 0$ (krzywa czerwona). Natomiast rozwiązanie z dodatnim warunkiem początkowym $x(0) > 0$ rozbiega się w skończonym czasie $t < 1/3x_0$. Gdyby to równanie miało opisywać ruch cząstki, to oznacza że w skończonym czasie cząstka przebywa nieskończoną odległość. To jest нефизyczne. Równanie to mogłoby opisywać proces wybuchu substancji: x mogłoby być objętością pęczniejącej substancji która wybuchu po skończonym czasie.

1.3.3 Przykład 3

Równanie

$$\frac{dx}{dt} = 2\sqrt{x}, \quad x(0) = x_0 \geq 0 \quad (1.38)$$

jest równaniem różniczkowym nieliniowym. Prawa strona tego równania jest określona dla nieujemnych wartości $x \geq 0$. Podobnie jak 2 poprzednie równania, można je rozwiązać metodą separacji zmiennych. Otrzymamy rozwiązanie

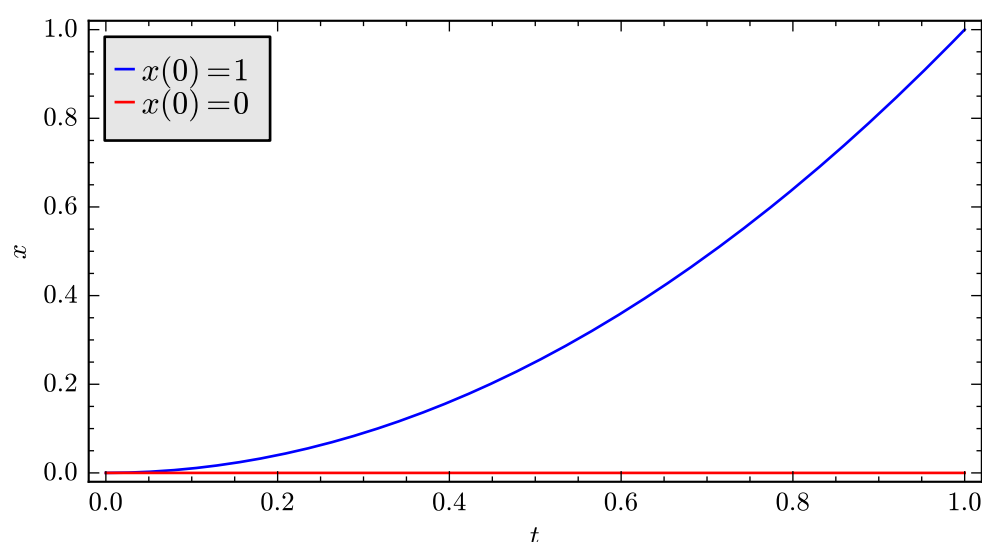
$$x(t) = (t + \sqrt{x_0})^2 \quad (1.39)$$

Funkcja ta jest określona dla wszystkich wartości czasu $t > 0$. Jest to jedyne rozwiązanie z wyjątkiem jednego warunku początkowego: $x(0) = 0$. Dla tego warunku początkowego istnieje jeszcze jedno rozwiązanie, a mianowicie $x(t) = 0$. Tak więc dla $x(0) = 0$ mamy 2 różne rozwiązania

$$x(t) = t^2, \quad x(t) = 0 \quad (1.40)$$

Jak przebiega ewolucja, gdy układ startuje ze stanu początkowego $x(0) = 0$? W tym przypadku rozwiązania są niejednoznaczne.

```
var('t')
p = plot(t**2, (t, 0, 1), legend_label=r"$x(0)=1$", color='blue')
p+=plot(0, (t, 0, 1), legend_label=r"$x(0)=0$", color='red')
p.show()
```



Rysunek 1.5: Rozwiązania równania (1.38).

Co jest takiego charakterystycznego w ostatnim przykładzie, że pojawia się niejednoznaczność rozwiązania równania różniczkowego? Na to pytanie daje odpowiedź twierdzenie o jednoznaczności rozwiązania równania różniczkowego. Potrzebna nam będzie własność funkcji:

Mówimy, że funkcja $f(x)$ spełnia warunek Lipschitza na zbiorze otwartym U jeżeli istnieje taka stała $L > 0$, że

$$|f(x_2) - f(x_1)| \leq L|x_2 - x_1| \quad (1.41)$$

dla wszystkich $x_1, x_2 \in U$.

Warunek Lipschitza można zapisać w postaci

$$|f(x+h) - f(x)| \leq Lh \quad \text{lub jako} \quad \left| \frac{f(x+h) - f(x)}{h} \right| \leq L \quad (1.42)$$

Z tego wynika że jeżeli $f(x)$ ma ograniczoną pochodną, to spełnia warunek Lipschitza. Są oczywiście nieróżniczkowalne funkcje, które spełniają warunek Lipschitza.

Twierdzenie Picarda Jeżeli funkcja $f(x)$ jest ciągła w U oraz spełnia warunek Lipschitza w U wówczas równanie różniczkowe

$$\frac{dx}{dt} = f(x), \quad x(0) = x_0 \quad (1.43)$$

ma dokładnie jedno rozwiązanie w U .

Istnieje kilka modyfikacji tego twierdzenia, ale na nasze potrzeby ta najprostsza wersja jest wystarczająca.

Teraz możemy odpowiedzieć, dlaczego w 3 przykładzie rozwiązanie jest niejednoznaczne: funkcja $f(x) = 2\sqrt{x}$ nie spełnia warunku Lipschitza ponieważ pochodna

$$\frac{df(x)}{dx} = \frac{1}{\sqrt{x}} \quad (1.44)$$

w punkcie $x = 0$ jest rozbieżna. W punktach $x > 0$ pochodna ma wartość skończoną i jest spełnione twierdzenie Picarda. Dlatego też rozwiązania są jednoznaczne.

1.3.4 Dodatek

Sage z powodzeniem jest w stanie rozwiązywać pewne równania różniczkowe zwyczajne. Zobaczmy jak poradzi sobie z powyższymi przykładami.

Przykład 1

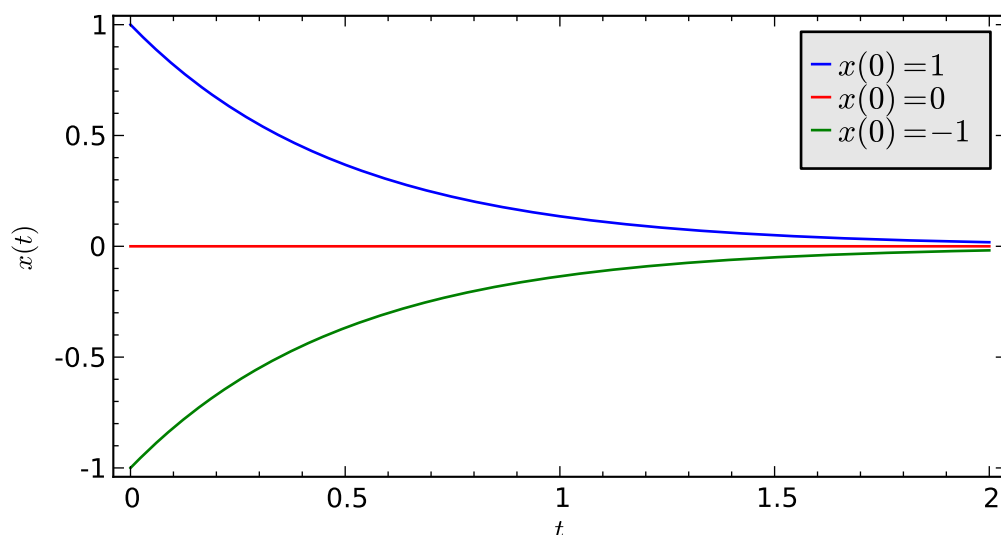
$$\frac{dx}{dt} = -2x, \quad x(0) = x_0 \quad (1.45)$$

z rozwiązaniem

$$x(t) = x_0 e^{-2t}. \quad (1.46)$$

Na początek zadamy sobie zmienne. Druga linijka mówi o tym, że zmienna x będzie funkcją parametru t (czasu). Zamiast używać nazwy g użyjemy świeżo obliczonego rozwiązania `rozv`.

```
var('t x_0')
x = function('x', t)
rrz = diff(x,t) == -2*x
rozv = desolve(rrz, x)
rozv = rozv.subs(c=x_0)
p=plot(rozv(x_0=1), (t,0,2), legend_label=r"$x(0)=1$", color='blue')
p+=plot(rozv(x_0=0), (t,0,2), legend_label=r"$x(0)=0$", color='green')
p+=plot(rozv(x_0=-1), (t,0,2), legend_label=r"$x(0)=-1$", color='red')
show(p)
```



Rysunek 1.6: Rozwiązania równania (1.45).

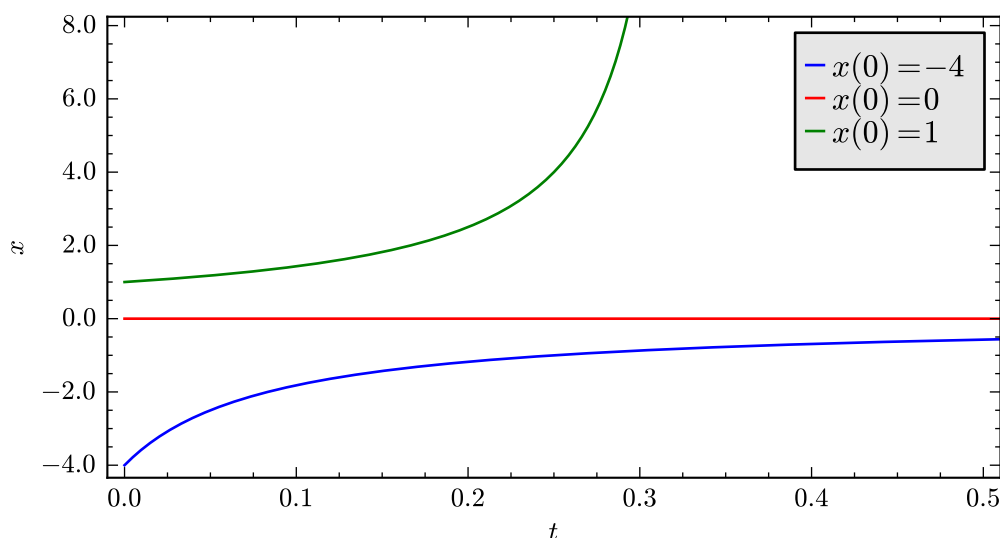
Przykład 2

$$\frac{dx}{dt} = 3x^2, \quad x(0) = x_0 \quad (1.47)$$

z rozwiązaniem

$$x(t) = \frac{x_0}{1 - 3x_0 t}. \quad (1.48)$$

```
var('t x_0 c')
x = function('x', t)
#Definiujemy równanie różniczkowe
rrz = diff(x,t) == 3*x^2
rozw2 = desolve(rrz, x)
#i je rozwiązujemy...
#krok 1\n obliczamy x(t) z poprzedniego kroku
rozw2 = solve(rozw2,x)[0].rhs()
#krok 2\n obliczamy x(0)
buf = rozw2(t=0) == x_0
show(buf)
#krok 3\n wyznaczamy stałą c
buf = solve(buf,c)[0].rhs()
#krok 4\n wstawiamy c do równania
rozw2 = rozw2.subs(c=buf).full_simplify()
#wyniki
w=plot(rozw2(x_0=-4), (t,0,1))
w+=plot(rozw2(x_0=0), (t,0,1))
w+=plot(rozw2(x_0=1), (t,0,1/3))
```



Rysunek 1.7: Rozwiązania równania (1.47).

Przykład 2

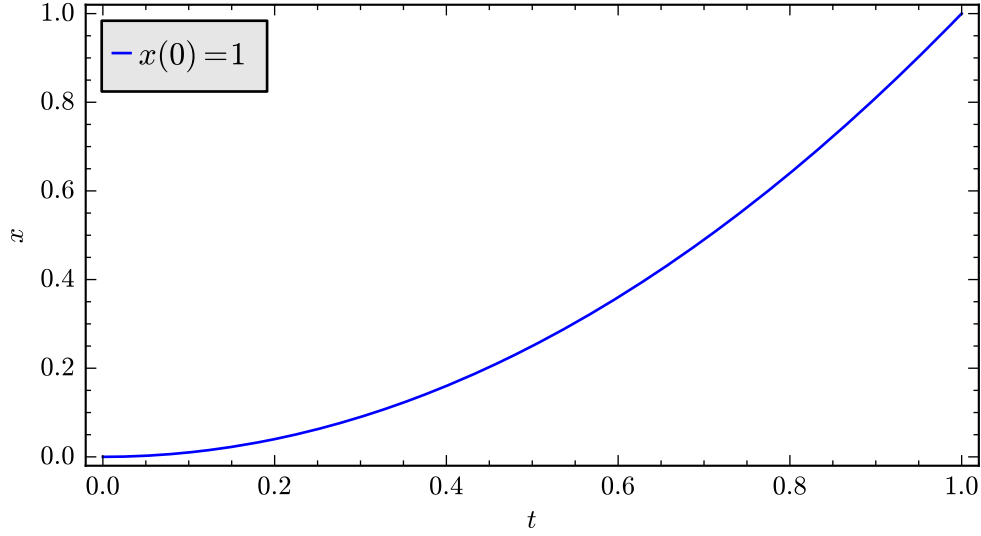
$$\frac{dx}{dt} = 2\sqrt{x}, \quad x(0) = x_0 \geq 0 \quad (1.49)$$

z rozwiązaniem

$$x(t) = (t + \sqrt{x_0})^2 \quad (1.50)$$

```
var('t x_0 c')
forget()
assume(x_0>=0)
assume(t+c>0)
#równanie
x = function('x', t)
rrz = diff(x,t) == 2*sqrt(x)
#i jego rozwiązanie
rozv3 = solve(desolve(rrz, x), x)[0]
#stała całkowania
buf = solve(x_0 == rozv3.rhs()(t=0), c)
#mamy dwa możliwe rozwiązania, wybieramy to z dodatnim c
buf = buf[1]
#i dostajemy ostatecznie
rozv3 = rozv3.subs(c=buf.rhs())
#wyniki
p1=plot(rozv3.rhs()(x_0=0), (t, 0, 1))
```

No tak, ale gdzie jest rozwiązanie $x(t) = 0$? Na chwilę obecną Sage nie rozróżni obu możliwych rozwiązań. Dlatego umiejętność analitycznego rozwiązania takich problemów wciąż jest niezbędna!



Rysunek 1.8: Rozwiązania równania (1.49).

1.4 Układy dynamiczne z czasem ciągłym

We Wstępie podaliśmy kilka przykładów układów opisanych za pomocą równań różniczkowych zwyczajnych. Pierwsza klasa układów to układy opisywane przez mechanikę klasyczną i jej równania Newtona. Inna klasa układów to równania fenomenologiczne opisujące procesy wzrostu, procesy kinetyki chemicznej, dynamiki populacyjnej w układach biologicznych, itd. Te dwie klasy układów opisywane są układem równań różniczkowych zwyczajnych zapisanych w ogólnej postaci jako układ

$$\begin{aligned}\frac{dx_1}{dt} &= F_1(x_1, x_2, x_3, \dots, x_n) \\ \frac{dx_2}{dt} &= F_2(x_1, x_2, x_3, \dots, x_n) \\ &\vdots \\ \frac{dx_n}{dt} &= F_n(x_1, x_2, x_3, \dots, x_n)\end{aligned}\tag{1.51}$$

Ten układ możemy zapisać w notacji wektorowej w postaci

$$\frac{d\vec{x}}{dt} = \vec{F}(\vec{x}), \quad \vec{x}(0) = \vec{x}_0\tag{1.52}$$

gdzie wektory

$$\vec{x} = [x_1, x_2, x_3, \dots, x_n], \quad \vec{F} = [F_1, F_2, F_3, \dots, F_n]\tag{1.53}$$

oraz dany jest zbiór warunków początkowych

$$\vec{x}(0) = [x_1(0), x_2(0), x_3(0), \dots, x_n(0)] = \vec{x}_0 = [x_1^{(0)}, x_2^{(0)}, x_3^{(0)}, \dots, x_n^{(0)}]\tag{1.54}$$

Wskaźnik n mówi, ile równań różniczkowych jest “ukrytych” w powyższym zapisie wektorowym. Innymi słowy, rozważamy układ n równań różniczkowych scharakteryzowanych przez n

funkcji skalarnych $F_i(x_1, x_2, x_3, \dots, x_n)$, ($i = 1, 2, 3, \dots, n$). Zauważmy, że rozważamy funkcje F_i które nie zależą w sposób jawny od czasu. W takim przypadku mówimy, że jest to układ autonomiczny n równań różniczkowych zwyczajnych. Ponadto wektor \vec{F} można traktować jako pole wektorowe stowarzyszone z układem równań różniczkowych lub pole wektorowe generowane przez układ równań różniczkowych. Do tej kwestii powrócimy jeszcze.

1.4.1 Przestrzeń fazowa

Jeszcze raz przepisujemy równania różniczkowe (1.52) w notacji:

$$\begin{aligned}\frac{dx_1}{dt} &= F_1(x_1, x_2, x_3, \dots, x_n) \\ \frac{dx_2}{dt} &= F_2(x_1, x_2, x_3, \dots, x_n) \\ &\vdots \\ \frac{dx_n}{dt} &= F_n(x_1, x_2, x_3, \dots, x_n)\end{aligned}\tag{1.55}$$

Powyższy układ równań różniczkowych definiuje pewien układ dynamiczny (matematyczna definicja układu dynamicznego może być bardzo abstrakcyjna, ale na nasze potrzeby wystarczy to, co napisaliśmy). Zbiór wszystkich możliwych wartości $\{x_1, x_2, x_3, \dots, x_n\}$ tworzy zbiór który nazywamy przestrzenią fazową układu (??). Wymiar tej przestrzeni wynosi n , czyli tyle ile jest równań różniczkowych.

W zależności od kontekstu, będziemy stosowali różne zapisy tych samych równań.

Przykłady:

1. Jedno równanie różniczkowe. Zwykle będziemy stosowali zapis

$$\frac{dx}{dt} = \dot{x} = f(x)\tag{1.56}$$

Przestrzeń fazowa jest 1-wymiarowa.

2. Dwa równania różniczkowe. Zwykle będziemy stosowali zapis

$$\begin{aligned}\frac{dx}{dt} &= \dot{x} = f(x, y) \\ \frac{dy}{dt} &= \dot{y} = g(x, y)\end{aligned}\tag{1.57}$$

Przestrzeń fazowa jest 2-wymiarowa.

3. Trzy równania różniczkowe. Zwykle będziemy stosowali zapis

$$\begin{aligned}\frac{dx}{dt} &= \dot{x} = f(x, y, z) \\ \frac{dy}{dt} &= \dot{y} = g(x, y, z) \\ \frac{dz}{dt} &= \dot{z} = h(x, y, z)\end{aligned}\tag{1.58}$$

Przestrzeń fazowa jest 3-wymiarowa.

4. Równanie Newtona dla cząstki poruszającej się tylko wzdłuż osi OX na którą działa siła $F(x)$ zależna tylko od położenia ma postać

$$m\ddot{x} = F(x)\tag{1.59}$$

gdzie m jest masą cząstki. Jest to równanie różniczkowe 2-go rzędu i jest ono równoważne układowi 2 równań różniczkowych 1-go rzędu:

$$\begin{aligned}\dot{x} &= v \\ \dot{v} &= \frac{1}{m}F(x)\end{aligned}\tag{1.60}$$

Przestrzeń fazowa jest 2-wymiarowa: jest to zbiór możliwych położeń i prędkości cząstki, $\{x, v\}$. Mimo swej prostoty, ten model jest niesłychanie ważny. Stanowi on punkt wyjścia dla zrozumienia wielu ważnych aspektów układów dynamicznych.

Geometryczne własności przestrzeni fazowej

Krzywa fazowe

Aby uniknąć na tym etapie abstrakcyjnych definicji, będziemy rozważać dla przykładu 2-wymiarowy układ dynamiczny

$$\begin{aligned}\dot{x} &= f(x, y), & x(0) &= x_0 \\ \dot{y} &= g(x, y), & y(0) &= y_0\end{aligned}\tag{1.61}$$

Przestrzeń fazowa jest 2-wymiarowa. Może to być płaszczyzna lub jej część. Ale może to być bardziej skomplikowany zbiór 2-wymiarowy. Na przykład może to być sfera (podobna do powierzchni piłki), może to być torus (podobny do dętki rowerowej). Mogą to być jeszcze bardziej skomplikowane obiekty 2-wymiarowe. Ale dla naszych celów wystarczy rozważać płaszczyznę. Na płaszczyźnie można estetycznie przedstawiać coś w formie rysunków. Wprowadzamy na płaszczyźnie kartezjański układ współrzędnych o osiach OX i OY . Warunek początkowy $\{x_0 = x(0), y_0 = y(0)\}$ jest punktem o odpowiednich współrzędnych. Rozwiązujemy powyższy układ równań różniczkowych numerycznie przy pomocy najprostszego schematu:

$$\begin{aligned}\frac{x(t+h) - x(t)}{h} &= f(x(t), y(t)) \\ \frac{y(t+h) - y(t)}{h} &= g(x(t), y(t))\end{aligned}\tag{1.62}$$

Przepiszemy to w postaci

$$\begin{aligned}x(t+h) &= x(t) + f(x(t), y(t))h \\ y(t+h) &= y(t) + g(x(t), y(t))h\end{aligned}\tag{1.63}$$

1. Obliczenia numeryczne musimy zacząć od warunku początkowego w chwili $t = 0$, czyli w pierwszym kroku obliczamy

$$\begin{aligned}x_1 &= x(h) = x(0) + f(x(0), y(0))h \\ y_1 &= y(h) = y(0) + g(x(0), y(0))h\end{aligned}\tag{1.64}$$

Na płaszczyźnie otrzymujemy punkt o współrzędnych $\{x_1, y_1\}$. Zaznaczmy go na płaszczyźnie. Teraz mamy już 2 punkty:

$$\{x_0, y_0\}, \quad \{x_1, y_1\}\tag{1.65}$$

2. W następnym kroku wybieramy czas $t = h$:

$$\begin{aligned}x_2 &= x(h+h) = x(2h) = x(h) + f(x(h), y(h))h \\ y_2 &= y(h+h) = y(2h) = y(h) + g(x(h), y(h))h\end{aligned}\tag{1.66}$$

Wykorzystamy oznaczenie jak wyżej: $x_1 = x(h)$, $y_1 = y(h)$ i przepiszemy te równania w postaci

$$\begin{aligned}x_2 &= x_1 + f(x_1, y_1)h \\ y_2 &= y_1 + g(x_1, y_1)h\end{aligned}\tag{1.67}$$

Na płaszczyźnie otrzymujemy punkt o współrzędnych $\{x_2, y_2\}$. Zaznaczmy go na płaszczyźnie. Teraz mamy już 3 punkty:

$$\{x_0, y_0\}, \quad \{x_1, y_1\}, \quad \{x_2, y_2\}\tag{1.68}$$

3. Widzimy od razu, że w 3 kroku otrzymujemy równania

$$\begin{aligned}x_3 &= x_2 + f(x_2, y_2)h \\ y_3 &= y_2 + g(x_2, y_2)h\end{aligned}\tag{1.69}$$

i otrzymujemy punkt o współrzędnych $\{x_3, y_3\}$.

4. Zauważamy, że w n -tym kroku otrzymujemy równania

$$\begin{aligned}x_n &= x_{n-1} + f(x_{n-1}, y_{n-1})h \\ y_n &= y_{n-1} + g(x_{n-1}, y_{n-1})h\end{aligned}\tag{1.70}$$

5. Częściej pisze się, co się otrzymuje w następnym kroku, czyli $n+1$:

$$\begin{aligned}x_{n+1} &= x_n + f(x_n, y_n)h \\ y_{n+1} &= y_n + g(x_n, y_n)h\end{aligned}\tag{1.71}$$

Otrzymujemy równania rekurencyjne, które pozwalają wyznaczyć ewolucję układu, czyli rozwiązanie numeryczne układu równań różniczkowych. Na płaszczyźnie XY otrzymujemy ciąg punktów o współrzędnych

$$\{x_n, y_n\}\tag{1.72}$$

Jeżeli przyrost czasu h jest odpowiednio mały, to ten ciąg punktów łączymy linią ciągłą i otrzymujemy krzywą na płaszczyźnie. Ta krzywa nazywa się krzywą fazową układu dynamicznego. Mając narysowaną taką krzywą fazową, możemy wnioskować o ewolucji układu i cechach charakterystycznych zachowania się układu w czasie t . Poniżej przedstawiamy dwa przykłady: krzywe fazowe dla oscylatora harmonicznego i oscylatora harmonicznego tłumionego.

Oscylator harmoniczny Przykładem oscylatora harmonicznego jest ciało o masie m przyłączone do sprężyny i poruszające się wzdłuż osi OX . Siła działająca na to ciało jest proporcjonalna do wychylenia x od położenia równowagi i przeciwnie skierowana do wychylenia; gdy rozciągamy sprężynę w kierunku większych dodatnich wartości x to siła działa w kierunku ujemnych wartości x :

$$F = -kx\tag{1.73}$$

gdzie k charakteryzuje “sprężystość” sprężyny. Równanie Newtona ma postać:

$$m\ddot{x} = -kx, \quad \text{lub w postaci} \quad \ddot{x} = -(k/m)x = -\omega^2 x\tag{1.74}$$

gdzie $\omega^2 = k/m$. To równanie drugiego rzędu jest równoważne 2 równaniom pierwszego rzędu:

$$\begin{aligned}\dot{x} &= y, & x(0) &= x_0 \\ \dot{y} &= -\omega^2 x, & y(0) &= y_0\end{aligned}\tag{1.75}$$

Tłumiony oscylator harmoniczny Jeżeli w poprzednim przykładzie założymy bardziej realistyczną sytuację, w której układ nie jest w próżni, ale znajduje się w środowisku (np. w powietrzu, w wodzie lub innej cieczy), to na ciało działa dodatkowa siła, a mianowicie siła tarcia (tłumienia) F_d . Siła tarcia jest proporcjonalna do prędkości cząstki i przeciwnie skierowana do kierunku ruchu

$$F_d = -\gamma_0 v = -\gamma_0 \dot{x}\tag{1.76}$$

gdzie γ_0 nazywa się współczynnikiem tarcia (tłumienia).

Siła tarcia jest związana z oddziaływaniem ciała z cząsteczkami otoczenia. Otoczenie stawia opór gdy ciało porusza się w nim i im większa jest prędkość ciała tym większy jest opór otoczenia. Doświadczamy to, gdy biegniemy cali zanurzeni w wodzie.

Uwzględniając siłę tarcia, równanie Newtona przyjmuje postać

$$m\ddot{x} = -kx - \gamma_0\dot{x} \quad (1.77)$$

lub w postaci

$$\ddot{x} = -\frac{k}{m}x - \frac{\gamma_0}{m}\dot{x} = -\omega^2x - \gamma\dot{x} \quad (1.78)$$

gdzie $\omega^2 = k/m$ oraz $\gamma = \gamma_0/m$. Równanie powyższe jest równoważne 2 równaniom pierwszego rzędu:

$$\begin{aligned} \dot{x} &= y, & x(0) &= x_0 \\ \dot{y} &= -\gamma y - \omega^2 x, & y(0) &= y_0 \end{aligned} \quad (1.79)$$

Oczywiście gdy $\gamma = 0$, wówczas otrzymujemy równanie oscylatora harmonicznego bez tarcia (nie tłumionego).

Poniżej przedstawiamy krzywe fazowe dla tych 2 przykładów.

Oscylator harmoniczny bez tarcia Przede wszystkim napiszemy funkcję realizującą schemat Eulera rozwiązywania (iterowania) równań różniczkowych zupełnie tak jak przedstawiono powyżej.

```
def schemat_eulera2D(vec, ics, Tlist):
    i = 0
    dx = [ics[0]]
    dy = [ics[1]]
    h = Tlist[i+1] - Tlist[i]

    iks(x,y) = vec[0R]*h
    igrek(x,y) = vec[1R]*h

    for time in Tlist:
        dx.append(dx[i] + iks(dx[i],dy[i]))
        dy.append(dy[i] + igrek(dx[i],dy[i]))
        i += 1

    return zip(dx,dy)
```

Jak widać funkcja ta nie jest ani zbyt trudna do zrozumienia, ani przesadnie optymalnie napisana. Jako argumenty wejściowe pobierzemy wektor prawych stron układu dwóch równań różniczkowych (np: oscylatora harmonicznego), wartości początkowe obu zmiennych oraz listę punktów czasowych w jakich dynamika ma być próbkowana. Aby ów algorytm zadziałał, odległości pomiędzy poszczególnymi chwilami czasowymi muszą być stałe. Tak zdefiniowana funkcja pozwoli nam na praktyczne takie samo wywołanie jak oryginalna funkcja Sage `desolve_odeint`.

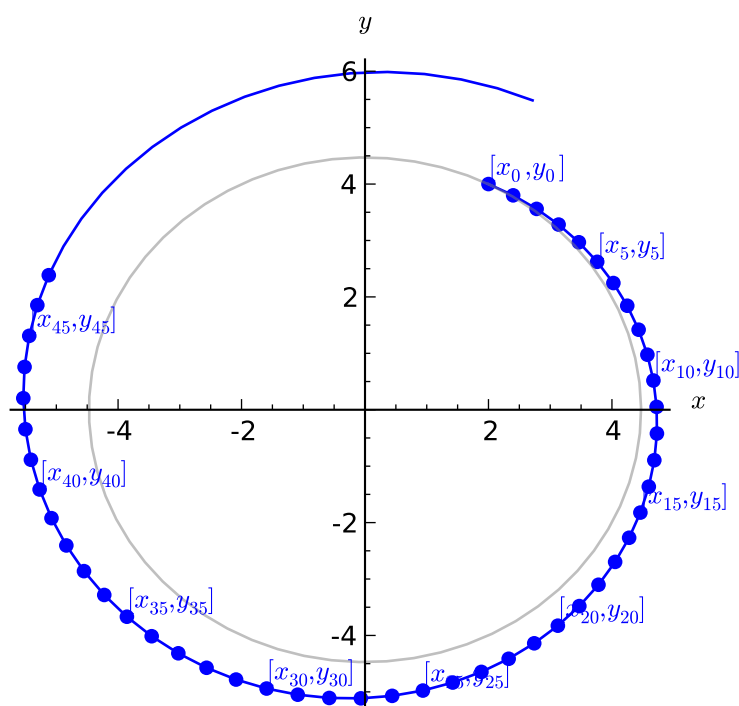
Następnie wykonamy symulacje dla konkretnego układu, tzn: dla ustalonych parametrów ω , x_0 , y_0 .

```

omega = 1
x0, y0 = 2, 4
h = 0.1
T = 2*pi/omega
l1T = srange(0,T,h,include_endpoint=True)
l2T = srange(0,T,h,include_endpoint=True)
de = [y,-omega^2*x]
bgr = desolve_odeint(vector(de),[x0,y0],l2T,[x,y])
oscylator = schemat_eulera2D(de,[x0,y0],l1T)

```

Teraz pozostaje nam już tylko wykreślić odpowiednie listy. Formę wizualizacji pozostawiamy czytelnikom, poniżej prezentując wyniki obliczeń. Krzywa zaprezentowana w tle to wynik użycia zaawansowanych algorytmów adaptatywnych. Krzywa bardziej widoczna to efekt obliczeń prostym schematem Eulera, który to pozostawia wiele do życzenia...



Rysunek 1.9: Numeryczne rozwiązania oscylatora harmonicznego.

W przypadku oscylatora nietłumionego, krzywe fazowe są zamknięte. Częstka z biegiem czasu porusza się tak, że położenie x i prędkość $v = y$ leżą na krzywej fazowej. Ponieważ jest to krzywa zamknięta, to po pewnym czasie cząstka znowu “przebiega” punkty, które się powtarzają. Powtarzanie się jest cechą charakterystyczną ruchu okresowego. Tak więc krzywa fazowa zamknięta przedstawia ruch okresowy (periodyczny). Okres takiego ruchu periodycznego to czas potrzebny na to, aby cząstka startując od punktu np. $\{x_0, y_0\}$ i poruszając się po krzywej fazowej dotarła znowu do tego samego punktu $\{x_0, y_0\}$.

W przypadku oscylatora tłumionego, krzywą fazową jest spirala zwijająca się do punktu $\{0, 0\}$. Ruch po spirali oznacza, że zarówno x jak i $v = y$ maleją w czasie i dla długich czasów położenie x oraz prędkość v dążą do zera, czyli cząstka zwalnia i na końcu zatrzymuje się. To jest ruch tłumiony: amplituda drgań maleje w czasie. To jest to, co obserwujemy w ruchu

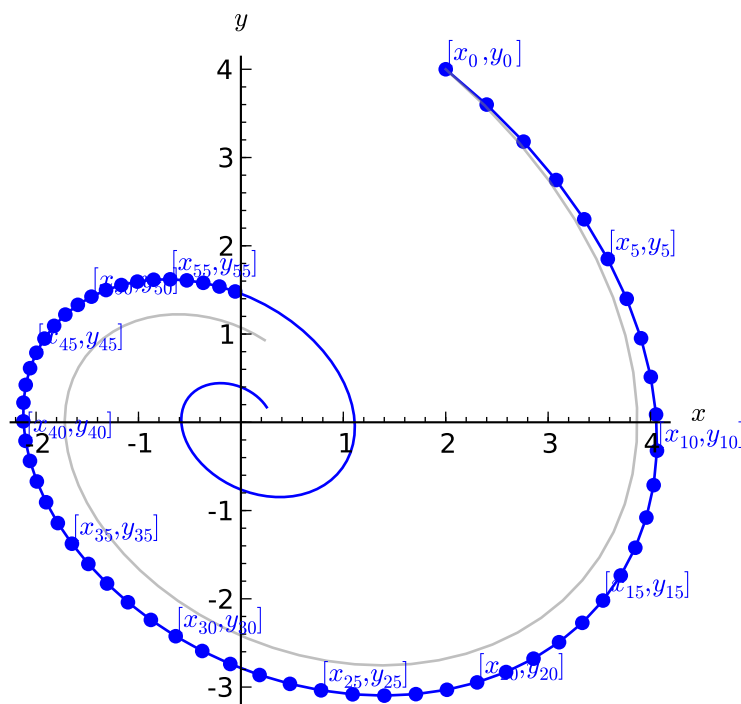
kulki zawieszonych na nitce: kulka wykonuje coraz to mniejsze drgania i po długim czasie wisi pionowo (to jest coś co nazywa się stanem równowagi lub położeniem stacjonarnym).

Gdy mamy bardziej skomplikowane krzywe fazowe, ich “rozszyfrowanie” może być trudniejsze. Ale ogólna zasada jest taka: gdy x rośnie to oznacza wzrost położenia cząstki. Gdy y maleje to oznacza, że maleje prędkość cząstki. Gdy x maleje to maleje współrzędna położenia cząstki. Gdy y rośnie to rośnie prędkość cząstki.

Tłumiony oscylator harmoniczny Powtórzmy dokładnie te same obliczenia co wcześniej. W zasadzie wystarczy tylko dodać parametr γ opisujący tłumienie oraz przemodelować lekko linijkę definiującą równania różniczkowe.

```
gamma = 0.5
de = [y, -omega^2*x-gama*y]
```

Cała reszta wykonuje się identycznie. Poniżej prezentujemy wyniki zabaw numerycznych.



Rysunek 1.10: Numeryczne rozwiązania tłumionego oscylatora harmonicznego.

1.4.2 Pole wektorowe

Prawe strony układu równań różniczkowych

$$\begin{aligned} \dot{x} &= f(x, y), & x(0) &= x_0 \\ \dot{y} &= g(x, y), & y(0) &= y_0 \end{aligned} \quad (1.80)$$

można traktować jak składowe pewnego pola wektorowego:

$$\vec{F} = [F_x, F_y] = [f(x, y), g(x, y)] \quad (1.81)$$

W każdym punkcie płaszczyzny o współrzędnych $\{x, y\}$ rysujemy wektor o składowych $[f(x, y), g(x, y)]$. W ten sposób otrzymujemy pole wektorowe. No dobrze, ale jaką informację o układzie można “wyciągnąć” z tego pola wektorowego. Wykonajmy takie oto ćwiczenie numeryczne: Startujemy z warunku początkowego $\{x_0, y_0\}$ i rysujemy w tym punkcie wektor o składowych $[f(x_0, y_0), g(x_0, y_0)]$ czyli

$$\text{w punkcie } \{x_0, y_0\} \text{ rysujemy wektor o składowych } [f(x_0, y_0), g(x_0, y_0)] \quad (1.82)$$

Jak poprzednio, rozwiązujemy numerycznie układ równań różniczkowych i obliczamy $\{x_1, y_1\}$:

$$\text{w punkcie } \{x_1, y_1\} \text{ rysujemy wektor o składowych } [f(x_1, y_1), g(x_1, y_1)] \quad (1.83)$$

Następnie obliczamy $\{x_2, y_2\}$:

$$\text{w punkcie } \{x_2, y_2\} \text{ rysujemy wektor o składowych } [f(x_2, y_2), g(x_2, y_2)] \quad (1.84)$$

W n -tym kroku iteracji obliczamy $\{x_n, y_n\}$:

$$\text{w punkcie } \{x_n, y_n\} \text{ rysujemy wektor o składowych } [f(x_n, y_n), g(x_n, y_n)] \quad (1.85)$$

Ponieważ wszystkie powyższe punkty $\{x_i, y_i\}$ leżą na krzywej fazowej, to wektory $[f(x_i, y_i), g(x_i, y_i)]$ są przyczepione do tych krzywych fazowych. Zauważamy, że wektory te są styczne do krzywej fazowej. Jeżeli $\{x_i, y_i\}$ miały by interpretację położenia cząstki na płaszczyźnie, to wektory $[f(x_i, y_i), g(x_i, y_i)]$ miały by interpretację prędkości ponieważ $\dot{x} = f(x, y)$ oraz $\dot{y} = g(x, y)$. Wiemy, że $\dot{x} = v_x$ jest x -ową składową prędkości, z kolei $\dot{y} = v_y$ jest y -ową składową prędkości. Innymi słowy, otrzymane pole wektorowe to pole prędkości fikcyjnej cząstki.

Wizualizacja pola wektorowego w systemie Sage może być dość prosto wykonane za pomocą wbudowanej funkcji `plot_vector_field`. Funkcja zwraca wykres unormowanego pola dla zadanego równania różniczkowego. My jednak zobrazujemy też powyższe idee i wykreśliśmy (w Sage oczywiście) pole dla kilku kolejnych kroków iteracji numerycznej.

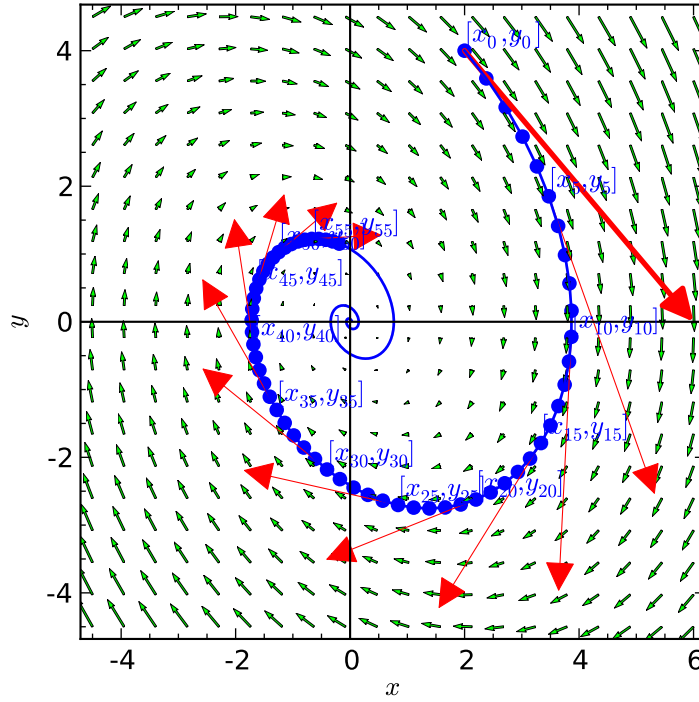
```
omega, gama = 1, 0.5
xm, ym = 4.5, 4.5
de = [y, -omega^2*x-gama*y]
plot_vector_field(de, (x, -xm, xm), (y, -ym, ym))
```

Zachęcamy do eksperymentowania z różnymi modelami czy też układami fizycznymi. Proszę tylko pamiętać, że aby powyższa numeryka zadziałała poprawnie należy poprawnie zdefiniować model fizyczny: ustalić wszystkie parametry i zmienne. Inwencja mile widziana (może warto wykorzystać dekorator `interact`)!

1.5 Dynamiczne układy zachowawcze i dysypatywne

1.5.1 Układy zachowawcze

Niektóre układy równań różniczkowych mają specyficzną strukturę i ukryte własności. Przykłady z fizyki mają taką specyficzną strukturę. Rozpatrzmy równanie Newtona dla cząstki o



Rysunek 1.11: Pole wektorowe oraz krzywe fazowe dla oscylatora tłumionego.

jednym stopniu swobody w postaci:

$$m\ddot{x} = F(x) = -\frac{dV(x)}{dx} = -V'(x) \quad (1.86)$$

gdzie siła $F(x)$ jest potencjalna, tzn. istnieje taka funkcja $V(x)$, że $F(x) = -V'(x)$. Oznaczenie $V'(x)$ jest pochodną funkcji $V(x)$ względem zmiennej x . Funkcja $V(x)$ nazywa się energią potencjalną, ale my będziemy pisali krótko - potencjał. Jeżeli znamy siłę $F(x)$ to potencjał można znaleźć obliczając całkę:

$$V(x) = -\int_a^x F(y)dy \quad (1.87)$$

gdzie a jest dowolną liczbą wybieraną tak jak nam wygodnie. Np. możemy wybrać tak, aby w pewnym punkcie potencjał był zerowy lub nieskończony.

Równanie Newtona jest równaniem 2-go rzędu, autonomicznym. Zapiszemy je w postaci

$$\begin{aligned} \dot{x} &= v, & x(0) &= x_0, \\ m\dot{v} &= F(x) = -V'(x), & v(0) &= v_0. \end{aligned} \quad (1.88)$$

Z tego wynika, że przestrzeń fazowa układu jest 2-wymiarowa $\{x, v\}$. W tej przestrzeni fazowej może analizować krzywe fazowe. Można zauważyć, że położenie $x(t)$ oraz prędkość $v(t)$ cząstki zmieniają się w czasie zgodnie z równaniem Newtona, to istnieje pewna funkcja (kombinacja) tych 2 funkcji $x(t)$ oraz $v(t)$, która nie zmienia się w czasie:

$$E[x(t), v(t)] = \frac{1}{2}mv^2(t) + V(x(t)) = \frac{1}{2}mv^2(0) + V(x(0)) = E[x(0), v(0)] \quad (1.89)$$

Wielkość E nazywa się w fizyce całkowitą energią układu i składa się z 2 części: energii kinetycznej cząstki $E_k = mv^2/2$ oraz energii potencjalnej cząstki $E_p = V(x)$. Jeżeli E nie zmienia się w czasie, to znaczy że jest to funkcja stała ze względu na czas i pochodna względem czasu powinna być zero. Sprawdźmy to:

$$\frac{dE}{dt} = \frac{d}{dt}E[x(t), v(t)] = \frac{\partial E}{\partial x} \frac{dx}{dt} + \frac{\partial E}{\partial v} \frac{dv}{dt} = V'(x)\dot{x} + mv\dot{v} = -F(x)v + vF(x) \quad (1.90)$$

gdzie skorzystaliśmy ze związku pomiędzy siłą i energią potencjalną oraz z równania Newtona.

Ponieważ E nie zmienia się w czasie, to mówimy że jest to stała ruchu lub całka ruchu, lub całka pierwsza układu (ostatnie nazwy wydają się być dziwaczne, bo w wyrażeniu dla E nie widać żadnej całki). Istnienie stałych czy też całek ruchu ułatwia analizę układów. Pokażemy to na przykładzie oscylatora harmonicznego dla którego postać siły jest dobrze znana:

$$F(x) = -kx = -m\omega^2 x, \quad V(x) = \frac{1}{2}kx^2 = \frac{1}{2}m\omega^2 x^2, \quad \omega^2 = \frac{k}{m} \quad (1.91)$$

Prawo zachowania energii mówi, że

$$E = \frac{1}{2}mv^2(t) + \frac{1}{2}kx^2(t) = \text{const.} = \frac{1}{2}mv^2(0) + \frac{1}{2}kx^2(0) \quad (1.92)$$

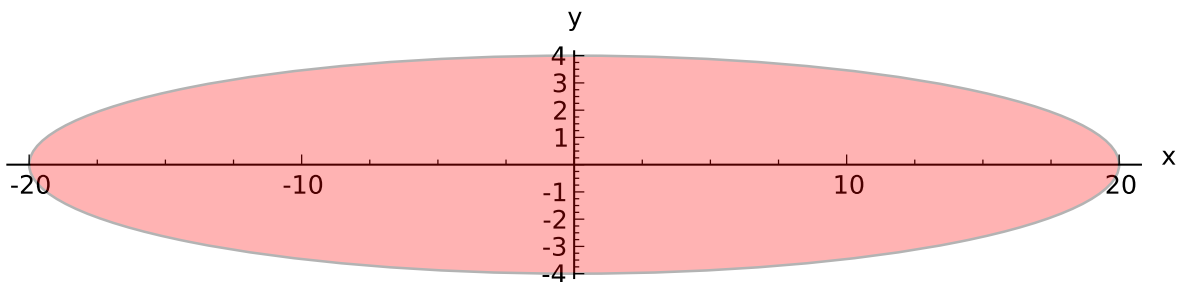
Ponieważ ta wielkość jest niezmienna w czasie, to określa równanie krzywej fazowej na płaszczyźnie XY . Łatwo zauważyć, że powyższe równanie w zmiennych $\{x, y = v\}$ ma postać

$$my^2 + kx^2 = 2E \quad (1.93)$$

Jest to równanie elipsy:

$$\frac{x^2}{(2E/k)} + \frac{y^2}{(2E/m)} = 1 \quad (1.94)$$

o osiach $a = 2E/k$ oraz $b = 2E/m$.



Rysunek 1.12: Krzywa fazowa (1.94) dla $k = 0.2$, $E = 2$, $m = 1$.

Elipsa jest krzywą zamkniętą, więc ruch jest periodyczny. Można sobie wyobrazić, że ruch cząstki w potencjale $V(x)$ jest podobny do ruchu cząstki we wnętrzu połówki sfery (w czasie). Nie jest to prawdą, ale takie wyobrażenie wyrabia w nas intuicję o własnościach ruchu. Poniżej przedstawiamy krok po kroku co zrobić, aby narysować krzywe fazowe układu.

Rysujemy wykres przedstawiający potencjał $V(x)$. Poniżej tego wykresu, z osią pionową ustawioną jak w wykresie dla potencjału, rysujemy 2 symetryczne krzywe zadane przez prawo zachowania energii $\frac{1}{2}mv^2 + V(x) = E$ czyli stąd wynika że $v = \pm \sqrt{\frac{2}{m}[(E - V(x))]}$. Te dwie krzywe $v = v(x, E)$ są krzywymi fazowymi. Cząstka porusza się w prawo gdy prędkość jest dodatnia $v > 0$ (zielona krzywa) i w lewo gdy prędkość jest ujemna $v < 0$ (czerwona krzywa). Prędkość jest zero wówczas, gdy $V(x) = E$. Wynika to z prawa zachowania energii (podstaw tam $v = 0$). Równanie $V(x) = E$ wyznacza punkty zwrotu x_i : cząstka w tych punktach ma zerową prędkość i zmienia kierunek ruchu (zawraca).

Spróbujemy, krok po kroku zanalizować równanie Newtona aby uzyskać krzywe fazowe.

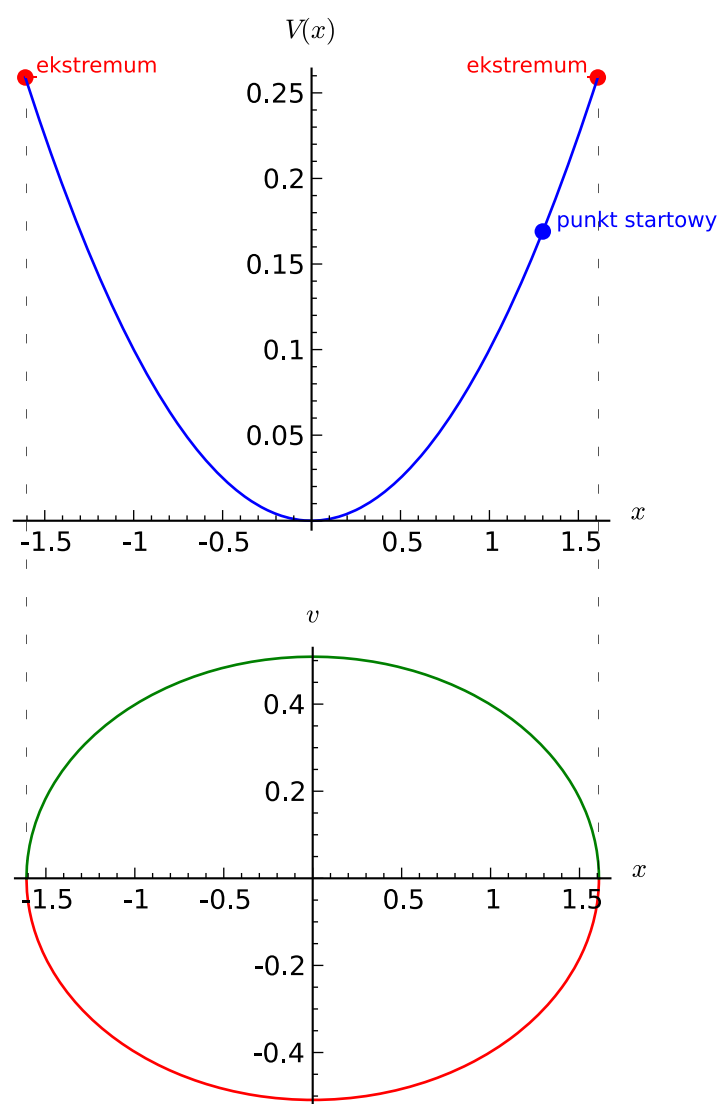
$$m\ddot{x} = F \quad (1.95)$$

Jeżeli siła będzie liniowa $F = -kx$ to dostaniemy wyżej opisane zagadnienie oscylatora harmonicznego. Na początku musimy zadeklarować nazwy zmiennych oraz parametrów użytych w modelu. Pamiętaj - każdorazowo, jeżeli chcesz obliczać coś symbolicznie, trzeba taką linijkę napisać i ją wykonać. W kolejnych liniijkach ustalimy parametry układu, zdefiniujemy siły z jakimi mamy do czynienia i obliczymy potencjał (całka z siły brana ze znakiem minus). W następnym kroku, z prawa zachowania energii, obliczymy teraz jak prędkość zależy od położenia (owe krzywe fazowe).

```
var('x y')
# parametry wizualizacji
x0, v0, k, m = 1.3, 0.3, 0.2, 1
# siły
F = -k*x
# krok 1 - wykres potencjału
V = -integral(F, x)
p1 = plot(V, xmin=-x0, xmax=x0)
#prawa zachowania energii
E = m*v0^2 + V(x=x0)
PZE = m*v^2 + V == E
#i jego rozwiązanie
rozv = solve(PZE, v)
v1=rozv[0].rhs()
v2=rozv[1].rhs()
#maksymalne wychylenia
rozv = solve(PZE(v=0), x)
xmin = rozv[0].rhs()
xmax = rozv[1].rhs()
#krzywe fazowe
p2 = plot(v1, (x,xmin,xmax), color='red')
p2 += plot(v2, (x,xmin,xmax), color='green')
```

1.5.2 Układy potencjalne

Układ o 1 stopniu swobody jest potencjalny (tzn. istnieje potencjał $V(x)$ pod warunkiem, że siła zależy tylko od położenia cząstki, tzn. $F = F(x)$). Jeżeli siła zależy także od prędkości cząstki, tzn. gdy $F = F(x, v)$, nie istnieje potencjał V taki aby $F = -V' = -dV/dx$. Dla



Rysunek 1.13: Krzywe fazowa oscylatora harmonicznego.

układów o wielu stopniach swobody, opisywanych układem równań Newtona

$$m_i \frac{d^2 \vec{r}_i}{dt^2} = \vec{F}_i(\vec{r}_1, \vec{r}_2, \vec{r}_3, \dots, \vec{r}_N) \quad (1.96)$$

dla N cząstek, układ jest potencjalny, gdy istnieje taka funkcja skalarna $V(\vec{r}_1, \vec{r}_2, \vec{r}_3, \dots, \vec{r}_N)$, że siła działająca na i -tą cząstkę jest gradientem potencjału ze znakiem minus. Prościej jest to wyjaśnić na przykładzie 1 cząstki poruszającej się w przestrzeni 3-wymiarowej:

$$\begin{aligned} m \frac{d^2 x}{dt^2} &= F_1(x, y, z) = -\frac{\partial}{\partial x} V(x, y, z), \\ m \frac{d^2 y}{dt^2} &= F_2(x, y, z) = -\frac{\partial}{\partial y} V(x, y, z), \\ m \frac{d^2 z}{dt^2} &= F_3(x, y, z) = -\frac{\partial}{\partial z} V(x, y, z). \end{aligned} \quad (1.97)$$

W ogólnym przypadku, gdy mamy zadane 3 składowe siły F_1, F_2 oraz F_3 , nie musi istnieć tylko jedna funkcja V taka aby powyższe równania były spełnione. Nasuwa się pytanie, czy istnieje proste kryterium mówiące, że układ jest potencjalny. Jeżeli

$$\vec{F} = -\text{grad } V \quad \text{to} \quad \text{rot } \vec{F} = -\text{rot grad } V = -\vec{\nabla} \times \vec{\nabla} V \equiv 0 \quad (1.98)$$

gdzie operator $\vec{\nabla}$ jest operatorem różniczkowania

$$\vec{\nabla} = \hat{e}_x \frac{\partial}{\partial x} + \hat{e}_y \frac{\partial}{\partial y} + \hat{e}_z \frac{\partial}{\partial z} \quad (1.99)$$

Wystarczy zatem sprawdzić, czy rotacja pola sił \vec{F} jest 0.

Zadanie

Sprawdzić, czy siły $\vec{F}(x, y, z)$ o zadanych składowych $F_i = F_i(x, y, z)$ są potencjalne.

1. Układ 1.

$$F_1 = \frac{y}{x^2 + y^2 + z^2}, \quad F_2 = -\frac{x}{x^2 + y^2 + z^2}, \quad F_3 = \frac{z}{x^2 + y^2 + z^2} \quad (1.100)$$

2. Układ 2.

$$F_1 = \frac{x - z}{x^2 + y^2}, \quad F_2 = xe^{-y^2}, \quad F_3 = z + 5 \quad (1.101)$$

3. Układ 3.

$$F_1 = 25x^4y - 3y^2, \quad F_2 = 5x^5 - 6xy - 5, \quad F_3 = 0 \quad (1.102)$$

Jeżeli układ jest potencjalny to łatwo sprawdzić, podobnie jak wyżej w przypadku układu o 1-stopniu swobody, że istnieje stała ruchu - całkowita energia układu:

$$E = \sum_i \frac{m\vec{v}_i^2}{2} + V(\vec{r}_1, \vec{r}_2, \vec{r}_3, \dots, \vec{r}_N) = \text{cons.}, \quad \frac{dE}{dt} = 0 \quad (1.103)$$

Dlatego takie pole sił nazywa się zachowawczym polem sił. Wszystkie siły związane z potencjalnym polem sił są siłami zachowawczymi. Istnieją jednak siły, które nie są siłami potencjalnymi, mimo to pozostają siłami zachowawczymi. Przykładem może być siła Lorentza działająca na naładowaną cząstkę poruszającą się w polu magnetycznym. Nie należy tego mylić z zachowawczymi układami dynamicznymi. Tę kwestię postaramy się teraz wyjaśnić.

1.5.3 Dynamiczne układy zachowawcze i dysypatywne

W teorii układów dynamicznych ważną rolę pełnią dwa pojęcia: zachowawcze układy dynamiczne i dysypatywne układy dynamiczne. Znowu dla jasności wywodu rozpatrzmy przykład układu o 3-wymiarowej przestrzeni fazowej:

$$\begin{aligned}\dot{x} &= F_1(x, y, z), & x(0) &= x_0, \\ \dot{y} &= F_2(x, y, z), & y(0) &= y_0, \\ \dot{z} &= F_3(x, y, z), & z(0) &= z_0.\end{aligned}\tag{1.104}$$

Wybieramy w przestrzeni fazowej obszar $D(0)$ o objętości $M(0)$. Zawiera on wszystkie możliwe warunki początkowe

$$\{x_0, y_0, z_0\} \in D(0)\tag{1.105}$$

Pod wpływem ewolucji każdy punkt (x_0, y_0, z_0) z tego obszaru przejdzie po czasie t do punktu $(x(t), y(t), z(t))$. Zbiór tych punktów w chwili t tworzy obszar $D(t)$ o objętości $M(t)$. Zachodzi pytanie:

$$\text{w jakich przypadkach } M(t) = M(0)\tag{1.106}$$

Innymi słowy, kiedy układ dynamiczny zachowuje objętość fazową. Zbadamy ten problem. Wprowadzimy nowe oznaczenia, aby ułatwić notację:

$$x_t = x(t), \quad y_t = y(t), \quad z_t = z(t)\tag{1.107}$$

Objętość fazowa warunków początkowych w chwili $t = 0$ wynosi

$$M(0) = \int \int \int_{D(0)} dx_0 dy_0 dz_0\tag{1.108}$$

Objętość fazowa w chwili t wynosi

$$M(t) = \int \int \int_{D(t)} dx_t dy_t dz_t\tag{1.109}$$

Ewolucja układu to nic innego jak zamiana zmiennych $(x_0, y_0, z_0) \rightarrow (x_t, y_t, z_t)$. Dokonajmy tej zamiany zmiennych w drugiej całce:

$$\begin{aligned}M(t) &= \int \int \int_{D(t)} dx_t dy_t dz_t = \\ &= \int \int \int_{D(0)} \frac{\partial(x_t, y_t, z_t)}{\partial(x_0, y_0, z_0)} dx_0 dy_0 dz_0 = \int \int \int_{D(0)} J(t) dx_0 dy_0 dz_0\end{aligned}\tag{1.110}$$

gdzie J jest jacobianem transformacji $(x_t, y_t, z_t) \rightarrow (x_0, y_0, z_0)$. Jeżeli objętość fazowa nie zmienia się w czasie (jest funkcją stałą), to jej pochodna

$$\frac{dM(t)}{dt} = \int \int \int_{D(0)} \frac{dJ(t)}{dt} dx_0 dy_0 dz_0 \quad (1.111)$$

wynosi zero. Jeżeli

$$\frac{dJ(t)}{dt} = 0 \quad \text{to} \quad \frac{dM(t)}{dt} = 0 \quad \text{czyli} \quad M(t) = M(0) \quad (1.112)$$

Więc rozpoczynamy obliczenia

$$\frac{dJ(t)}{dt} = \frac{d}{dt} \frac{\partial(x_t, y_t, z_t)}{\partial(x_0, y_0, z_0)} = \frac{d}{dt} \begin{bmatrix} \frac{\partial x_t}{\partial x_0} & \frac{\partial x_t}{\partial y_0} & \frac{\partial x_t}{\partial z_0} \\ \frac{\partial y_t}{\partial x_0} & \frac{\partial y_t}{\partial y_0} & \frac{\partial y_t}{\partial z_0} \\ \frac{\partial z_t}{\partial x_0} & \frac{\partial z_t}{\partial y_0} & \frac{\partial z_t}{\partial z_0} \end{bmatrix} \quad (1.113)$$

Należy powyższy wyznacznik rozwinąć i pamiętać, że rozwiązania równań różniczkowych

$$x_t = x_t(x_0, y_0, z_0), \quad y_t = y_t(x_0, y_0, z_0), \quad z_t = z_t(x_0, y_0, z_0) \quad (1.114)$$

zależą od warunków początkowych $\{x_0, y_0, z_0\}$. Po rozwinięciu wyznacznika pojawiają się wyrażenia typu

$$\frac{d}{dt} \frac{\partial x_t}{\partial z_0} = \frac{\partial}{\partial z_0} \frac{dx_t}{dt} = \frac{\partial}{\partial z_0} \dot{x}_t = \frac{\partial}{\partial z_0} F_1(x_t, y_t, z_t) = \frac{\partial F_1}{\partial x_t} \frac{\partial x_t}{\partial z_0} + \frac{\partial F_1}{\partial y_t} \frac{\partial y_t}{\partial z_0} + \frac{\partial F_1}{\partial z_t} \frac{\partial z_t}{\partial z_0} \quad (1.115)$$

Jak widać, w tym prostym przypadku musimy przeprowadzić uciążliwe rachunki. Znacznie lepiej jest posłużyć się rachunkiem symbolicznym z wykorzystaniem SAGE. Aby przeprowadzić dowód, najlepiej jest obejść ograniczenia operacji na wyrażeniach z pochodnymi w Sage. Pochodna wyznacznika jest zrobiona automatycznie, potem jest ręcznie wykonane podstawienie:

$$\frac{\partial}{\partial z_0} \dot{x}_t = \frac{\partial F_1}{\partial x_t} \frac{\partial x_t}{\partial z_0} + \frac{\partial F_1}{\partial y_t} \frac{\partial y_t}{\partial z_0} + \frac{\partial F_1}{\partial z_t} \frac{\partial z_t}{\partial z_0} \quad (1.116)$$

```
for i, (v, lv) in enumerate(xy_wsp):
    for j, (u, lu) in enumerate(xy_wsp):
        J[i, j] = var("d%sd%s"%(v, u),
            latex_name = \
                r'\displaystyle\frac{\partial %s_t}{\partial %s_0}'%(lv, lu)
        )
        var("dF%sd%s"%(v, u),
            latex_name = \
                r'\displaystyle\frac{\partial F_%s}{\partial %s_t}'%(lv, lu)
        )
#
to_fun = dict()
for v in J.list():
    vars()[str(v).capitalize()] = function(str(v).capitalize(), t)
    var("%sd"%str(v))
```

```

to_fun[v]=vars()[str(v).capitalize()]
to_fun[vars()[str(v)+"d"]]=vars()[str(v).capitalize()].diff()
to_var = dict((v,k) for k,v in to_fun.items())
#
to_rhs = dict()
for i, (v,lv) in enumerate(xy_wsp):
    for j, (u,lu) in enumerate(xy_wsp):
        to_rhs[vars()[str(v)+"d"+str(u)]] = \
            sum([vars()[str(v)+"d"+str(w)]*vars()[str(w)+"d"+str(u)] \
                for w,wl in xy_wsp])

#macierz Jakobiego
show(J)
#Wszystkie pochodne cząstkowe są reprezentowane przez niezależne
#zmienne, aby policzyć pochodną wyznacznika, zamieniamy je
#podstawiając słownik zamieniający zmienne na funkcje:
show(J.subs(to_fun))
#Liczymy wyznaczniki pochodną, oraz wracamy do zmiennych
#symbolicznych
J.subs(to_fun).det().diff(t).subs(to_var).show()
#
#Używając słownika to_rhs, podstawiamy prawe strony ODE
J.subs(to_fun).det().diff(t).subs(to_var).subs(to_rhs).show()
#
#Ostatecznie dzielimy otrzymany wzór przez Jacobian
final = \
    J.subs(to_fun).det().diff(t).subs(to_var).subs(to_rhs)/J.det()
final.simplify_full().show()

```

Ostatecznie otrzymamy wyrażenie

$$\frac{dJ(t)}{dt} = J(t) \left[\frac{\partial F_1}{\partial x_t} + \frac{\partial F_2}{\partial y_t} + \frac{\partial F_3}{\partial z_t} \right] = J(t) \operatorname{div} \vec{F} \quad (1.117)$$

To, co jest w nawiasie kwadratowym nazywa się dywergencją pola wektorowego \vec{F} . Wstawiamy to wyrażenie do równania (1.111) i otrzymamy

$$\begin{aligned} \frac{dM(t)}{dt} &= \int \int \int_{D(0)} \frac{dJ(t)}{dt} dx_0 dy_0 dz_0 = \\ &= \int \int \int_{D(0)} J(t) \operatorname{div} \vec{F} dx_0 dy_0 dz_0 = \\ &= \int \int \int_{D(t)} \operatorname{div} \vec{F} dx_t dy_t dz_t \end{aligned} \quad (1.118)$$

gdzie dokonaliśmy odwrotnego przejścia (z prawej strony na lewą stronę) jak w równaniu (1.110).

Można teraz uogólnić ten wynik na dowolną ilość wymiarów przestrzeni fazowej dla układu równań

$$\frac{d\vec{x}}{dt} = \vec{F}(\vec{x}), \quad \vec{x} = [x_1, x_2, x_3, \dots, x_n], \quad \vec{F} = [F_1, F_2, F_3, \dots, F_n] \quad (1.119)$$

i otrzymamy

Twierdzenie Jeżeli dywergencja pola wektorowego \vec{F} danego równania różniczkowego jest zero,

$$\operatorname{div} \vec{F} = \sum_i \frac{\partial F_i}{\partial x_i} = 0 \quad (1.120)$$

wówczas objętość fazowa jest zachowana, $M(t) = M(0)$. Takie układy dynamiczne nazywamy zachowawczymi. Jeżeli objętość fazowa maleje w czasie, to układ nazywamy dyssypatywnym. Innymi słowy, układ jest dyssypatywny gdy objętość $M(t) < M(0)$ dla $t > 0$. Oznacza to, że dla układów dyssypatywnych

$$\frac{dM(t)}{dt} < 0 \quad (1.121)$$

Gdyby

$$\operatorname{div} \vec{F} = C_0 = \text{const.} \quad (1.122)$$

wówczas z równania (1.118) otrzymujemy prostą relację

$$\frac{dM(t)}{dt} = C_0 M(t) \quad (1.123)$$

która pozwala rozstrzygnąć czy układ jest dyssypatywny.

Przykład 1: Oscylator harmoniczny tłumiony

$$\begin{aligned} \dot{x} &= y = F_1(x, y), & x(0) &= x_0, \\ \dot{y} &= -\gamma y - \omega^2 x = F_2(x, y), & y(0) &= y_0. \end{aligned} \quad (1.124)$$

Łatwo obliczyć dywergencję pola

$$\operatorname{div} \vec{F} = \frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} = -\gamma < 0 \quad (1.125)$$

Równanie (1.122) przyjmuje postać

$$\frac{dM(t)}{dt} = -\gamma M(t), \quad \text{jego rozwiązaniem jest funkcja malejąca} \quad M(t) = M(0)e^{-\gamma t} \quad (1.126)$$

czyli objętość fazowa (w tym przypadku powierzchnia fazowa) maleje w czasie i dlatego jest to dyssypatywny układ dynamiczny.

Przykład 2: Model Lorenza

$$\begin{aligned} \dot{x} &= \sigma(y - x) = F_1(x, y, z), & x(0) &= x_0, \\ \dot{y} &= x(\rho - z) - y = F_2(x, y, z), & y(0) &= y_0, \\ \dot{z} &= xy - \beta z = F_3(x, y, z), & z(0) &= z_0. \end{aligned} \quad (1.127)$$

gdzie wszystkie parametry są dodatnie: $\sigma, \rho, \beta > 0$.

Obliczmy dywergencję 3-wymiarowego pola $\vec{F} = [F_1, F_2, F_3]$. Proste rachunki pokazują, że

$$\operatorname{div} \vec{F} = \frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} + \frac{\partial F_3}{\partial z} = -\sigma - 1 - \beta < 0 \quad (1.128)$$

Objętość fazowa (w tym przypadku faktycznie objętość w 3 wymiarowej przestrzeni) maleje eksponencjalnie w czasie, podobnie jak w poprzednim przykładzie. Dlatego też jest to dysypatywny układ dynamiczny.

1.6 Stany stacjonarne i ich stabilność

Czasami rozwiązanie równań różniczkowych dąży do stałej wartości dla długich czasów, $t \rightarrow \infty$. Mówimy wówczas, że istnieje rozwiązanie stacjonarne (stan stacjonarny, punkt stały, punkt równowagi):

$$\lim_{t \rightarrow \infty} \vec{x}(t) = \vec{x}_s \quad (1.129)$$

Może istnieć kilka stanów stacjonarnych, a nawet nieskończenie wiele stanów stacjonarnych. Który z tych stanów się realizuje, zależy to od warunku początkowego $\vec{x}(0) = \vec{x}_0$. Rozwiązanie stacjonarne \vec{x}_s nie zależy od czasu. Nazwa “rozwiązanie stacjonarne” nie jest bezpodstawne. Faktycznie jest to rozwiązanie układu równań różniczkowych

$$\frac{d\vec{x}}{dt} = \vec{F}(\vec{x}), \quad \vec{x}(0) = \vec{x}_0 \quad (1.130)$$

z warunkiem początkowym $\vec{x}(0) = \vec{x}_s$. Jeżeli $\vec{x}(t) = \vec{x}_s$ jest rozwiązaniem stacjonarnym to musi spełniać układ (1.130), czyli

$$\text{jeżeli } \vec{x}(t) = \vec{x}_s \quad \text{to} \quad \frac{d\vec{x}(t)}{dt} = \frac{d\vec{x}_s}{dt} = 0 = \vec{F}(\vec{x}(t)) = \vec{F}(\vec{x}_s) = 0 \quad (1.131)$$

Innymi słowy, stan stacjonarny wyznaczany jest z warunku:

$$\vec{F}(\vec{x}_s) = 0 \quad (1.132)$$

Powyższy warunek to układ n -równań algebraicznych. Zwykle udaje nam się go rozwiązać analitycznie w niewielu przypadkach, w szczególności gdy wymiar przestrzeni fazowej $n > 1$. Natomiast możemy taki układ rozwiązywać numerycznie. Jeżeli już wyznaczymy stan stacjonarny, to nasuwa się pytanie na ile jest on stabilny, tzn. gdy nieco wytrącimy układ z tego stanu to czy powróci on do poprzedniego stanu stacjonarnego czy też oddali się od niego. Na przykład stanem stacjonarnym kulki poruszającej się na nitce w polu ziemskim jest położenie pionowe. Gdy kulkę wychylimy z tego położenia, po dostatecznie długim czasie powróci ona do pozycji pionowej i tak tam pozostanie nieruchoma. Jest to stabilny stan równowagi. Rozważmy teraz kulkę mogącą poruszać się tylko po sferze. Gdy umieścimy kulkę na biegunie północnym sfery w polu ziemskim to nieskończenie małe zaburzenie spowoduje, że kulka spadnie z tego położenia i nigdy do niego nie powróci. W obu tych przypadkach zakładamy rzeczywiste warunki ruchu z tarciem. Pominięcie tarcia spowoduje radykalnie różne zachowanie. Te dwa przykłady pozwalają nabyć intuicję, co to znaczy że stan stacjonarny jest stabilny

lub jest niestabilny. Z grubsza można powiedzieć, że stan stacjonarny \vec{x}_s jest stabilny jeśli każda trajektoria startującego z punktu bliskiego \vec{x}_s pozostaje blisko \vec{x}_s wraz z upływem czasu. Natomiast \vec{x}_s jest niestabilny gdy każda trajektoria startującego z punktu bliskiego \vec{x}_s oddala się od tego punktu gdy $t \rightarrow \infty$. Można podać bardziej precyzyjne definicje.

Definicja Mówimy, że stan stacjonarny \vec{x}_s jest stabilny jeżeli dla dowolnego $\epsilon > 0$ istnieje takie $\delta(\epsilon) > 0$, że dla każdego \vec{x}_0 takiego że $|\vec{x}_0 - \vec{x}_s| < \delta$ rozwiązanie $\vec{x}(t)$ spełnia nierówność: $|\vec{x}(t) - \vec{x}_s| < \epsilon$ dla dowolnych czasów $t > 0$. Jeżeli dodatkowo $\lim_{t \rightarrow \infty} \vec{x}(t) = \vec{x}_s$ to stan stacjonarny \vec{x}_s jest asymptotycznie stabilny.

Innymi słowy dla stabilnych stanów rozwiązanie $\vec{x}(t)$ jest cały czas blisko rozwiązania stacjonarnego \vec{x}_s , a dla asymptotycznie stabilnych stanów rozwiązanie $\vec{x}(t)$ dąży do \vec{x}_s gdy czas $t \rightarrow \infty$.

1.6.1 Przypadek A: Jedno równanie różniczkowe

Przykład: Równanie różniczkowe liniowe

$$\dot{x} = \lambda x = f(x), \quad \lambda \in R \quad (1.133)$$

Stan stacjonarny znajdujemy jako rozwiązanie równania

$$f(x_s) = \lambda x_s = 0 \quad \text{stąd otrzymujemy stan stacjonarny} \quad x_s = 0 \quad (1.134)$$

Pytamy, czy ten stan jest stabilny. Musimy zaburzyć rozwiązanie stacjonarne $x(t) = x_s = 0$ i wystartować z dostatecznie bliskiego w stosunku do $x_s = 0$ warunku początkowego X_0 . Rozwiązaniem równania różniczkowego liniowego jest funkcja

$$x(t) = X_0 e^{\lambda t} \quad (1.135)$$

Jeżeli

$$\lambda < 0 \quad \text{to} \quad x(t) \rightarrow 0 \quad \text{dla wszystkich } X_0 \text{ bliskich } 0 \quad (1.136)$$

Wówczas stan stacjonarny $x_s = 0$ jest stabilny i dodatkowo jest asymptotycznie stabilny ponieważ rozwiązanie to dąży do $x_s = 0$ gdy $t \rightarrow \infty$.

Jeżeli

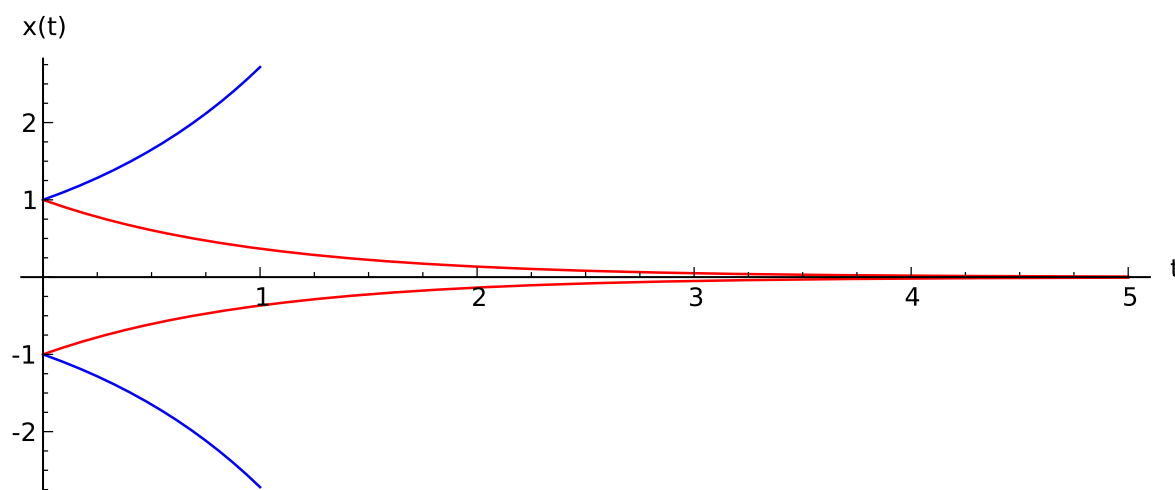
$$\lambda > 0 \quad \text{to} \quad x(t) \rightarrow \infty \quad \text{dla wszystkich } X_0 \text{ bliskich } 0 \quad (1.137)$$

Wówczas stan stacjonarny $x_s = 0$ jest niestabilny.

```
p1=plot(exp(-x),(x,0,5),color="red")
p2=plot(exp(x),(x,0,1))
p3=plot(-exp(-x),(x,0,5),color="red")
p4=plot(-exp(x),(x,0,1))
show(p1+p2+p3+p4,axes_labels=['t','x(t)'])
```

Na rysunku przedstawiono zagadnienie stabilności: czerwone krzywe dążą do 0 gdy $t \rightarrow \infty$. Niebieskie krzywe rozbiegają się (uciekają od 0) gdy $t \rightarrow \infty$.

Warunki początkowe $x(0) = \pm 0.05$ są blisko stanu stacjonarnego $x_s = 0$. Przykład ten sugeruje nam metodę badania stabilności stanu stacjonarnego. Teraz podamy tę metodę.



Rysunek 1.14: Stabilność punktów stacjonarnych

1.6.2 Metoda linearyzacji badania stabilności

Rozpatrujemy układ 1-wymiarowy:

$$\frac{dx}{dt} = \dot{x} = f(x), \quad f(x_s) = 0 \quad (1.138)$$

Aby zbadać stabilność stanu x_s , analizujemy zaburzenie

$$h(t) = x(t) - x_s, \quad |h(0)| = |x(0) - x_s| < \delta \quad (1.139)$$

Funkcja $h(t)$ powinna być mała, gdy stan x_s jest stabilny. Jak daleko jest rozwiązanie $x(t)$ od rozwiązania x_s . Zobaczmy, jakie równanie różniczkowe spełnia $h(t)$:

$$\begin{aligned} \frac{dh}{dt} &= \frac{d}{dt}[x(t) - x_s] \\ &= \frac{dx}{dt} = f(x) = f(x_s + h) = f(x_s) + f'(x_s)h + \frac{1}{2}f''(x_s)h^2 + \frac{1}{3!}f'''(x_s)h^3 + \dots \end{aligned} \quad (1.140)$$

Ponieważ $f(x_s) = 0$, otrzymujemy równanie różniczkowe dla odchylenia $h(t)$ od stanu stacjonarnego

$$\frac{dh}{dt} = f'(x_s)h + \frac{1}{2}f''(x_s)h^2 + \frac{1}{3!}f'''(x_s)h^3 + \dots \quad (1.141)$$

Jeżeli $f'(x_s) \neq 0$, to pierwszy istotny wyraz w tym równaniu jest liniowy względem h . Wyrazy h^2, h^3 i wyższych potęg są pomijalnie małe. Jeżeli np. $h = 10^{-2}$ to $h^2 = 10^{-4}, h^3 = 10^{-6}$. Wówczas h^2, h^3 i wyższe potęgi h można pominąć jako bardzo małe. Otrzymujemy równanie różniczkowe liniowe

$$\frac{dh}{dt} = \lambda h, \quad \lambda = f'(x_s) \quad (1.142)$$

z rozwiązaniem

$$h(t) = h(0)e^{\lambda t} \quad (1.143)$$

Wiemy już z powyższego przykładu, że gdy $\lambda < 0$ to $h(t) \rightarrow 0$ gdy $t \rightarrow \infty$. To oznacza, że zaburzenie $x(t) \rightarrow x_s$ gdy $t \rightarrow \infty$ i wówczas stan stacjonarny x_s jest stabilny. Jeżeli $\lambda > 0$ to $|h(t)| \rightarrow \infty$ gdy $t \rightarrow \infty$. To oznacza, że zaburzenie $x(t)$ ucieka od x_s gdy $t \rightarrow \infty$ i wówczas stan stacjonarny x_s jest niestabilny. Otrzymujemy następujące kryterium na stabilność stanu stacjonarnego:

- Jeżeli $\lambda = f'(x_s) < 0$ to stan stacjonarny jest stabilny,
- jeżeli $\lambda = f'(x_s) > 0$ to stan stacjonarny jest niestabilny,
- jeżeli $\lambda = f'(x_s) = 0$ to nie wiem nic na temat stabilności. Musimy badać następne niezerowe wyrazy. Jeżeli $f''(x_s) \neq 0$ zatrzymujemy pierwszy nieznikający wyraz czyli badamy równanie

$$\frac{dh}{dt} = \gamma h^2, \quad \gamma = \frac{1}{2}f''(x_s) \quad (1.144)$$

Jeżeli $f'(x_s) = 0$ oraz $f''(x_s) = 0$ to bierzemy następny nieznikający wyraz i badamy równanie

$$\frac{dh}{dt} = \nu h^3, \quad \nu = \frac{1}{3!}f'''(x_s) \quad (1.145)$$

Jeżeli w tych przypadkach $h(t) \rightarrow 0$ gdy $t \rightarrow \infty$, to stan stacjonarny x_s jest stabilny. W przeciwnym przypadku - nie jest stabilny.

1.6.3 Metoda potencjału badania stabilności

W jednym wymiarze, równanie różniczkowe zawsze można przedstawić w postaci

$$\frac{dx}{dt} = \dot{x} = f(x) = -\frac{dV(x)}{dx} = -V'(x), \quad f(x_s) = 0 \quad (1.146)$$

gdzie funkcja

$$V(x) = -\int f(x)dx \quad (1.147)$$

nazywana jest “potencjałem”. W ogólności to nie jest potencjał fizyczny który pojawia się w równaniu Newtona dla cząstki z tłumieniem:

$$m\ddot{x} + \gamma\dot{x} = -V'(x) \quad (1.148)$$

Jeżeli ruch cząstki odbywa się w środowisku o dużym tarcu, w tzw. reżimie przetłumionym, w którym przyspieszenie cząstki jest znikomo małe (formalnie gdy $m = 0$), wówczas równanie Newtona ma postać

$$\gamma\dot{x} = -V'(x) \quad (1.149)$$

które po przeskalowaniu ma postać:

$$\dot{x} = -\frac{1}{\gamma}V'(x) = -\tilde{V}'(x) \quad (1.150)$$

Stąd też “historycznie” wywodzi się nazwa potencjał dla abstrakcyjnego układu dynamicznego:

$$\dot{x} = f(x) == -V'(x) \quad (1.151)$$

Stan stacjonarny x_s określony równaniem

$$f(x_s) = -V'(x_s) = 0 \quad (1.152)$$

to punkt ekstremalny potencjału (o ile pochodna parzystego rzędu $V^{(2k)}(x_s) \neq 0$). Punkt x_s jest stabilny gdy

$$\lambda = f'(x_s) = -V''(x_s) < 0 \quad (1.153)$$

czyli $V''(x_s) > 0$. To odpowiada minimum potencjału. Natomiast punkt x_s niestabilny odpowiada maksimum potencjału. Mamy więc klarowny obraz: Rysujemy potencjał $V(x)$. Minima potencjału to stabilne stany stacjonarne; maksima potencjału to niestabilne stany stacjonarne.

Zadania

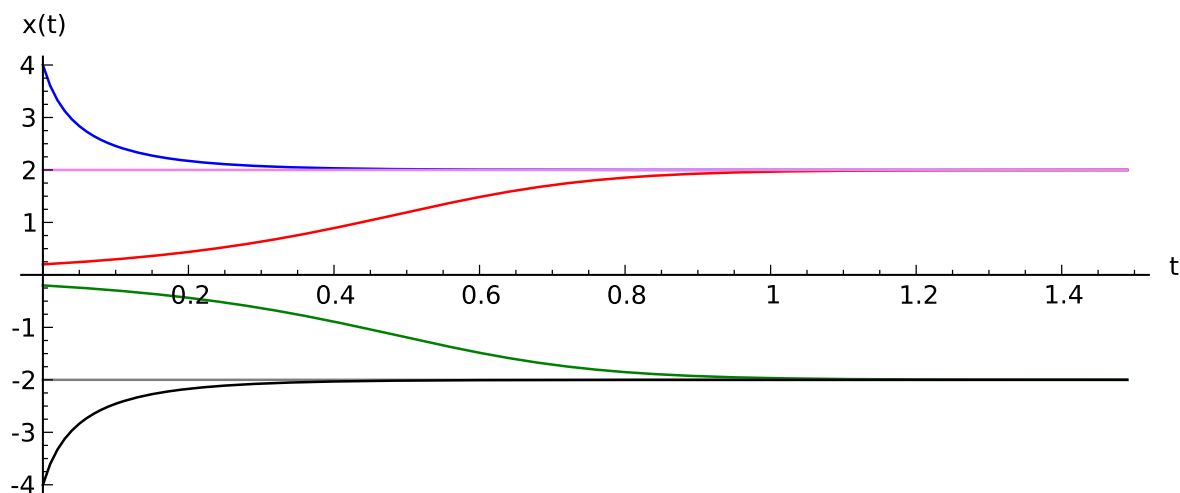
Wyznacz stany stacjonarne i zbadaj ich (asymptotyczną) stabilność. Korzystaj z metody linearyzacji i metody potencjału.

1. $\dot{x} = 4x - x^3$
2. $\dot{x} = 1 + x^4$
3. $\dot{x} = 3 \sin(x)$
4. $\dot{x} = 2x \sin(x)$
5. $\dot{x} = -x^2 \sin(x)$

Poniżej pokazujemy wyniki dla zadania 1. Są 3 stany stacjonarne: $x_{s1} = 2, x_{s2} = 0, x_{s3} = -2$. Stany $x_{s1} = 2$ oraz $x_{s3} = -2$ są asymptotycznie stabilne (rozwiązania dążą do tych stanów). Stan $x_{s2} = 0$ jest niestabilny (rozwiązania uciekają od tego stanu).

```
var('x,y,z,u,Z,Y,t')
T0 = srange(0,1.5,0.01)
f11 = 4*x-x^3
f12 = 4*y-y^3
f13 = 4*z-z^3
f15 = 4*u-u^3
f16 = 0
de = vector([f11,f12,f13,0,0,f15])
ic = [4,0.2,-0.2,2,-2,-4]
sol5=desolve_odeint(de,ic,T0,[x,y,z,Z,Y,u])
line(zip(T0,sol5[:,0]),figsize=(7,4))+\
line(zip(T0,sol5[:,1]),color='red')+ \
line(zip(T0,sol5[:,2]),color='green')+ \
line(zip(T0,sol5[:,4]),color='gray')+ \
line(zip(T0,sol5[:,5]),color='black')+ \
line(zip(T0,sol5[:,3]),color='violet')
```

Z powyższego przykładu zauważmy następujące cechy:



Rysunek 1.15: Stabilność punktów stacjonarnych

1. Zbiór warunków początkowych dzieli się na dwa podzbiory $A_1 = (-\infty, 0)$ oraz $A_2 = (0, \infty)$. Dla warunków początkowych ze zbioru A_1 rozwiązania $x(t) \rightarrow -2$ dla $t \rightarrow \infty$, a dla warunków początkowych ze zbioru A_2 rozwiązania $x(t) \rightarrow 2$ dla $t \rightarrow \infty$.
2. Krzywe $x(t)$ są funkcjami monotonicznymi: albo cały czas rosną w czasie, albo cały czas maleją gdy czas rośnie. Dlaczego? Rozważmy 2 przykłady warunków początkowych widocznych na rysunku:
 1. gdy $x(0) = 4$ to $\dot{x}(x = 4) = 4 \cdot 4 - 4^3 = -48 < 0$. Pochodna jest ujemna, a to oznacza że funkcja maleje. Podobnie jest dla wszystkich warunków początkowych $x(0) > 2$
 2. gdy $x(0) = 0.2$ to $\dot{x}(x = 0.2) = 4 \cdot 0.2 - (0.2)^3 > 0$. Pochodna jest dodatnia, a to oznacza że funkcja rośnie. Podobnie jest dla wszystkich warunków początkowych $x(0) \in (0, 2)$
3. Różne krzywe $x(t)$ nigdy się nie przecinają. Wynika to z tego, że rozwiązania są jedyne i jednoznaczne. Jeżeli by się przecinały, to z punktu przecięcia wychodziłoby kilka rozwiązań. A to jest niemożliwe na mocy twierdzenia o jednoznaczności rozwiązań.

1.6.4 Przypadek B: Układ 2 równań różniczkowych

Dla jasności prezentacji, rozpatrujemy układ 2 równań różniczkowych

$$\begin{aligned}\dot{x} &= f(x, y), \\ \dot{y} &= g(x, y).\end{aligned}\tag{1.154}$$

Stany stacjonarne (x_s, y_s) określone są jako rozwiązania układu równań algebraicznych

$$\begin{aligned}f(x_s, y_s) &= 0, \\ g(x_s, y_s) &= 0.\end{aligned}\tag{1.155}$$

Gdyby istniał potencjał $V(x, y)$, powyżej przedstawione wnioski otrzymane z metody potencjału są także prawdziwe dla układów wielowymiarowych. Ponieważ w ogólnym przypadku

“potencjał” nie musi istnieć, zbadamy stabilność stanów (x_s, y_s) stosując metodę linearyzacji. Definiujemy nowe funkcje

$$\begin{aligned} X(t) &= x(t) - x_s, \\ Y(t) &= y(t) - y_s. \end{aligned} \quad (1.156)$$

Charakteryzują one odchylenie od stanu stacjonarnego, które są małe gdy stan stacjonarny jest stabilny. Zobaczmy, jakie równania różniczkowe spełniają te funkcje:

$$\begin{aligned} \dot{X}(t) &= \dot{x}(t) - \dot{x}_s = \dot{x}(t) = \\ &= f(x(t), y(t)) = f(x_s + X(t), y_s + Y(t)) = f(x_s, y_s) + \frac{\partial f}{\partial x} X + \frac{\partial f}{\partial y} Y + \dots \end{aligned} \quad (1.157)$$

$$\begin{aligned} \dot{Y}(t) &= \dot{y}(t) - \dot{y}_s = \dot{y}(t) = \\ &= g(x(t), y(t)) = g(x_s + X(t), y_s + Y(t)) = g(x_s, y_s) + \frac{\partial g}{\partial x} X + \frac{\partial g}{\partial y} Y + \dots \end{aligned} \quad (1.158)$$

Wszystkie pochodne cząstkowe są obliczane w punkcie (x_s, y_s) . Ponieważ w punkcie stacjonarnym $f(x_s, y_s) = g(x_s, y_s) = 0$ otrzymujemy zlinearyzowane równania różniczkowe w postaci

$$\dot{X} = a_{11}X + a_{12}Y \quad (1.159)$$

$$\dot{Y} = a_{21}X + a_{22}Y \quad (1.160)$$

gdzie macierz współczynników

$$J = \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \quad \text{obliczona w punkcie } (x_s, y_s) \quad (1.161)$$

jest macierzą Jacobiego. Ponieważ otrzymaliśmy liniowy układ równań różniczkowych, jego rozwiązania poszukujemy w postaci funkcji

$$X(t) = Ae^{\lambda t}, \quad Y(t) = Be^{\lambda t} \quad (1.162)$$

gdzie stałe A oraz B są określone przez warunki początkowe $(X(0), Y(0))$.

Zauważamy, że

$$\dot{X} = \lambda X, \quad \dot{Y} = \lambda Y \quad (1.163)$$

Wstawiamy te relacje do układu równań zlinearyzowanych:

$$\lambda A = a_{11}A + a_{12}B \quad (1.164)$$

$$\lambda Y = a_{21}A + a_{22}B \quad (1.165)$$

Jest to zagadnienie własne dla macierzy Jacobiego, gdzie λ to są wartości własne macierzy Jacobiego. To jest także równoważne liniowemu układowi równań algebraicznych

$$(a_{11} - \lambda)A + a_{12}B = 0 \quad (1.166)$$

$$a_{21}A + (a_{22} - \lambda)B = 0 \quad (1.167)$$

Taki układ ma niezerowe rozwiązanie dla A oraz B gdy wyznacznik

$$\begin{aligned} \det(J - \lambda I) &= \begin{vmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{vmatrix} = (a_{11} - \lambda)(a_{22} - \lambda) - a_{12}a_{21} = \\ &= \lambda^2 - (a_{11} + a_{22})\lambda + a_{11}a_{22} - a_{12}a_{21} = 0 \end{aligned} \quad (1.168)$$

Macierz I jest macierzą jednostkową, tzn. diagonalną 2×2 i wszystkie elementy diagonalne to 1. Z powyższej relacji otrzymujemy równanie kwadratowe dla nieznanych wartości własnych $\lambda = \lambda_1$ oraz $\lambda = \lambda_2$.

Rozwiązanie zlinearyzowanego układu jest w postaci kombinacji liniowej :

$$X(t) = A_1 e^{\lambda_1 t} + A_2 e^{\lambda_2 t}, \quad Y(t) = B_1 e^{\lambda_1 t} + B_2 e^{\lambda_2 t} \quad (1.169)$$

Pytanie o stabilność stanu stacjonarnego (x_s, y_s) jest równoważne pytaniu: kiedy

$$\lim_{t \rightarrow \infty} X(t) = 0 \quad \lim_{t \rightarrow \infty} Y(t) = 0 \quad (1.170)$$

Funkcje eksponencjalne dążą do zera jeżeli część rzeczywista wartości własnych macierzy Jacobiego λ_i są ujemne:

$$Re(\lambda_1) < 0, \quad Re(\lambda_2) < 0 \quad (1.171)$$

Wówczas stan stacjonarny jest asymptotycznie stabilny. Jeżeli

$$Re(\lambda_1) > 0, \quad \text{i/lub} \quad Re(\lambda_2) > 0 \quad (1.172)$$

to stan stacjonarny jest niestabilny. Jeżeli

$$Re(\lambda_1) = Re(\lambda_2) = 0 \quad (1.173)$$

to stan stacjonarny jest stabilny, ale nie jest asymptotycznie stabilny. Jeżeli wartości własne macierzy Jacobiego wynoszą zero, metoda linearyzacji nie rozstrzyga o stabilności.

Zamiast wyznaczyć wartości własne (λ_1, λ_2) tej macierzy, wystarczy sprawdzić, kiedy część rzeczywista wartości własnych jest ujemna (lub dodatnia). Ponieważ macierz Jacobiego jest macierzą 2×2 , więc otrzymujemy równanie kwadratowe dla λ . Aby wartości własne miały część rzeczywistą ujemną muszą zachodzić dwie relacje:

$$\lambda_1 + \lambda_2 = a_{11} + a_{22} < 0 \quad \text{oraz} \quad \lambda_1 \lambda_2 = a_{11}a_{22} - a_{12}a_{21} > 0, \quad (1.174)$$

to oznacza że

$$\text{Tr } J < 0, \quad \det J > 0 \quad (1.175)$$

gdzie Tr oznacza ślad macierzy, czyli sumę elementów diagonalnych macierzy oraz \det jest wyznacznikiem macierzy.

Przykład 1

Rozważmy układ równań różniczkowych

$$\begin{aligned}\dot{x} &= 1 - xy = f(x, y), \\ \dot{y} &= x - y^2 = g(x, y).\end{aligned}\tag{1.176}$$

Łatwo znaleźć stany stacjonarne

$$\begin{aligned}1 - xy &= 0, \\ x - y^2 &= 0.\end{aligned}\tag{1.177}$$

Z drugiego równania $x = y^2$ wstawiamy do pierwszego równania: $1 - y^3 = 0$ czyli $y^3 = 1$. Stąd $y = 1$ oraz $x = 1$. Otrzymujemy stan stacjonarny

$$(x_s, y_s) = (1, 1)\tag{1.178}$$

Obliczmy elementy macierzy Jacobiego

$$J = \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} \end{bmatrix} = \begin{bmatrix} -y & -x \\ 1 & -2y \end{bmatrix} = \begin{bmatrix} -1 & -1 \\ 1 & -2 \end{bmatrix} \quad \text{w punkcie } (x = 1, y = 1)\tag{1.179}$$

Ślad macierz J , czyli suma elementów diagonalnych $\text{Tr}J = -3 < 0$ oraz wyznacznik $\det(J) = 3 > 0$. Spełnione są relacje dla stany stacjonarnego który jest asymptotycznie stabilny. Wniosek: stan stacjonarny $(x_s, y_s) = (1, 1)$ jest asymptotycznie stabilny.

Można sprawdzić to, wyliczając explicite wartości własne macierzy Jacobiego:

$$\lambda_1 = \frac{1}{2}(-3 + i\sqrt{3}), \quad \lambda_2 = \frac{1}{2}(-3 - i\sqrt{3})\tag{1.180}$$

Ich części rzeczywiste są ujemne: $\Re(\lambda_1) = -3/2, \Re(\lambda_2) = -3/2$.

1.6.5 Przypadek C: Stabilność stanów stacjonarnych układów wielowymiarowych

Dla układu równań różniczkowych o wymiarze większym niż 2 stosujemy te same metody co dla układu 2 równań różniczkowych. Oczywiście istnieje więcej różnych przypadków i większe bogactwo własności stanów stacjonarnych. Nie będziemy przedstawiali tu przypadku o dowolnym wymiarze. Rozważymy przypadek 3 równań, aby pokazać podobieństwo do przypadku 2 równań. Analizujemy układ 3 równań różniczkowych:

$$\begin{aligned}\dot{x} &= F_1(x, y, z), & x(0) &= x_0, \\ \dot{y} &= F_2(x, y, z), & y(0) &= y_0, \\ \dot{z} &= F_3(x, y, z), & z(0) &= z_0.\end{aligned}\tag{1.181}$$

Stany stacjonarne są określone przez rozwiązania układu równań algebraicznych:

$$F_1(x, y, z) = 0, \quad F_2(x, y, z) = 0, \quad F_3(x, y, z) = 0\tag{1.182}$$

Z równań tych otrzymujemy stan stacjonarny $(x, y, z) = (x_s, y_s, z_s)$. Następnie obliczymy macierz Jacobiego

$$J = \begin{bmatrix} \frac{\partial F_1}{\partial x} & \frac{\partial F_1}{\partial y} & \frac{\partial F_1}{\partial z} \\ \frac{\partial F_2}{\partial x} & \frac{\partial F_2}{\partial y} & \frac{\partial F_2}{\partial z} \\ \frac{\partial F_3}{\partial x} & \frac{\partial F_3}{\partial y} & \frac{\partial F_3}{\partial z} \end{bmatrix} \quad (1.183)$$

Obliczymy macierz Jacobiego w punkcie stacjonarnym $J = J(x_s, y_s, z_s)$. W ten sposób otrzymujemy macierz liczbową. Kolejnym krokiem jest znalezienie wartości własnych $\lambda = \lambda_i (i = 1, 2, 3)$ tej macierzy, czyli rozwiązanie wielomianu 3-go stopnia dla λ :

$$\det(J - \lambda I) = \begin{vmatrix} \frac{\partial F_1}{\partial x} - \lambda & \frac{\partial F_1}{\partial y} & \frac{\partial F_1}{\partial z} \\ \frac{\partial F_2}{\partial x} & \frac{\partial F_2}{\partial y} - \lambda & \frac{\partial F_2}{\partial z} \\ \frac{\partial F_3}{\partial x} & \frac{\partial F_3}{\partial y} & \frac{\partial F_3}{\partial z} - \lambda \end{vmatrix} = 0 \quad (1.184)$$

Macierz I jest macierzą jednostkową, tzn. diagonalną 3×3 i wszystkie elementy diagonalne to 1.

Jeżeli części rzeczywiste wszystkich wartości własnych macierzy Jacobiego λ_i są ujemne:

$$\Re(\lambda_1) < 0, \quad \Re(\lambda_2) < 0, \quad \Re(\lambda_3) < 0 \quad (1.185)$$

to stan stacjonarny jest asymptotycznie stabilny. Jeżeli chociaż jedna z wartości własnych λ_i ma część rzeczywistą dodatnią

$$\Re(\lambda_1) > 0, \quad \text{i/lub} \quad \Re(\lambda_2) > 0, \quad \text{i/lub} \quad \Re(\lambda_3) > 0 \quad (1.186)$$

to stan stacjonarny jest niestabilny.

W przypadkach wielowymiarowych wygodnie jest stosować metody numeryczne. Obliczanie wartości własnych macierzy jest numerycznie zadaniem łatwym.

Przykład 2: Model Lorenza

$$\begin{aligned} \dot{x} &= \sigma(y - x) = F_1(x, y, z), & x(0) &= x_0, \\ \dot{y} &= x(\rho - z) - y = F_2(x, y, z), & y(0) &= y_0, \\ \dot{z} &= xy - \beta z = F_3(x, y, z), & z(0) &= z_0. \end{aligned} \quad (1.187)$$

gdzie wszystkie parametry są dodatnie: $\sigma, \rho, \beta > 0$.

Znajdujemy stany stacjonarne czyli rozwiązujemy układ równań algebraicznych

$$\begin{aligned} \sigma(y - x) &= 0, \\ x(\rho - z) - y &= 0, \\ xy - \beta z &= 0. \end{aligned} \quad (1.188)$$

Możemy posłużyć się programem w Sage, ale układ ten jest na tyle prosty, że możemy go rozwiązać sami. Z pierwszego równania wynika, że

$$y = x, \quad (1.189)$$

z trzeciego równania otrzymujemy

$$z = \frac{1}{\beta}x^2. \quad (1.190)$$

Teraz wstawiamy x oraz z do drugiego równania otrzymując relację:

$$x\rho - x\frac{1}{\beta}x^2 - x = 0, \quad (1.191)$$

czyli

$$x[\rho - 1 - \frac{1}{\beta}x^2] = 0. \quad (1.192)$$

Ta relacja jest prosta i wynika z niej że

$$x = x_1 = 0 \quad \text{lub} \quad x = x_2 = \sqrt{\beta(\rho - 1)} \quad \text{lub} \quad x = x_3 = -\sqrt{\beta(\rho - 1)} \quad (1.193)$$

Otrzymujemy 3 stany stacjonarne:

$$\begin{aligned} P_1 &= (x_1, y_1, z_1) = (0, 0, 0), \\ P_2 &= (x_2, y_2, z_2) = (\sqrt{\beta(\rho - 1)}, \sqrt{\beta(\rho - 1)}, \rho - 1), \\ P_3 &= (x_3, y_3, z_3) = (-\sqrt{\beta(\rho - 1)}, -\sqrt{\beta(\rho - 1)}, \rho - 1). \end{aligned} \quad (1.194)$$

Dla każdego stanu stacjonarnego musimy zbadać jego stabilność analizując zagadnienie własne dla macierzy Jacobiego. No to do dzieła...

```
#kilka zmiennych
var('x y z sigma rho beta alpha')
#i kilka zalozen
assume(sigma>0)
assume(rho>0)
assume(beta>0)
#definiujemy rownania rozniczkowe
F1 = sigma*(y-x)
F2 = x*(rho-z) - y
F3 = x*y - beta*z
#stany stacjonarne
rozv = solve([F1==0,F2==0,F3==0],[x,y,z])
P1 = rozv[2]
P2 = rozv[0]
P3 = rozv[1]
#macierz Jacobiego
J = matrix([
[diff(F1,x),diff(F1,y),diff(F1,z)],
[diff(F2,x),diff(F2,y),diff(F2,z)],
[diff(F3,x),diff(F3,y),diff(F3,z)]
])
#analiza stabilnosci
punkt='P1' # lub 'P2','P3'
##automatycznie
```



```

P = dict(zip(['P1', 'P2', 'P3'], [P1, P2, P3])) [punkt]
J = matrix([
[
diff(F1, x) (x=P[0].rhs(), y=P[1].rhs(), z=P[2].rhs()),
diff(F1, y) (x=P[0].rhs(), y=P[1].rhs(), z=P[2].rhs()),
diff(F1, z) (x=P[0].rhs(), y=P[1].rhs(), z=P[2].rhs())
],
[
diff(F2, x) (x=P[0].rhs(), y=P[1].rhs(), z=P[2].rhs()),
diff(F2, y) (x=P[0].rhs(), y=P[1].rhs(), z=P[2].rhs()),
diff(F2, z) (x=P[0].rhs(), y=P[1].rhs(), z=P[2].rhs())
],
[
diff(F3, x) (x=P[0].rhs(), y=P[1].rhs(), z=P[2].rhs()),
diff(F3, y) (x=P[0].rhs(), y=P[1].rhs(), z=P[2].rhs()),
diff(F3, z) (x=P[0].rhs(), y=P[1].rhs(), z=P[2].rhs())
]
])
#zagadnienie własne macierzy
dJ = det(J - alpha*matrix(3,3,1)) == 0
rozwdJ1 = solve(dJ, alpha)
b, s, r = 1, 2, 3
i, j = 0, 0
for a in rozwdJ1:
    j+=1
    buf = real(a.rhs() (rho=r, beta=b, sigma=s))
    if buf > 0: i += 1
print "Dla beta=%s, sigma=%s, rho=%s"%(b, s, r)
print "%s %s jest punktem stabilnym"%(punkt, "" if i == 0 else "nie ")

```

1.7 Atraktory

Atraktor A to taki zbiór w przestrzeni fazowej układu dynamicznego, że wiele trajektorii startujących nawet bardzo daleko od tego zbioru dąży w miarę upływu czasu do tego zbioru A . Najprościej jest to zrozumieć na przykładzie oscylatora tłumionego. Realizacją takiego oscylatora może być wahadło matematyczne czyli kulka na nieważkim pręcie zawieszona w jakimś środowisku (byleby nie w próżni). Na kulkę działa siła przyciągania ziemskiego i siła tarcia. Kulke wychylamy o dowolny kąt od pozycji pionowej i obserwujemy trajektorię kulki. Kulka wykonuje coraz to mniejsze wahania i po dostatecznie długim czasie zatrzymuje się w pozycji pionowej. Kulke możemy wychylać o dowolny kąt, nadawać jej dowolną prędkość, a i tak po pewnym czasie kulka zatrzyma się w pozycji pionowej, która odpowiada zerowemu wychyleniu kulki. Ten zerowy kąt wychylenia jest atraktorem. W tym przypadku atraktorem jest punkt w przestrzeni fazowej. Ponieważ przestrzeń fazowa oscylatora harmonicznego tłumionego jest 2-wymiarowa położenie-prędkość, atraktorem jest punkt $A = (pooenie = 0, prdko = 0)$.

Podamy teraz bardziej formalną definicję.

Atraktor Atraktorem nazywamy ograniczony zbiór w przestrzeni fazowej, do którego dążą

asymptotycznie w czasie (tzn. gdy $t \rightarrow \infty$) obszary warunków początkowych o niezerowej objętości w przestrzeni fazowej. Atraktor to inaczej zbiór przyciągający: przyciąga on trajektorie o różnych warunkach początkowych. Ale nie musi on przyciągać wszystkich trajektorii. Dla danego układu dynamicznego może istnieć wiele atraktorów, nawet nieskończenie wiele. Atraktory mogą mieć nieskomplikowaną strukturę: to może być punkt, kilka punktów, krzywa taka jak okrąg czy zdeformowana elipsa, część płaszczyzny, torus (podobny do dętki), część przestrzeni. Atraktory mogą też mieć skomplikowaną strukturę: może to być zbiór fraktalny, tzn. zbiór o niecałkowitym wymiarze, np. 0.63, 2.06. Taki atraktor nazywa się *dziwnym atraktorem*.

Z atraktorami związane są obszary przyciągania lub baseny przyciągania $B(A)$. Basenem przyciągania atraktora A nazywamy zbiór warunków początkowych x_0 , dla których trajektorie są przyciągane do A , czyli

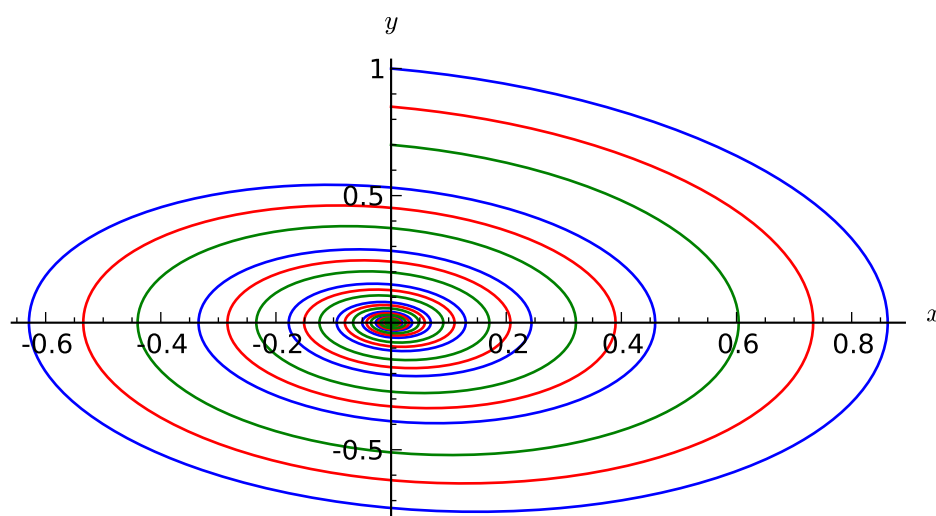
$$B(A) = \{x_0 : \lim_{t \rightarrow \infty} x(t; x_0) \in A\} \quad (1.195)$$

gdzie $x(t; x_0)$ jest trajektorią startującą z warunku początkowego x_0 , np. rozwiązaniem układu równań różniczkowych z odpowiednimi warunkami początkowymi \vec{x}_0 .

1.7.1 Przykład 1: oscylator harmoniczny tłumiony

Jest tylko jeden atraktor: to punkt $(0, 0)$. Basenem przyciągania jest cała płaszczyzna fazowa.

```
var('x,y')
T = srange(0,50,0.01)
de = vector([y,-x-0.2*y])
sol1 = desolve_odeint(de,[0,1],T,[x,y]).tolist()
sol2 = desolve_odeint(de,[0,0.85],T,[x,y]).tolist()
sol3 = desolve_odeint(de,[0,0.7],T,[x,y]).tolist()
p = list_plot(sol1) +\
    list_plot(sol2) +\
    list_plot(sol3)
```

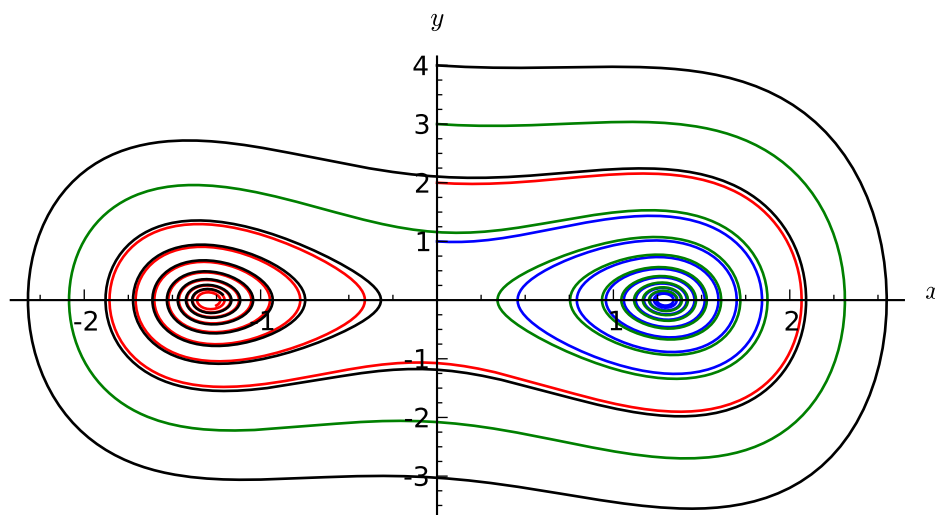


Rysunek 1.16: Trajektorie zbiegają się do jednego basenu przyciągania $(x, y) = (0, 0)$.

1.7.2 Przykład 2: oscylator nieliniowy (bistabilny) tłumiony

Są dwa atraktory: punkt $(-x_s, 0)$ oraz symetryczny punkt $(x_s, 0)$, gdzie x_s jest stanem stacjonarnym. Płaszczyzna fazowa dzieli się na 2 baseny przyciągania, które są “pasiastymi” naprzemiennymi zbiorami. Przejrzysta wizualizacja jest opracowana na [naszej stronie internetowej](#).

```
var('x,y')
T = srange(0,30,0.01)
de = vector([y, 2*x-1.2*x^3 -0.2*y])
sol1=desolve_odeint(de, [0,1], T, [x,y])
sol2=desolve_odeint(de, [0,2], T, [x,y])
sol3=desolve_odeint(de, [0,3], T, [x,y])
sol4=desolve_odeint(de, [0,4], T, [x,y])
g = list_plot(sol1.tolist()) +\
    list_plot(sol2.tolist()) +\
    list_plot(sol3.tolist()) +\
    list_plot(sol4.tolist())
```



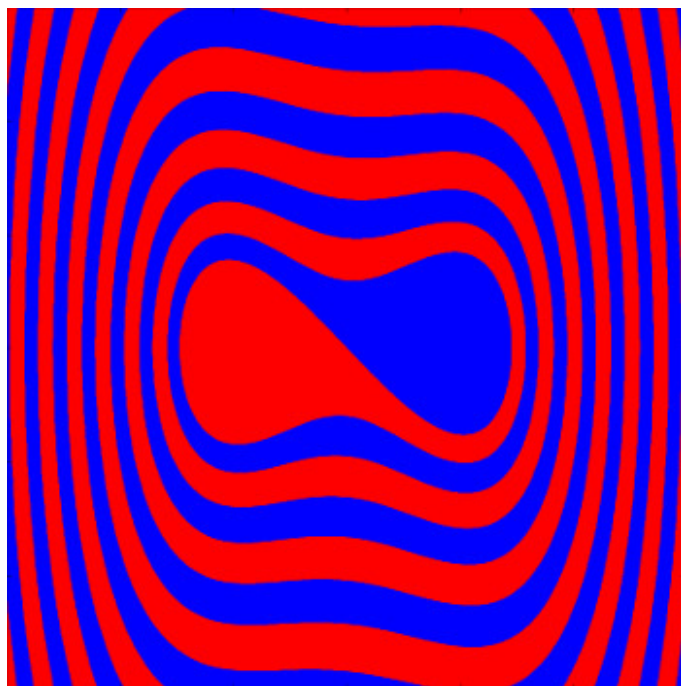
Rysunek 1.17: Trajektorie zbiegają się do dwóch basenów przyciągania usytuowanych w dwóch minimach potencjału bistabilnego $(-x_s, 0)$ oraz $(x_s, 0)$.

Jeżeli warunek początkowy leży w obszarze o kolorze niebieskim, trajektorie dążą do prawego minimum potencjału; dokładniej do stanu stacjonarnego $(x_s, 0)$ który jest atraktorem. Jeżeli warunek początkowy leży w obszarze o kolorze czerwonym, trajektorie dążą do lewego minimum potencjału; dokładniej do stanu stacjonarnego $(-x_s, 0)$ który jest atraktorem.

1.7.3 Przykład 3: Cykl graniczny

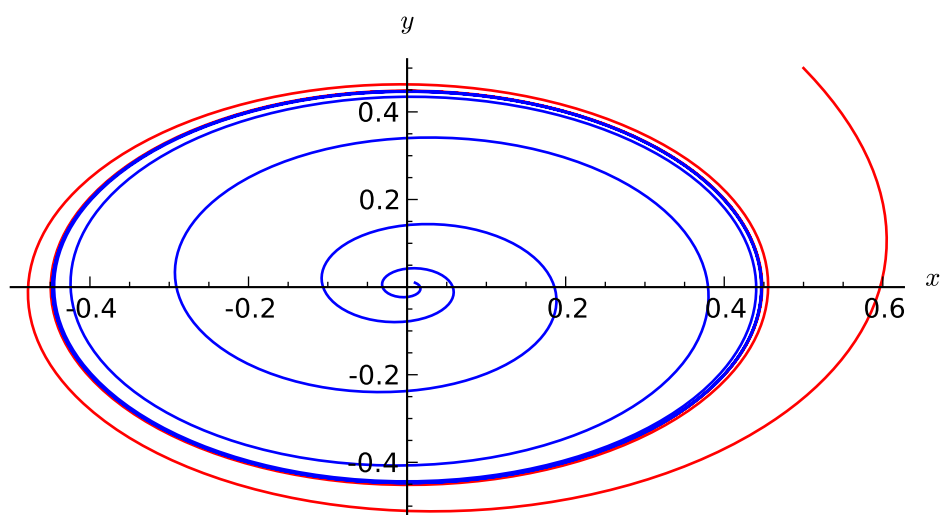
Atraktorem jest krzywa zamknięta (okrąg, elipsa, inne dowolne krzywe zamknięte). Poniżej przedstawiamy dwa przykłady zaczerpnięte z modeli biologicznych.

Przykład 1



Rysunek 1.18: Dwa obszary przyciągania dla układu bistabilnego

```
var('x,y')
T = srange(0,50,0.01)
de1 = y + x*(0.2 - (x^2 + y^2))
de2 = -x + y*(0.2 - (x^2 + y^2))
s1 = desolve_odeint(vector([de1,de2]),[0.5,0.5],T,[x,y])
s2 = desolve_odeint(vector([de1,de2]),[0.01, 0.01],T,[x,y])
p1 = list_plot(s1.tolist())
p2 = list_plot(s2.tolist())
```



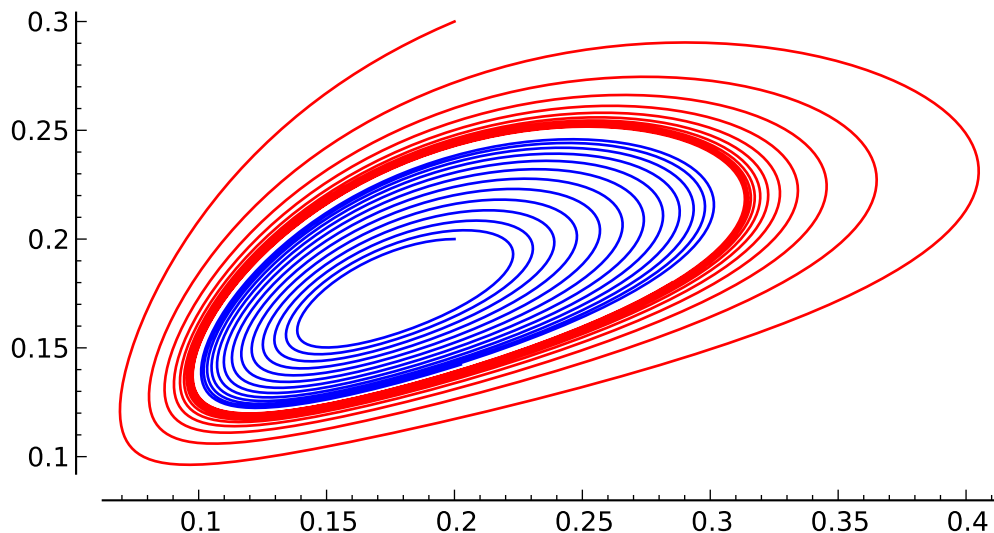
Rysunek 1.19: Cykl graniczny.

Przykład 2

```

var('x,y')
a, b, d = 1.3, 0.33, 0.1
F(x,y) = x*(1-x) - a*x*y/(x+d)
G(x,y) = b*y*(1-y/x)
T = srange(0,80,0.01)
sl1 = desolve_odeint(vector([F,G]), [0.2,0.3], T, [x,y])
sl2 = desolve_odeint(vector([F,G]), [0.2,0.2], T, [x,y])
j1 = list_plot(sl1.tolist())
j2 = list_plot(sl2.tolist())

```



Rysunek 1.20: Cykl graniczny.

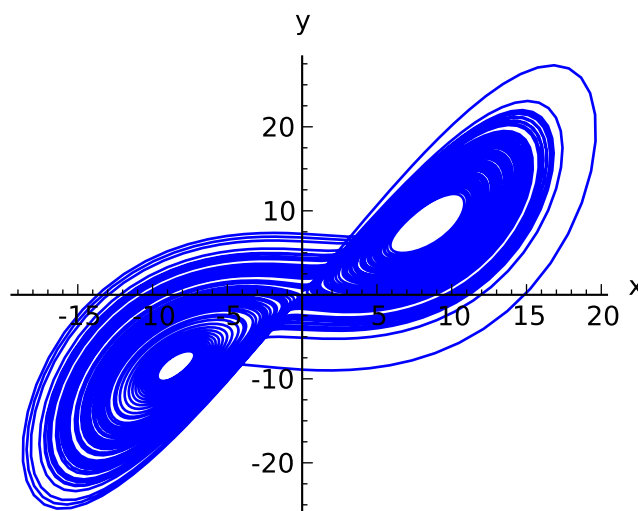
1.7.4 Przykład 4: Atraktor Lorenza

Jest to przykład tak zwanego dziwnego atraktora. Najprostsza jego definicja to taka, że posiada on strukturę fraktala. O układzie Lorenza generującym ten fraktal można poczytać w poprzednim rozdziale tego skryptu, traktującym o stanach stacjonarnych.

```

var('x y z')
rho, sigma, beta = 28, 10, 8/3
F1 = sigma*(y-x)
F2 = x*(rho-z) - y
F3 = x*y - beta*z
de = vector([F1,F2,F3])
T = srange(0,100,0.01)
atraktor_lorenza = desolve_odeint(de, [0,0.5,1], T, [x,y,z])
p2d = list_plot(zip(atraktor_lorenza[:,0],atraktor_lorenza[:,1]))
p3d = list_plot(atraktor_lorenza.tolist(), viewer='tachyon')

```



Rysunek 1.21: Atraktor Lorenza.

1.8 Dyskretne układy dynamiczne

We Wprowadzeniu pokazaliśmy w jaki sposób mogą pojawiać się zagadnienia modelowane równaniami rekurencyjnymi

$$x_{n+1} = f(x_n), \quad x_0 = c \quad (1.196)$$

Z jednej strony mogą to być przybliżenia równań różniczkowych, gdy pochodną aproksymujemy skończonym ilorazem różniczkowym. Z drugiej strony, mogą to być układy dynamiczne w których czas jest dyskretny i jest mierzony co pewien ustalony interwał czasowy (co sekundę, co godzinę, co dzień, co miesiąc, itd.). Tego typu modele spotykamy w biologii. Kolejnym przykładem jest algorytm Newtona pozwalający rekurencyjnie znajdować pierwiastki równania

$$F(x) = 0 \quad (1.197)$$

W metodzie Newtona kolejne przybliżenie pierwiastka dane jest przez równanie

$$x_{n+1} = x_n + \frac{F(x_n)}{F'(x_n)} \equiv f(x_n) \quad (1.198)$$

gdzie zakładamy, że pochodna $F'(x)$ istnieje (w przeciwnym przypadku metoda ta nie może być stosowana). Oznacza to, że nie możemy wpaść w ekstremum funkcji $F(x)$. W szczególności dla $F(x) = x^2 - 2$ ciąg (1.198) w zależności od warunku początkowego zmierza do liczby $x_+ = \sqrt{2}$ lub $x_- = -\sqrt{2}$ (co jest oczywiste skądinąd).

Wizualizacja iteracji Używając komputera, kolejne wyrazy x_n jest łatwo uzyskać. Dla przykładu, niech

$$x_{n+1} = x_n - 0.5x_n^2 + 1.7, \quad x_0 = 1 \quad (1.199)$$

Poniżej prezentujemy program w SAGE:

```
# warunek początkowy
x0 = 1
# tu można wstawić dowolną funkcję f(x)
f(x) = x - 0.5*x^2 + 1.7
for i in range(20):
    print x0
    x0 = f(x0)
```

który wyprodukuje iteracyjnie następujące liczby

```
1
2.2000000000000000
1.4800000000000000
2.0848000000000000
1.6116044800000000
2.01296998002196
1.68694590978715
1.96405265851335
1.73530123580667
1.92966604631059
1.76786052116862
1.90519511001533
1.79031090640217
1.88770433561089
1.80599050626881
1.87518965190227
1.81702153660159
1.86623790436459
1.82481594652102
1.85983932718231
```

Poniżej jest prezentacja graficzna dla ogólniejszego przypadku

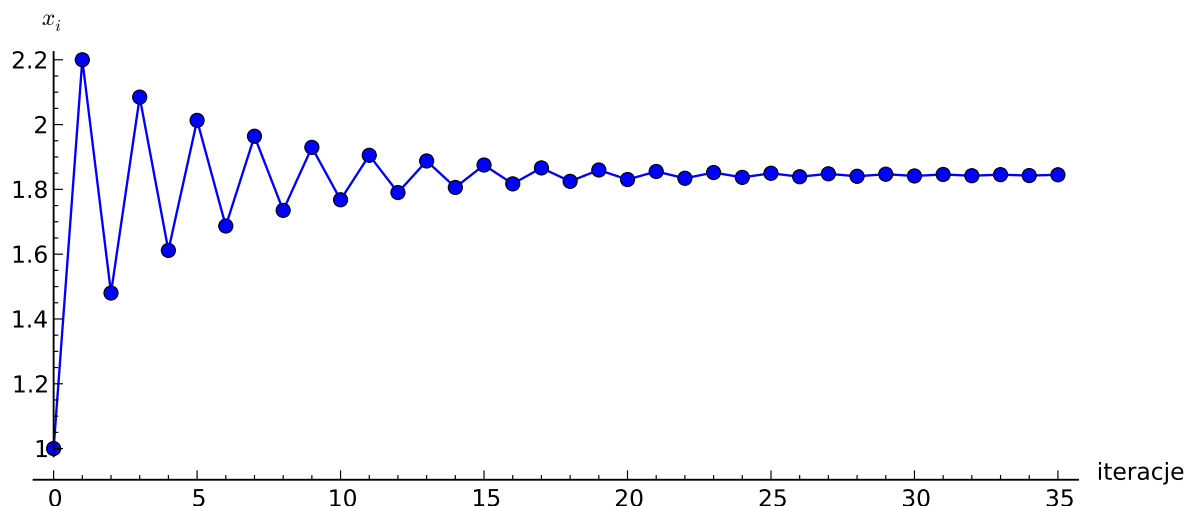
$$x_{n+1} = x_n - bx_n^2 + 1.7, \quad x_0 = 1 \quad (1.200)$$

z dowolnym parametrem b . Zachęcamy do zabawy: zmieniajcie wartość parametru b oraz warunek początkowy. Czytelnik łatwo może podać swoje własne równanie, zmieniając odpowiednie wyrażenia w programie.

```
def ne(b,X):
    return X - b*X*X + 1.7

def pophis(startp,b,length):
    his = [startp]
    for i in range(length):
        his.append(ne(b,his[i]))
    return his

@interact
def _(b=slider(0.05,3,0.05,default=0.5,label='Factor b')):
    show(list_plot(pophis(1,b,35),plotjoined=True,marker='o'))
```

Rysunek 1.22: Graficzna prezentacja równania (1.200), dla $b = 0.5$.

1.8.1 Metoda pajęczyny

Metoda pajęczyny pozwala na wykresie śledzić własności równań rekurencyjnych. Za pomocą tego wykresu można obserwować kolejne kroki iteracji.

Sposób rysowania:

1. Dla danego równania $x_{n+1} = f(x_n)$, rysujemy wykres funkcji $y = f(x)$ oraz linię prostą $y = x$. Prosta ta pozwala przenosić wartość x_{n+1} z osi OY na oś OX.
1. Na osi OX zaznaczamy warunek początkowy x_0 . Znajdujemy graficznie punkt $x_1 = f(x_0)$, który jest na osi OY.
1. Przy pomocy prostej $y = x$ przenosimy teraz punkt x_1 na oś OX.
2. Mając na osi OX punkt x_1 , traktujemy go jako następny warunek początkowy i znajdujemy graficznie na osi OY punkt $x_2 = f(x_1)$.
1. Powtarzamy kroki od 2,3,4 z ostatnim otrzymanym punktem jako początkowym x_0 .

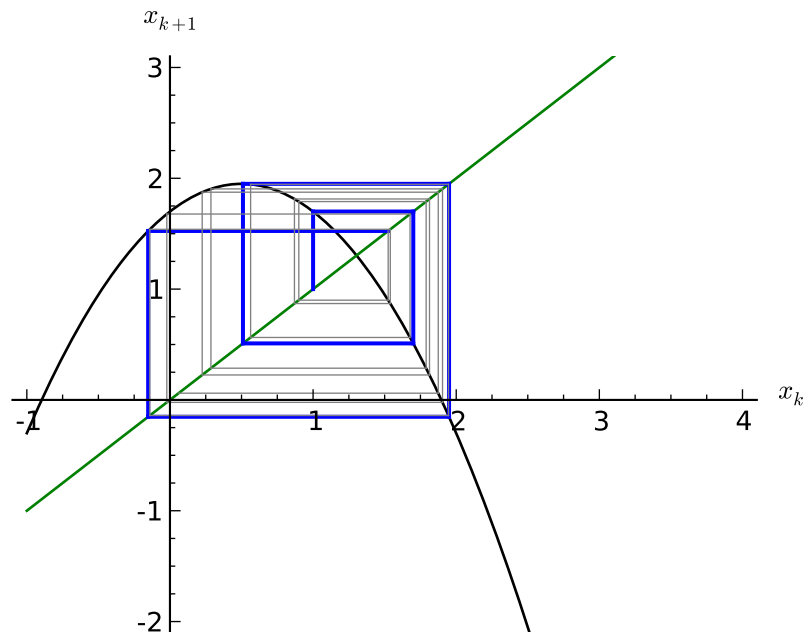
```
var('r,x0')
@interact
def cobweb(a=slider(0.4,1.4,0.1,default=1),
           x0=slider(0,1,0.1,default=1)):
    f(x) = x - a*x**2 + 1.7
    p = plot(f(x)==0, (x,-1,3), color='black')
    p += plot(x, (x,-1,4), color='green', figsize=5)
    for k in range(20):
        th = 1
        if k>45:
            th, color = 1.5, 'red'
        elif k < 5:
            th, color = 1.5, 'blue'
```



```

else:
    th, color = 0.5, 'grey'
    l1 = line([(x0, x0), (x0, f(x0))], color=color, thickness=th)
    l2 = line([(x0, f(x0)), (f(x0), f(x0))], color=color, thickness=th)
    p += l1 + l2
    x0 = f(x0)
p.axes_labels(["$x_k$", "$x_{k+1}$"])
p.show()

```



Rysunek 1.23: Przykład wykorzystania metody pajęczyny dla (1.200) dla domyślnych parametrów $a = 1, x_0 = 1$ z `interact`.

1.8.2 Stany stacjonarne (punkty stałe) o okresie 1

Układ dynamiczny z czasem dyskretnym ma postać równania rekurencyjnego

$$x_{n+1} = f(x_n), \quad x_0 \text{ znany} \quad (1.201)$$

Wartość funkcji w następnym kroku x_{n+1} jest obliczana z wartości funkcji x_n w kroku poprzednim.

Z punktu widzenia modelowania, chcielibyśmy wiedzieć, czy układ w trakcie ewolucji dąży do jakiegoś stanu stacjonarnego i czy ten stan stacjonarny jest stabilny. Pojawia się też pytanie, jakiego typu stany stacjonarne mogą pojawiać się dla układów modelowanych równaniami rekurencyjnymi.

A. Stany stacjonarne (punkty stałe o okresie 1)

Przypomnijmy sobie, co oznacza istnienie stanu stacjonarnego dla układów modelowanych równaniem różniczkowym:

$$\frac{dx(t)}{dt} = F(x(t)) \quad (1.202)$$

Stan stacjonarny to taki stan, który nie zmienia się w trakcie ewolucji, nie zmienia się wraz ze zmianą czasu, czyli $x(t) = x_s$ jest wielkością stałą, niezmienną. Skoro tak, to

$$\frac{dx(t)}{dt} = \frac{dx_s}{dt} = 0 \quad (1.203)$$

Aby lewa strona równania różniczkowego była równa prawej stronie, musi zachodzić równość:

$$F(x_s) = 0 \quad (1.204)$$

To jest równanie (warunek), z którego wyznaczamy stan stacjonarny.

Podobnie rozumiemy w przypadku równań rekurencyjnych: Stan stacjonarny to taki stan, który nie zmienia się w trakcie ewolucji (dyskretnej), czyli z punktu x_s otrzymujemy znowu x_s . Dla równania dyskretnego ta niezmiennosc oznacza, że

$$\text{jeżeli } x_n = x_s \text{ to } x_{n+1} = x_s \quad (1.205)$$

i równanie rekurencyjne

$$x_{n+1} = f(x_n) \text{ ma postać } x_s = f(x_s) \quad (1.206)$$

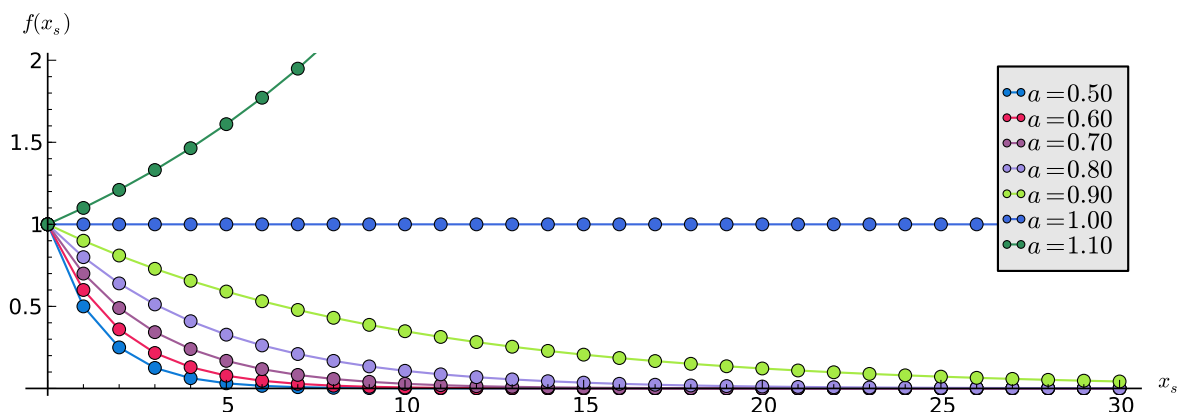
Stąd otrzymujemy równanie dla stanu stacjonarnego

$$x_s = f(x_s) \text{ lub w uproszczonym zapisie } x = f(x) \quad (1.207)$$

Zapamiętajmy to równanie! Matematycy zamiast nazwy “stan stacjonarny” stosują nazwę “punkt stały odwzorowania” lub “punkt stały o okresie 1”.

Poniżej przedstawiamy to graficznie. Po wielu krokach iteracji obserwujemy powtarzanie się tej samej wartości: kolejna iteracji już nic nie zmienia, jest ta sama, stała.

```
def newpop(a, prevpop):  
    return a * prevpop  
  
def p_history(startpop, a, length):  
    history = [startpop]  
    for i in range(length):  
        history.append(newpop(a, history[i]))  
    return history  
  
@interact  
def _(a=slider(0.5, 1.1, 0.05, default=0.5, label='Malthus Factor a')):  
    myplot=list_plot(p_history(1, a, 30), plotjoined=True, marker='o')  
    myplot.show()
```



Rysunek 1.24: Punkt stały odwzorowania.

Zobaczmy jak to “działa” w przypadku tzw. równania logistycznego gdy $f(x) = ax(1 - x)$:

$$x = f(x) \quad \text{oznacza} \quad x = ax = ax - ax^2 \quad (1.208)$$

Jest to równanie kwadratowe:

$$ax^2 - ax + x = 0 \quad \text{czyli} \quad ax^2 + x = x[ax + 1] = 0 \quad (1.209)$$

Otrzymujemy dwa rozwiązania

$$x_1 = 0 \quad \text{oraz} \quad x_2 = \frac{a - 1}{a} = 1 - \frac{1}{a} \quad (1.210)$$

Są to dwa stany stacjonarne układu. Występuje tu podobieństwo do ciągłego modelu Verhulsta, które też posiada dwa stany stacjonarne $N_1 = 0$ (niestabilne) oraz $N_2 = K$ (stabilne).

Który ze stanów stacjonarnych x_1 i x_2 jest stabilny, a który niestabilny. Zbadamy teraz ten problem.

B. Stabilność stanów stacjonarnych

Stabilność stanu stacjonarnego x_s jest podobnie definiowana jak w przypadku równań różniczkowych zwyczajnych: dowolnie małe odchylenie od stanu stacjonarnego jest w trakcie ewolucji redukowane i odchylenie dąży do zera. Innymi słowy: jeżeli stan początkowy x_0 mało różni się od stanu stacjonarnego x_s , to ciąg liczb:

$$x_0, \quad x_1 = f(x_0), \quad x_2 = f(x_1), \quad x_3 = f(x_2), \quad x_4 = f(x_3), \dots \quad (1.211)$$

dąży do stanu stacjonarnego x_s .

Innymi słowy, jeżeli $|x_0 - x_s| < \delta$ dla odpowiednio małej liczby δ to $|x_n - x_s| \rightarrow 0$ gdy $n \rightarrow \infty$. Wyprowadzimy kryterium które pozwala badać stabilność x_s .

Wprowadzamy oznaczenie dla małego odchylenia od stanu stacjonarnego:

$$y_n = x_n - x_s \ll 1 \quad (1.212)$$

stąd

$$x_n = x_s + y_n. \quad (1.213)$$

Wówczas

$$y_{n+1} = x_{n+1} - x_s = f(x_n) - x_s = f(x_s + y_n) - x_s \quad (1.214)$$

Ponieważ y_n jest małą wielkością, to funkcję $f(x_s + y_n)$ rozwijamy w szereg Taylora:

$$\begin{aligned} y_{n+1} = f(x_s + y_n) - x_s &\approx [f(x_s) + f'(x_s)y_n + \dots] - x_s = \\ &= f(x_s) - x_s + f'(x_s)y_n + \dots = \lambda y_n + \dots, \end{aligned} \quad (1.215)$$

gdzie

$$\lambda = f'(x_s) \quad (1.216)$$

jest liczbą. Wykorzystaliśmy tu równość dla stanu stacjonarnego: $x_s = f(x_s)$. W ten sposób otrzymaliśmy równanie rekurencyjne dla odchylenia od stanu stacjonarnego:

$$y_{n+1} = \lambda y_n \quad (1.217)$$

Jeżeli $y_n \rightarrow 0$ to stan stacjonarny x_s jest stabilny. Musimy teraz zbadać, dla jakich wartości liczby $\lambda = f'(x_s)$ równanie dla y_n ma rozwiązania dążące do zera dla warunku początkowego $|y_0| < 1$.

Rozpatrzmy 4 przypadki:

1. Przypadek $\lambda > 1$

Aby łatwiej zrozumieć, przyjmijmy $\lambda = 2$. Wówczas równanie na odchylenie od stanu stacjonarnego ma postać

$$y_{n+1} = 2y_n \quad (1.218)$$

Otrzymamy ciąg liczbowy

$$\begin{aligned} y_0, \\ y_1 = 2y_0, \\ y_2 = 2y_1 = 2 \times 2y_0 = 2^2 y_0, \\ y_3 = 2y_2 = 2^3 y_0, \\ y_4 = 2y_3 = 2^4 y_0, \\ \dots \end{aligned} \quad (1.219)$$

Widzimy, że kolejne liczby rosną, ponieważ są mnożone przez czynnik 2 i ciąg liczbowy jest rozbieżny. Otrzymujemy stąd wniosek, że dla $\lambda > 1$, stan stacjonarny nie jest stabilny: nieskończenie małe odchylenie od wartości stacjonarnej rośnie wraz z kolejnym krokiem iteracji.

2. Przypadek $\lambda < -1$

Aby łatwiej zrozumieć, przyjmijmy $\lambda = -2$. Wówczas równanie na odchylenie od stanu stacjonarnego ma postać

$$y_{n+1} = -2y_n \quad (1.220)$$

Otrzymamy ciąg liczbowy

$$\begin{aligned} y_0, \\ y_1 &= -2y_0, \\ y_2 &= -2y_1 = (-2) \times (-2)y_0 = 2^2y_0, \\ y_3 &= (-2)y_2 = -2^3y_0, \\ y_4 &= (-2)y_3 = 2^4y_0, \\ &\dots \end{aligned} \quad (1.221)$$

Widzimy, że wartości bezwzględne kolejnych liczby rosną, ponieważ są mnożone przez czynnik -2 i ciąg liczbowy jest rozbieżny. Otrzymujemy stąd wniosek, że dla $\lambda < -1$, stan stacjonarny nie jest stabilny: nieskończenie małe odchylenie od wartości stacjonarnej rośnie wraz z kolejnym krokiem iteracji.

3. Przypadek $\lambda \in (-1, 1)$

Aby łatwiej zrozumieć, przyjmijmy $\lambda = (1/2)$. Wówczas równanie na odchylenie od stanu stacjonarnego ma postać

$$y_{n+1} = \frac{1}{2}y_n \quad (1.222)$$

Otrzymamy ciąg liczbowy

$$\begin{aligned} y_0, \\ y_1 &= \frac{1}{2}y_0, \\ y_2 &= \frac{1}{2}y_1 = \frac{1}{2} \times \frac{1}{2}y_0 = \frac{1}{2^2}y_0, \\ y_3 &= \frac{1}{2}y_2 = \frac{1}{2^3}y_0, \\ y_4 &= \frac{1}{2}y_3 = \frac{1}{2^4}y_0, \\ &\dots \end{aligned} \quad (1.223)$$

Widzimy, że kolejne liczby maleją, ponieważ są mnożone przez czynnik $1/2$ i ciąg liczbowy dąży do zera. Otrzymujemy stąd wniosek, że dla $\lambda \in (-1, 1)$, stan stacjonarny jest stabilny: nieskończenie małe odchylenie od wartości stacjonarnej maleje do zera wraz z kolejnym krokiem iteracji.

Wniosek z tego jest następujący:

Uwaga: Stan stacjonarny x_s jest stabilny jeżeli $\lambda = f'(x_s) \in (-1, 1)$.

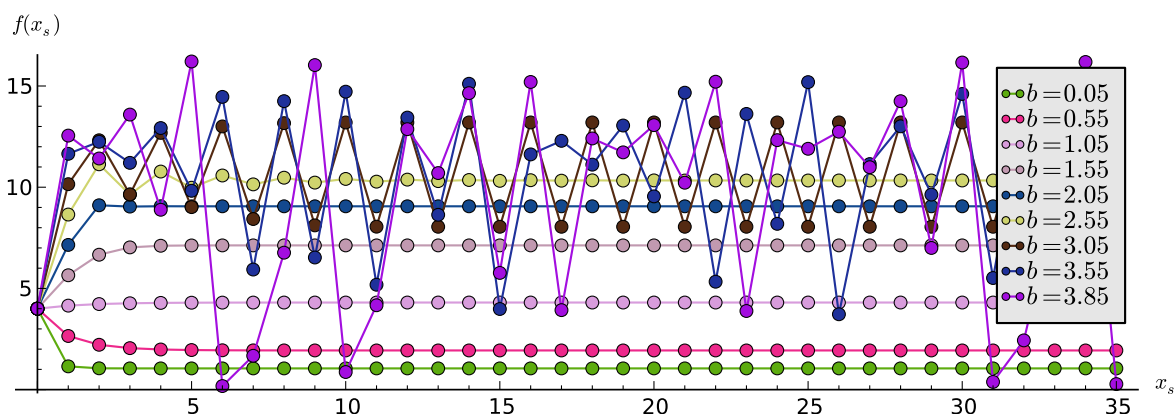
1.8.3 Stany stacjonarne o okresie 2

Jeżeli dla długich iteracji (formalnie dla $n \rightarrow \infty$) otrzymujemy ciągle tę samą wartość, to mówimy o stanie stacjonarnym o okresie 1. Pokazano to na powyższym rysunku. Jednakże mogą pojawić się inne stany stacjonarne. Dla przykładu, zobaczmy jak zachowuje się układ dla $n \rightarrow \infty$, który jest pokazany poniżej.

```
def ne(b,pre):
    return 1 + b*pre*(1-(1/16)*pre)

def pophis(startp,b,length):
    his = [startp]
    for i in range(length):
        his.append(ne(b,his[i]))
    return his

@interact
def _(b=slider(0.05,3.8,0.05,default=3.05,label='Factor b')):
    p = list_plot(pophis(4,b,35),plotjoined=True,marker='o')
    p.show(figsize=[8,3],axes_labels=[r'$x_s$', r'$f(x_s)$'])
```



Rysunek 1.25: Stany stacjonarne o okresie 2.

Obserwujemy, że dla dużych wartości iteracji $n \gg 1$, dwie wartości iteracji powtarzają się i układ “skacze” pomiędzy dwoma stanami. Mówimy wówczas o stanie stacjonarnym o okresie 2. Możemy także powiedzieć, że jest to stan periodyczny. Jak wyznaczyć takie stany? Skorzystamy ze wzoru na kolejne iteracje:

$$x_{n+2} = f(x_{n+1}), \quad x_{n+1} = f(x_n) \quad (1.224)$$

Należy zauważyć, że wartość $x_{n+2} = x_n$ ponieważ co drugi krok jest ten sam stan. Dlatego też

$$x_{n+2} = f(x_{n+1}) = f[f(x_n)] = x_n \quad (1.225)$$

Jeżeli oznaczymy stan stacjonarny x_s o okresie 2 jako x^* to powyższe równanie przepisujemy jako

$$f[f(x^*)] = x^* \quad (1.226)$$

Ale to samo zachodzi dla stanu $x_{n+3} = x_{n+1}$. Dlatego równanie powyższe ma 2 rozwiązania $x^* = x_1^*$ oraz $x^* = x_2^*$.

W praktyce rozwiązujemy równanie w postaci bardziej przyjaznej:

$$x = f[f(x)] \equiv g(x) \quad (1.227)$$

Pamiętajmy, że jest to złożenie 2 funkcji (co prawda takich samych funkcji, ale to jest drugorzędne).

To jest bardzo ważne równanie!

Napisanie tego równania w jawnej postaci nastrocza duże kłopoty przeciętnemu studentowi. Dlatego podamy 1 przykład. Niech układ dynamiczny będzie opisany równaniem

$$x_{n+1} = 1 - 2x_n^2, \quad f(x_n) = 1 - 2x_n^2 \quad \text{czyli} \quad f(x) = 1 - 2x^2 \quad (1.228)$$

Ile wynosi $f[f(x)]$? Obliczamy

$$f[f(x)] = 1 - 2[f(x)]^2 = 1 - 2[1 - 2x^2]^2 = 1 - 2[1 - 4x^2 + 4x^4] = -8x^4 + 8x^2 - 1 \quad (1.229)$$

Dlatego równanie które określa stan stacjonarny o okresie 2 ma postać

$$x = -8x^4 + 8x^2 - 1 = g(x) \quad (1.230)$$

Jest to wielomian 4-go stopnia.

Stabilność stanów stacjonarnych o okresie 2

Badanie stabilności stanów o okresie 2 jest analogiczne do badania stabilności stanów stacjonarnych o okresie 1. Jeżeli stan stacjonarny jest określony przez równanie

$$x^* = g(x^*) \quad (1.231)$$

to stan ten jest stabilny gdy

$$\lambda_2 = g'(x^*) \in (-1, 1) \quad (1.232)$$

Ponieważ funkcja $g(x)$ jest funkcją złożoną, więc należy stosować reguły różniczkowania funkcji złożonej:

$$\lambda_2 = g'(x^*) = [\{f[f(x)]\}']_{x=x^*} = \left[\frac{df(u)}{du} \right]_{u=f(x^*)} \left(\frac{df(x)}{dx} \right)_{x=x^*} \in (-1, 1) \quad (1.233)$$

Powyższy związek można przepisać w postaci:

$$\lambda_2 = f'(u^*)f'(x^*) \in (-1, 1) \quad (1.234)$$

gdzie $u^* = f(x^*)$

Warto coś tu zauważyć i uprościć. Ponieważ stan x^* jest o okresie 2 to jak nadmieniliśmy powyżej, faktycznie są 2 stany: $x^* = x_1^*$ (np. górny) oraz $x^* = x_2^*$ (np. dolny), co widać

doskonale z powyższego rysunku. Dlatego też dolny stan przechodzi w górny i następnie górny stan przechodzi w dolny. Można to zapisać jako:

$$x_2^* = f(x_1^*), \quad x_1^* = f(x_2^*) \quad (1.235)$$

Stąd otrzymujemy relację:

$$\lambda_2(x_1^*) = f'(x_2^*)f'(x_1^*) \in (-1, 1) \quad (1.236)$$

Podobna relacja zachodzi dla drugiego stanu

$$\lambda_2(x_2^*) = f'(x_1^*)f'(x_2^*) \in (-1, 1) \quad (1.237)$$

Jest to ta sama relacja co dla x_1^* . Więc oznacza to, że wystarczy zbadać wielkość $f'(x_1)f'(x_2)$, gdzie x_1 oraz x_2 to dwa rozwiązania równania $x = f(f(x))$.

Podobnie można badać punkty stałe o dowolnym okresie n :

$$x^* = f\{f[\dots f(x^*)\dots]\} = G(x^*) \quad (1.238)$$

gdzie mamy n -krotne złożenie odwzorowania f .

Stan x^* jest stabilny gdy

$$\lambda_n = G'(x^*) \in (-1, 1) \quad (1.239)$$

Ponieważ funkcja $G(x)$ jest funkcją złożoną, więc należy stosować reguły różniczkowania funkcji złożonej, podobnie jak to jest pokazane powyżej w przypadku $n = 2$.

Chaos w układach dynamicznych

2.1 Determinizm a przewidywalność

Determinizm jest terminem wieloznacznym. W odniesieniu do nauk przyrodniczych możemy mówić o tym że :

Każde zjawisko jest wyznaczone przez swoje przyczyny i całokształt warunków. Każde zjawisko podlega prawidłowościom przyrody, które wyrażane są w postaci odpowiednich praw nauki. Jeżeli dysponujemy odpowiednią wiedzą (o przyczynach, prawidłowościach), to można (przynajmniej w zasadzie) przewidywać przyszły bieg zdarzeń. Poznawcze znaczenie zasady przyczynowości sprowadza się do możliwości przewidywania. Determinizm i indeterminizm występują już w starożytnej filozofii przyrody. Demokryt twierdził, że nie ma w świecie zdarzeń przypadkowych, lecz „wszystko dzieje się wskutek konieczności”. Determinizm łączono z poglądem, że prawa przyrody mają charakter jednoznaczny. Okazało się jednak że wiele zjawisk podlega jedynie prawom statystycznym. Pogląd, że pewne zjawiska przyrody nie podlegają prawom jednoznacznym, ale jedynie statystycznym to indeterminizm. Doniosła rola praw statystycznych w fizyce współczesnej wymaga rozszerzenia formuły determinizmu: każde zdarzenie podlega prawom przyrody jednoznacznym (determinizm jednoznaczny, mocny) bądź statystycznym (determinizm probabilistyczny, umiarkowany). Wówczas indeterminizm (umiarkowany) mieści się w ramach determinizmu w szerszym ujęciu. Wydawało się, że w mechanice, stworzonej przez Galileusza i Newtona, znając położenie i prędkość (lub położenie i pęd) ciała materialnego można przewidzieć cały przyszły ruch tego ciała (a więc podać położenie i pęd w każdej chwili późniejszej), a także ustalić, jaki był ruch tego ciała w przeszłości. Wszystko to jest słuszne przy założeniu, że znane są wszystkie siły działające na ciało w przeszłości i w przyszłości. Wynikało to z twierdzeń matematycznych o jednoznaczności rozwiązań równań różniczkowych, a takimi są równania Newtona.

Jednakże historia pokazała, że w fizyce klasycznej są znane teorie statystyczne, niedeterministyczne. Rozważmy dla przykładu gaz zamknięty w jakimś naczyniu. Wiemy, że w warunkach równowagi termodynamicznej dwie równe co do objętości części tego naczynia będą zawierać jednakową liczbę cząsteczek gazu. Nie wiemy jednak, które cząsteczki znajdą się w której z dwu połówek naczynia. Sytuacja pozornie przypomina prawo rozpadu: można podać taki czas, w którym rozpadnie się połowa atomów, i w ten sposób podzielić wszystkie atomy na dwie równe części - te, które się w tym czasie rozpadną, i te, które się nie rozpadną, podobnie

jak podzieliliśmy cząsteczki gazu według kryterium, w której połowie naczynia się znajdują. W klasycznej fizyce statystycznej znamy prawa rządzące zachowaniem pojedynczych cząstek (są nimi z założenia prawa mechaniki newtonowskiej), a nasza niewiedza co do tego zachowania jest spowodowana po pierwsze niemożliwością śledzenia ruchu wielu miliardów obiektów, a po drugie, brakiem potrzeby, aby to czynić: wystarczy nam znać właśnie tylko pewne wielkości średnie, które ujawniają się fenomenologicznie, na przykład jako temperatura gazu, czy też jego ciśnienie. Tak więc rzeczywisty kompletny opis stanu gazu musiałby zawierać informację dotyczącą N wektorów położenia i N wektorów pędu (N - liczba cząsteczek gazu), co jest liczbą ogromną, podczas gdy opis statystyczny ogranicza się do kilku potrzebnych liczb. Opis statystyczny odnosi się do ogromnej liczby cząstek i jest to opis oparty o teorię prawdopodobieństwa i teorię procesów stochastycznych. Z definicji jest to opis niedeterministyczny. Ale jak powiedzieliśmy, opis jednej cząstki jest w pełni deterministyczny. Twierdzenia o jednoznaczności rozwiązań równań różniczkowych dawały nadzieję na totalny determinizm i przewidywalność ruchu pojedynczych cząstek. Nadzieja ta z praktycznego punktu widzenia okazała się mrzonką. W latach 50-tych XX wieku pokazano, że z praktycznego punktu widzenia determinizm mechaniki Newtona jest złudny i ugruntowana wiara w przewidywalność zachowania się prostych układów mechanicznych załamała się. Pojawiły się liczne przykłady, a później teoria matematyczna, pokazujące niemożliwość przewidywania czasowej ewolucji prostych układów mechanicznych. Podkreślamy, że chodzi tu o praktyczne aspekty przewidywalności. Z matematycznego punktu widzenia, przewidywalność jest ciągle słuszna. Dobrym przykładem nieprzewidywalności w praktyce jest prognoza pogody, co udowadnia codzienne życie. Poniżej przedstawimy zagadnienia, które ukażą nam, co oznacza nieprzewidywalność w teorii deterministycznej. Pokażemy, dlaczego ewolucja określona przez determinizm równań Newtona jest nieprzewidywalna. Ta deterministyczna nieprzewidywalność ma swoją nazwę: deterministyczny chaos.

2.2 Model chaosu. Układ bistabilny (oscylator Duffing)

Aby dobrze zrozumieć istotę chaotycznego zachowania się układów dynamicznych, posłużymy się prostym przykładem z mechaniki klasycznej Newtona. Rozpatrujemy jednowymiarowy ruch cząstki wzdłuż osi OX opisany równaniem Newtona:

$$m\ddot{x} = F(x, \dot{x}, t) = ax - bx^3 - \gamma\dot{x} + A \cos(\Omega t), \quad x(0) = x_0, \quad \dot{x}(0) = v(0) = v_0 \quad (2.1)$$

Model ten wydaje się być banalnie prosty.

jest to ruch cząstki w polu siły $F(x) = ax - bx^3$ ($a, b > 0$) jest to ruch tłumiony tarciem $F(v) = -\gamma v = -\gamma\dot{x}$ oraz na cząstkę działa siła periodyczna w czasie $F(t) = A \cos(\Omega t)$. Zauważmy, że średnia wartość siły $F(t)$ na jednym okresie wynosi zero, czyli średnio działa zerowa siła! Siła $F(x) = x - x^3$ to jest siła potencjalna. Dlatego warto wykreślić potencjał $V(x)$ odpowiadający tej sile:

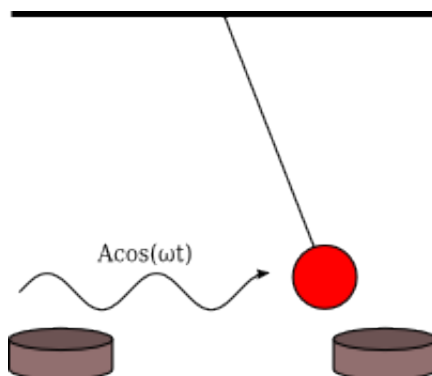
$$V(x) = - \int F(x) dx = \frac{1}{4}bx^4 - \frac{1}{2}ax^2 \quad (2.2)$$

Potencjał ten nazywa się potencjałem bistabilnym i jest jednym z najważniejszych modelowych potencjałów w fizyce, począwszy od teorii Newtona, poprzez teorię przejść fazowych,

teorię aktywacji reakcji chemicznych, a na teorii cząstek elementarnych (mechanizm Higgsa) skończywszy. Poniżej pokazujemy wykres tego potencjału (mówiąc precyzyjnie: energii potencjalnej).

Powyższe równanie Newtona ma kilka realizacji.

1 Pierwszy przykład realizacji: Metalowa kulka zawieszona na nieważkim pręcie w polu dwóch magnesów (które modelują bistabilny potencjał). Na kulkę działa w kierunku poziomym okresowe pole elektryczne $A \cos(\omega t)$. Układ ten jest przedstawiony na rysunku.

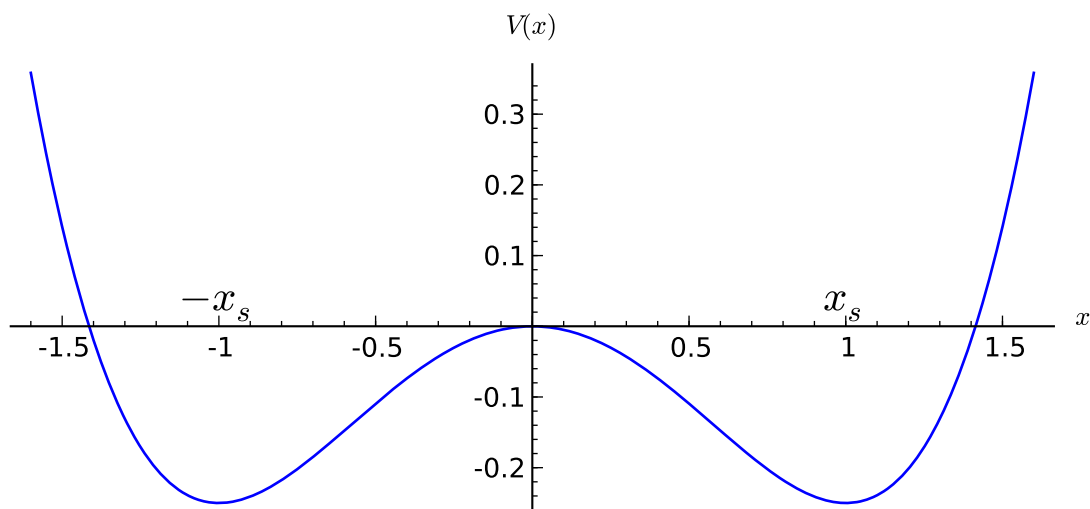


Rysunek 2.1: Przykład realizacji układu bistabilnego.

2 Drugi przykład realizacji: Obwód elektroniczny, który jest szczegółowo opisany w pracy:

B. K. Jones and G. Trefan, Am. J. Phys. 69 (2001) str. 464. “The Duffing oscillator: A precise electronic analog chaos demonstrator for the undergraduate laboratory”

```
V = x^4/4 - x^2/2
plot(V, (x, -1.6, 1.6))
```



Rysunek 2.2: Potencjał bistabilny.

2.2.1 Skalowanie

Układ opisany powyżej zawiera 6 parametrów. Część parametrów można wyeliminować poprzez przeskalowanie równania do postaci bezwymiarowej. Istnieje kilka możliwości. Zwykle zaczynamy od skalowania czasu i położenia. Nowy bezwymiarowy czas τ ma postać:

$$s = \frac{t}{\tau_0}, \quad \tau_0^2 = \frac{m}{a} \quad (2.3)$$

Nowe bezwymiarowe położenie definiujemy jako

$$X = \frac{x}{L}, \quad L^2 = \frac{a}{b} \quad (2.4)$$

Wówczas bezwymiarowa postać równania ruchu jest następująca:

$$\ddot{X} = X - X^3 - \gamma_0 \dot{X} + A_0 \cos(\omega_0 s), \quad X(0) = X_0, \quad \dot{X}(0) = \dot{X}_0 \quad (2.5)$$

Obecnie występują 3 przeskalowane parametry:

$$\gamma_0 = \frac{\tau_0^2}{mL} \gamma, \quad A_0 = \frac{\tau_0^2}{mL} A, \quad \omega_0 = \tau_0 \Omega \quad (2.6)$$

W dalszej części będziemy posługiwali się tylko i wyłącznie przeskalowanym równaniem. Dla tego wygodnie będzie używać “starych” oznaczeń: Będziemy analizowali równanie w postaci

$$\ddot{x} = x - x^3 - \gamma \dot{x} + A \cos(\omega_0 t), \quad x(0) = x_0, \quad \dot{x}(0) = \dot{y}_0 = v_0 \quad (2.7)$$

gdzie przeskalowany potencjał

$$V(x) = - \int F(x) dx = \frac{1}{4} x^4 - \frac{1}{2} x^2 \quad (2.8)$$

Przeskalowane równanie jest w takiej postaci, że przyjmujemy wartości parametrów $m = 1, a = 1, b = 1$.

2.2.2 Krok 1. Układ zachowawczy

W pierwszym kroku rozpatrujemy najprostszy przypadek (pamiętajmy o przeskalowanej postaci, w której masa cząstki $m = 1$):

$$\ddot{x} = x - x^3 = -V'(x), \quad x(0) = x_0, \quad \dot{x}(0) = v(0) = v_0 \quad (2.9)$$

Jest on równoważny układowi 2 równań różniczkowych, autonomicznych, pierwszego rzędu:

$$\begin{aligned} \dot{x} &= v, & x(0) &= x_0, \\ \dot{v} &= x - x^3, & v(0) &= v_0. \end{aligned} \quad (2.10)$$

Oznacza to, że przestrzeń fazowa jest 2-wymiarowa.

Taki przypadek był już rozpatrywany: jest to układ zachowawczy o jednym stopniu swobody. Istnieje jedna stała ruchu (jedna całka ruchu), a mianowicie całkowita energia układu:

$$\frac{1}{2}\dot{x}^2(t) + V(x(t)) = \text{const.} = E = E_k + E_p = \frac{1}{2}\dot{x}^2(0) + V(x(0)) = \frac{1}{2}v_0^2 + V(x_0) \quad (2.11)$$

na którą składa się energia kinetyczna E_k oraz energia potencjalna E_p . Stała E jest określona przez warunki początkowe $x(0) = x_0$ oraz $v(0) = v_0$. Ponieważ jest zachowana całkowita energia układu, ruch jest periodyczny. Nie istnieją atraktory i nie istnieją asymptotycznie stabilne stany stacjonarne. Krzywe fazowe są zamknięte co oznacza że cząstka porusza się periodycznie w czasie. W zależności od warunków początkowych, amplituda drgań jest większa lub mniejsza, ponieważ warunki początkowe wyznaczają wartość stałej ruchu E . Jeżeli dwa warunki początkowe (x_{01}, v_{01}) oraz (x_{02}, v_{02}) nieznacznie się różnią, np. w sensie odległości:

$$|[x_{01}^2 + v_{01}^2] - [x_{02}^2 + v_{02}^2]| \ll 1 \quad (2.12)$$

to krzywe fazowe nieznacznie się różnią i ruch cząstki dla tych dwóch warunków początkowych nieznacznie się różni. Mówimy wówczas, że układ jest nieczuły na zmianę warunków początkowych. Jak widać z powyższego wzoru, dwa różne warunki początkowe oznaczają, że układ ma dwie różne energie E . To z kolei oznacza, że częstotści ruchu periodycznego także będą różne. Różnica częstotści powoduje, że cząstki będą się powoli oddalać od siebie, ale tempo oddalania będzie liniowe w czasie. Gdyby tempo oddalania było znacznie szybsze, a mianowicie rośnie eksponencjalnie w czasie, zachowanie takie nazwalibyśmy chaotycznym. Do tego problemu powrócimy poniżej, ponieważ jest on kluczowym dla zrozumienia chaotycznego zachowania się układu.

Poniżej przedstawiamy potencjał i krzywe fazowe dla tego przypadku.

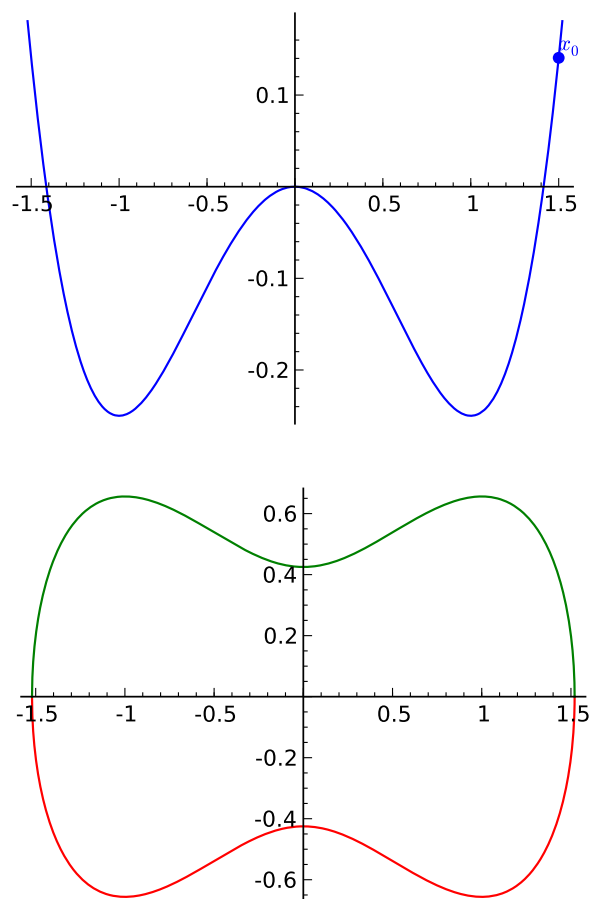
```
#parametry dla wizualizacji
var('x v')
x0, v0 = 1.5, 0.2
V = x^4/4 - x^2/2
E = V(x=x0) + v0^2

#prawo zachowania energii
PZE = v^2 + V == E

#wychylenia ekstremalne
rozv = solve(PZE(v=0), x)
xmin = min([i.rhs() for i in rozv if imag(i.rhs()) == 0])
xmax = max([i.rhs() for i in rozv if imag(i.rhs()) == 0])

#i jego rozwiązanie wobec v
rozv = solve(PZE, v)
v1=rozv[0].rhs()
v2=rozv[1].rhs()
vmax = abs(v1(x=0))

#potencjał
p1 = plot(V, (x,xmin,xmax))
#krzywe fazowe
p2 = plot(v1, (x,xmin,xmax), color='red')
p2+= plot(v2, (x,xmin,xmax), color='green')
```



Rysunek 2.3: Potencjał (górna krzywa) oraz krzywe fazowe (dolna) dla układu zachowawczego.

2.2.3 Krok 2. Układ dysypatywny czyli wpływ tarcia.

W drugim kroku dodajemy tarcie i rozpatrujemy równanie ruchu w postaci:

$$\ddot{x} = x - x^3 - \gamma \dot{x}, \quad x(0) = x_0, \quad \dot{x}(0) = v(0) = v_0 \quad (2.13)$$

Jest on równoważny układowi 2 równań różniczkowych, autonomicznych, pierwszego rzędu:

$$\begin{aligned} \dot{x} &= v, & x(0) &= x_0, \\ \dot{v} &= x - x^3 - \gamma v, & v(0) &= v_0. \end{aligned} \quad (2.14)$$

Oznacza to, że przestrzeń fazowa jest 2-wymiarowa.

Taki przypadek był także rozpatrywany: jest to układ dyssypatywny o jednym stopniu swobody. Nie istnieje już stała ruchu E . Całkowita energia układu maleje w czasie. W tym układzie istnieją 3 stany stacjonarne. Stany te określone są przez równanie:

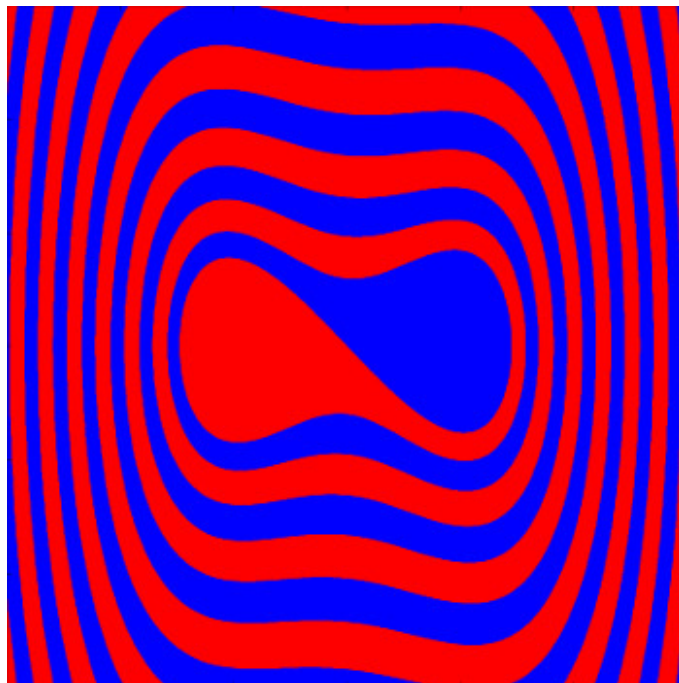
$$x - x^3 = 0, \quad \text{stąd} \quad x_{s0} = 0, \quad x_{s1} = 1, \quad x_{s2} = -1 \quad (2.15)$$

Stany stacjonarne $x_{s1} = 1$ oraz $x_{s2} = -1$ są stabilne. Stan $x_{s0} = 0$ jest niestabilny. Istnieją 2 atraktory $A_1 = x_{s1} = 1$ oraz $A_2 = x_{s2} = -1$ i 2 obszary przyciągania $B(A_1)$ oraz $B(A_2)$,

których suma mnogościowa $B(A_1) \cup B(A_2) = R^2$ jest całą płaszczyzną. Krzywe fazowe zawsze dążą do jednego z atraktorów lub do niestabilnego stanu stacjonarnego. Jeżeli dwa warunki początkowe (x_{01}, v_{01}) oraz (x_{02}, v_{02}) nieznacznie się różnią np. w sensie odległości:

$$|[x_{01}^2 + v_{01}^2] - [x_{02}^2 + v_{02}^2]| \ll 1 \quad (2.16)$$

i są w tym samym obszarze przyciągania, to krzywe fazowe nieznacznie się różnią i ruch cząstki dla tych dwóch warunków początkowych nieznacznie się różni. Mówimy wówczas, że układ jest nieczuły na zmianę warunków początkowych. Natomiast jeżeli dwa warunki początkowe $(x_{01}, v_{01}) \in B(A_1)$ oraz $(x_{02}, v_{02}) \in B(A_2)$ nieznacznie się różnią, ale są w dwóch obszarach przyciągania $B(A_1)$ oraz $B(A_2)$, to trajektorie zaczną po pewnym czasie różnić się znacznie, będą przyciągane do dwóch różnych atraktorów i będą dążyć do dwóch różnych stanów stacjonarnych $x_{s1} = 1$ oraz $x_{s2} = -1$. Tym niemniej, w takiej sytuacji mówimy, że układ jest nieczuły na zmianę warunków początkowych w sensie o którym mowa powyżej.



Rysunek 2.4: Diagram basenów przyciągania dla potencjału bistabilnego

Kolor niebieski to obszar warunków początkowych które są “przyciągane” do atraktora $(1, 0)$, do prawego minimum potencjału. Kolor czerwony to obszar warunków początkowych które są “przyciągane” do atraktora $(-1, 0)$, do lewego minimum potencjału. W zależności od wartości stałej tłumienia γ , diagram ten przybiera nieco inne kształty, ale struktura dwu-kolorowych pasów pozostaje. Brzeg obszarów przyciągania jest gładką krzywą, której wymiar wynosi 1. Jeżeli warunki początkowe są położone dokładnie na tym brzegu, to cząstka porusza się do niestabilnego stanu stacjonarnego $(x = 0, v = 0)$ (maksimum potencjału).

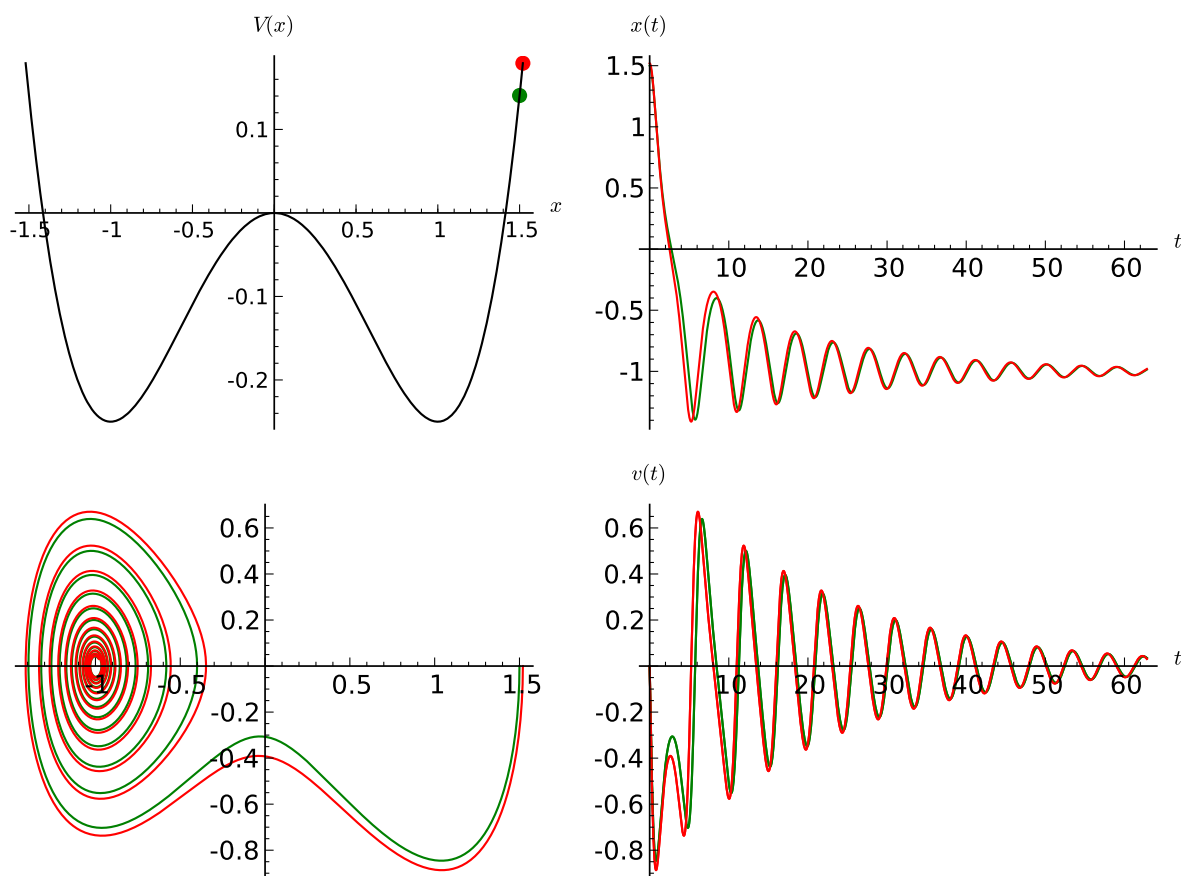
```
# wykresy dla przypadku z tłumieniem
var('x v')
x01, v01 = 1.50, 0
x02, v02 = 1.52, 0
```

```
# siła i potencjał
F = x-x^3
V = -integrate(F,x)

# tarcie: parametr gamma
g = 0.1

# numeryczne rozwiązanie równań ruchu
T = srange(0,20*pi,0.01)
de = vector([v,F-g*v])
num1 = desolve_odeint(de, [x01,v01], T, [x,v])
num2 = desolve_odeint(de, [x02,v02], T, [x,v])

#krzywe fazowe
lt = plot(V)
lb = list_plot(num1.tolist())
lb += list_plot(num2.tolist())
rt = list_plot(zip(T,num1[:,0].tolist()))
rt += list_plot(zip(T,num2[:,0].tolist()))
rb = list_plot(zip(T,num1[:,1].tolist()))
rb += list_plot(zip(T,num2[:,1].tolist()))
```

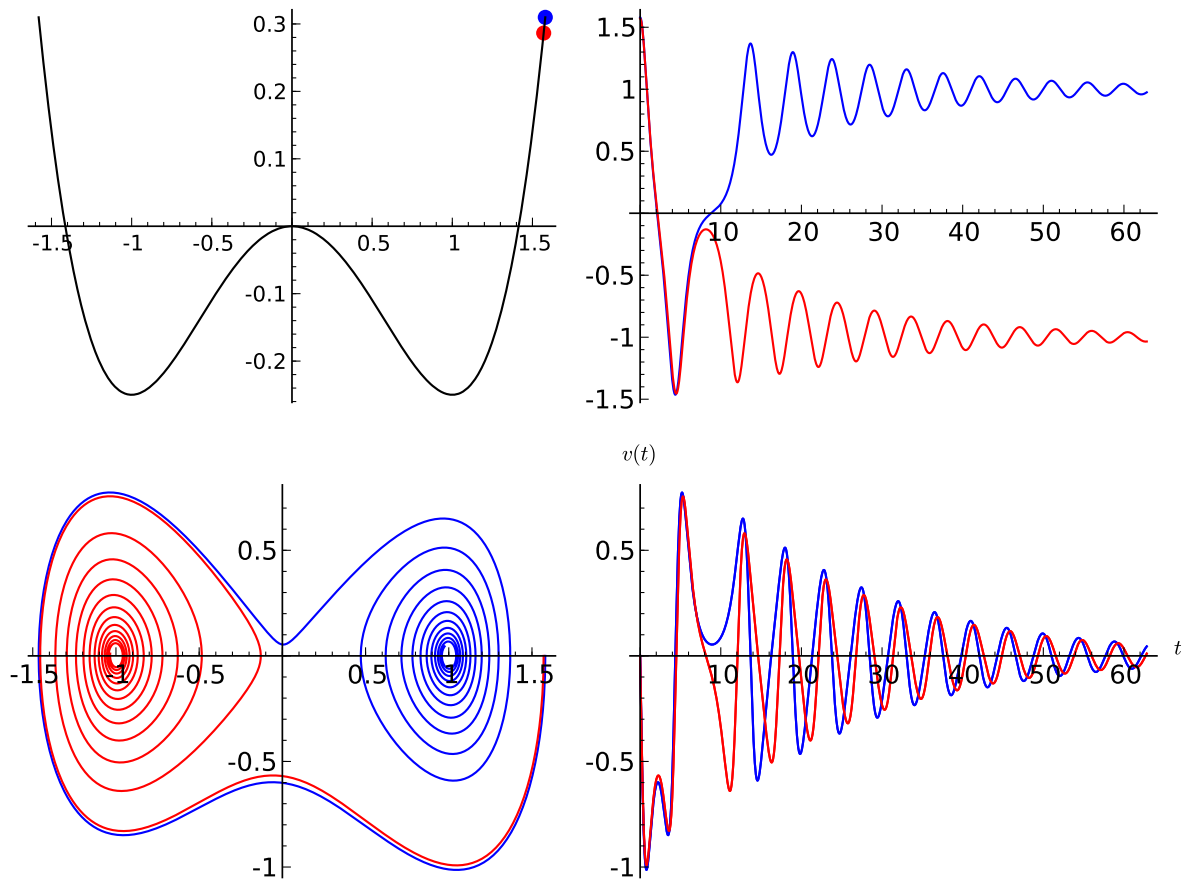


Rysunek 2.5: Potencjał (górny lewy panel), krzywe fazowe (lewy dolny panel), trajektorie (prawy górny panel) oraz prędkość chwilowa (prawy dolny panel) dla układu z tłumieniem.

Na powyższym zestawie rysunków, 2 warunki początkowe leżą w tym samym obszarze przyciągania atraktora $(-1, 0)$. Oznacza to, że 2 warunki początkowe są umiejscowione w czerwonym obszarze na diagramie basenów przyciągania pokazanym powyżej. Układ nie jest czuły na zmianę warunków początkowych, gdy leżą one w tym samym basenie przyciągania.

Teraz zobaczmy jak to wygląda dla innego zestawu warunków początkowych. W powyższym kodzie wystarczy podmienić dwie linijki.

```
x01, v01 = 1.58, 0
x02, v02 = 1.57, 0
```



Rysunek 2.6: Potencjał (górny lewy panel), krzywe fazowe (lewy dolny panel), trajektorie (prawy górny panel) oraz prędkość chwilowa (prawy dolny panel) dla układu z tłumieniem.

Na powyższym zestawie rysunków, 2 warunki początkowe leżą w dwóch różnych obszarach przyciągania. Oznacza to, że 1 warunek początkowy leży w niebieskim obszarze na diagramie basenów przyciągania, natomiast 2 warunek początkowy leży w czerwonym obszarze na diagramie basenów przyciągania. Te dwa warunki początkowe leżą blisko brzegu 2 basenów przyciągania. Dlatego układ jest czuły na zmianę warunków początkowych, pod warunkiem że leżą one w dwóch różnych basenach przyciągania. Ale to nie jest jeszcze kryterium własności chaotycznych układu.

2.2.4 Krok 3. Układ z tarciem i periodyczną siłą.

W trzecim kroku dodajemy siłę periodyczną w czasie i rozpatrujemy równanie ruchu w wyjściowej pełnej postaci:

$$\ddot{x} = x - x^3 - \gamma \dot{x} + A \cos(\omega_0 t), \quad x(0) = x_0, \quad \dot{x}(0) = v(0) = v_0 \quad (2.17)$$

Jest on równoważny układowi 3 równań różniczkowych, autonomicznych, pierwszego rzędu:

$$\begin{aligned} \dot{x} &= v, & x(0) &= x_0, \\ \dot{v} &= x - x^3 - \gamma v + A \cos z, & v(0) &= v_0, \\ \dot{z} &= \omega_0, & z(0) &= 0. \end{aligned} \quad (2.18)$$

Oznacza to, że przestrzeń fazowa jest 3-wymiarowa.

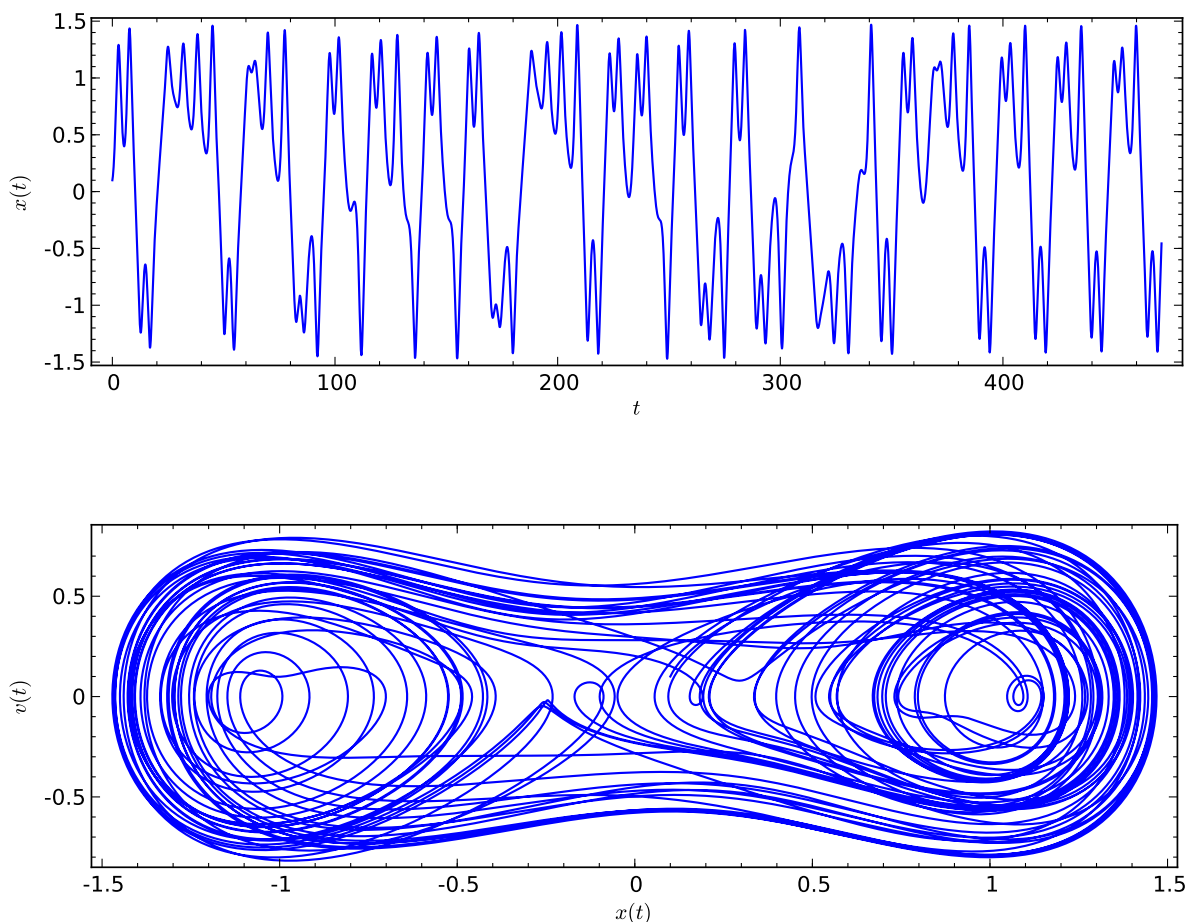
Matematycy wolą przepisać powyższy układ równań dla “tradycyjnych” 3 zmiennych (x, y, z) w postaci:

$$\begin{aligned} \dot{x} &= y, & x(0) &= x_0, \\ \dot{y} &= x - x^3 - \gamma y + A \cos z, & y(0) &= y_0, \\ \dot{z} &= \omega_0, & z(0) &= 0. \end{aligned} \quad (2.19)$$

czyli prędkość cząstki v jest teraz oznaczona jako $v = y$.

Okazuje się, że pełny układ wykazuje radykalnie inne własności od poprzednich 2 przypadków. Z punktu widzenia fizyki mamy taki oto proces: Cząstka porusza się w bistabilnym potencjale. Ponieważ potencjał dąży do nieskończoności gdy położenie dąży do nieskończoności, ruch cząstki jest ograniczony; cząstka jest uwięziona w potencjale i nie może uciec do nieskończoności. Siła tarcia pcha cząstkę do jednego ze (“starych”) stanów stacjonarnych x_{s1} lub x_{s2} . Z kolei zewnętrzna siła periodyczna w czasie pompuje energię do układu i przeciwdziała sile tarcia. Cząstka już nie dąży do stanu stacjonarnego, nie zatrzyma się dla długich czasów ale będzie ciągle poruszać się i nigdy już nie spocznie. Istotne stają się efekty inercjalne związane z masą cząstki, które są odzwierciedlone w wyrazie \dot{y} , czyli przyspieszeniu cząstki. Istotne jest to, że nie jest to ruch przetłumiony. W konsekwencji układ nie posiada stanu stacjonarnego w postaci punktu w przestrzeni fazowej jak to było w przypadku 2. Wszystkie te powyższe czynniki stają się istotne dla zrozumienia skomplikowanych i złożonych własności ewolucji cząstki.

```
var('x y z')
T = srange(0, 150*pi, 0.01)
de = vector([y, x-x^3-0.26*y+0.3*cos(z), 1])
sol = desolve_odeint(de, [0.1, 0.1, 0], T, [x, y, z])
t = line(zip(T, sol[:, 0]))
b = line(zip(sol[:, 0], sol[:, 1]))
```



Rysunek 2.7: Trajektoria (górny panel) oraz krzywa fazowe (dolny panel), dla układu wykazującego chaos.

2.2.5 Ruch periodyczny o okresie 1

W modelu występują 3 bezwymiarowe parametry: współczynnik tarcia γ , amplituda zewnętrznej siły A oraz częstość drgań ω_0 siły periodycznej w czasie. Poniżej pokażemy kilka charakterystycznych trajektorii układu. Zaczniemy od prostej periodycznej ewolucji, ruchu okresowego o tzw. okresie 1.

Założmy następujące wartości parametrów:

$$\gamma = 0.15, \quad A = 0.3, \quad \omega_0 = 1 \quad (2.20)$$

W tym przypadku obserwujemy regularny ruch. Jeżeli nieco zaburzymy warunki początkowe, to nowy ruch jest także regularny (trzeba być ostrożnym, gdy mówimy “nieco zaburzymy”).

```
var('x y z')
x0, y0, z0 = 0.1, 0.1, 0
kolor = 'green'
```

```
# siła
```

```

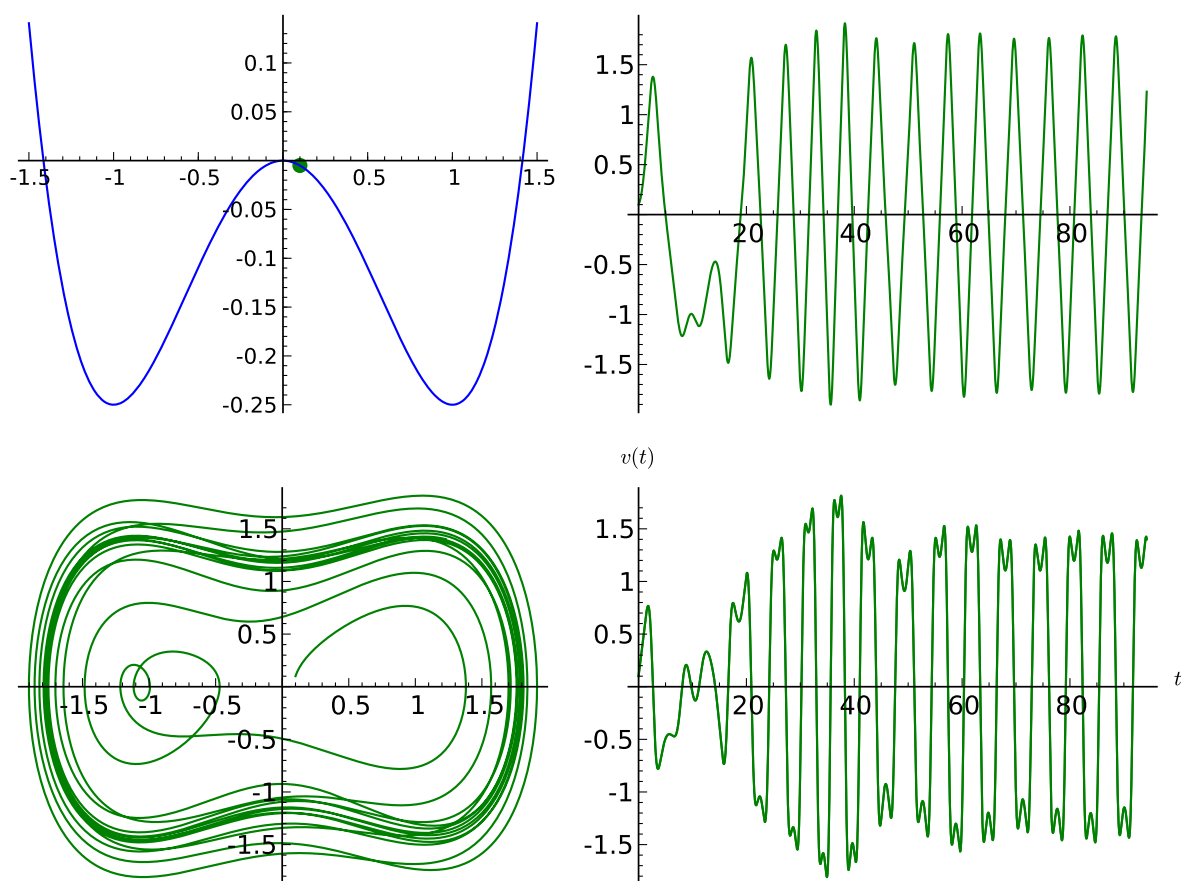
F = x-x^3
V = -integrate(F,x)

# tarcie: parametr gamma
g = 0.1
A = 0.3
w = 1

# układ różniczkowych równań ruchu
dx = y
dy = F - g*y + A*cos(z)
dz = w

# numeryczne rozwiązanie równań ruchu
T = srange(0,30*pi,0.01)
num = desolve_odeint(vector([dx,dy,dz]), [x0,y0,z0], T, [x,y,z])

```



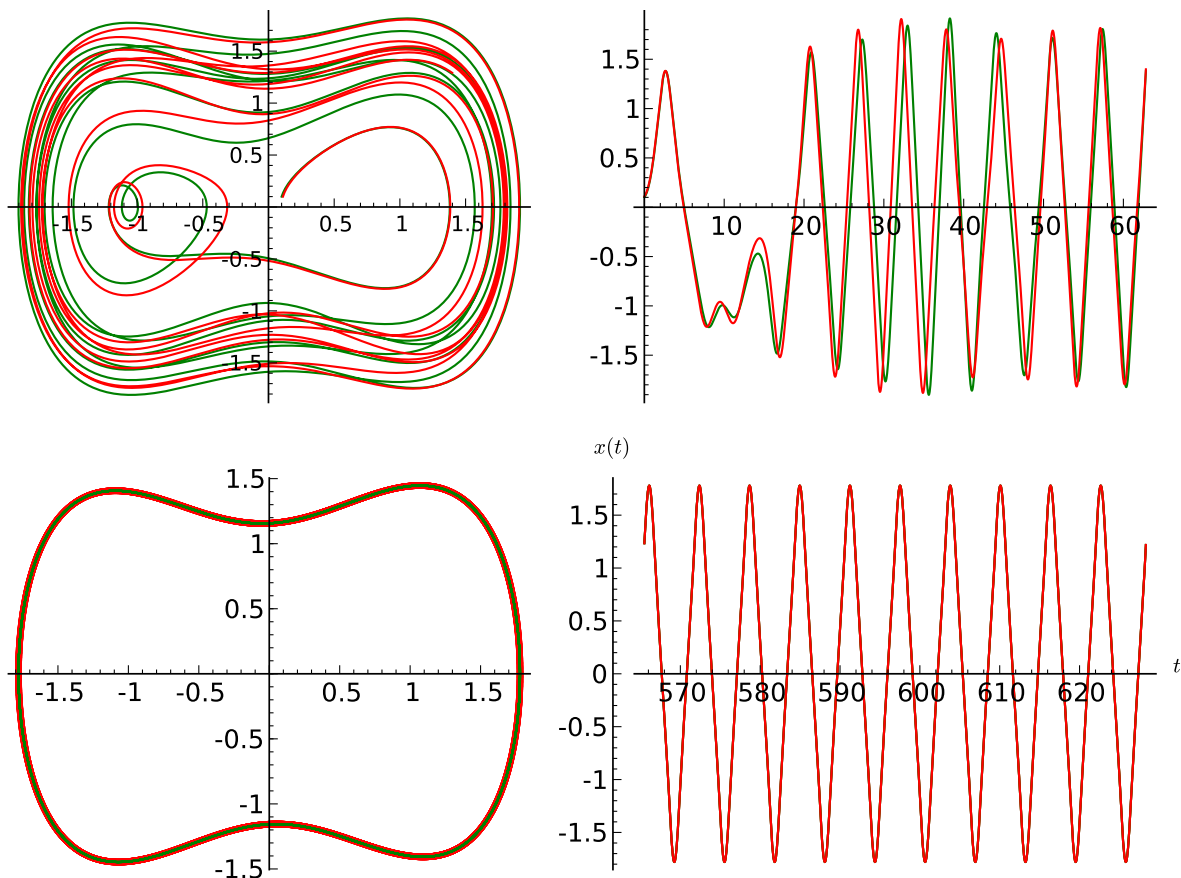
Rysunek 2.8: Potencjał (górny lewy panel), krzywe fazowe (lewy dolny panel), trajektorie (prawy górny panel) oraz prędkość chwilowa (prawy dolny panel) dla układu o okresowości 1.

Przyjrzyjmy się teraz dwóm trajektoriom startującym z bliskich warunków początkowych. Rozpatrzmy ich początkową i asymptotyczną (dla długich czasów) ewolucję.

```

var('x y z')
x01, y01, z01 = 0.1, 0.1, 0
x02, y02, z02 = 0.11, 0.1, 0
F = x - x^3
g, A, w = 0.1, 0.3, 1
dx, dy, dz = y, F - g*y + A*cos(z), w
T = srange(0, 200*pi, 0.01)
num1 = desolve_odeint(vector([dx, dy, dz]), [x01, y01, z01], T, [x, y, z])
num2 = desolve_odeint(vector([dx, dy, dz]), [x02, y02, z02], T, [x, y, z])
lnum = int(len(num1[:, 0])/10)
# ewolucja początkowa
trans1 = num1[:lnum]
trans2 = num2[:lnum]
# ewolucja asymptotyczna
asyp1 = num1[-lnum:]
asyp2 = num2[-lnum:]

```



Rysunek 2.9: W lewej kolumnie zaprezentowano krzywe fazowe. W prawej kolumnie wi-
dać trajektorie. Górny rząd dopowiada ewolucji początkowej a dolny asymptotycznej układu
Duffinga pracującego o okresowości 1.

Na dwóch górnych diagramach przedstawioną reżim krótkich czasów. Ponieważ 2 warunki po-
czątkowe nieco się różnią, więc początkowa ewolucja nieco się różni. Kolor czerwony i zielony
jest rozróżnialny na prawym górnym rysunku pokazującym ewolucję $x(t)$ dla krótkich czasów.
Jeżeli przyjrzymy się reżimowi długich czasów (dwa dolne diagramy) to zauważymy duże

podobieństwo w ewolucji: krzywe fazowe są zamknięte więc jest to prosty ruch periodyczny, przypominający nieco zdeformowaną funkcję typu $\sin(\alpha t)$ czy też $\cos(\alpha t)$. Jest to funkcja okresowa z charakterystycznym jednym jedynym okresem T . Dlatego mówimy, że jest ruch periodyczny o okresie 1. Dwie krzywe $x(t)$ na dolnym prawym rysunku nie są rozróżnialne.

Można zrobić doświadczenie numeryczne i wybierać różne warunki początkowe. Zobaczymy, że trajektorie dążą do tego samego okresowego rozwiązania, są przyciągane do tego okresowego rozwiązania. Innymi słowy, ta krzywa fazowa o okresie 1 jest ATRAKTOREM. Atraktor ten nazywa się periodycznym atraktorem o okresie 1 lub 1-okresowym atraktorem. Można by postawić pytanie: jak wygląda basen przyciągania dla tego atraktora. Aby dać odpowiedź na to pytanie należy zbadać numerycznie np. kwadrat warunków początkowych (x_0, y_0) i wybrać te warunki początkowe które dążą do powyższej krzywej fazowej o okresie 1. Okazuje się, że basen przyciągania jest “porządnym” zbiorem, którego brzeg jest gładką krzywą, podobnie jak w przypadku zilustrowanym powyżej dla układu tylko z tarcie, bez siły okresowej.

2.2.6 Ruch periodyczny o okresie 3

Założmy następujące wartości parametrów:

$$\gamma = 0.22, \quad A = 0.3, \quad \omega_0 = 1 \quad (2.21)$$

W tym przypadku obserwujemy także periodyczny ruch, ale nieco bardziej skomplikowany. Nie jest to prosty periodyczny ruch, ale tzw. ruch o okresie 3, tzn. teraz okres jest 3 razy dłuższy niż w poprzednim przypadku.

I znów zobaczymy, jak początkowa ewolucja różni się od tej po długim czasie.

Dla długich czasów, krzywe fazowe są zamknięte, ale nie są to krzywe typu zdeformowana elipsa. To są krzywe z 2 pętelkami. Tym niemniej, ruch jest periodyczny.

Podobnie jak poprzednim przypadku, można zrobić doświadczenie numeryczne i wybierać różne warunki początkowe. Zobaczymy, że wiele trajektorii dąży do tej samej okresowej orbity, są one przyciągane do tej zamkniętej krzywej fazowej. Innymi słowy, ta krzywa fazowa o okresie 3 jest ATRAKTOREM. Atraktor ten nazywa się periodycznym atraktorem o okresie 3 lub 3-okresowym atraktorem. Basen przyciągania dla tego atraktora na płaszczyźnie warunków początkowych (x_0, y_0) jest “porządnym” zbiorem o wymiarze 2 (czyli kawałek płaszczyzny), którego brzeg jest gładką krzywą.

2.2.7 Ruch chaotyczny

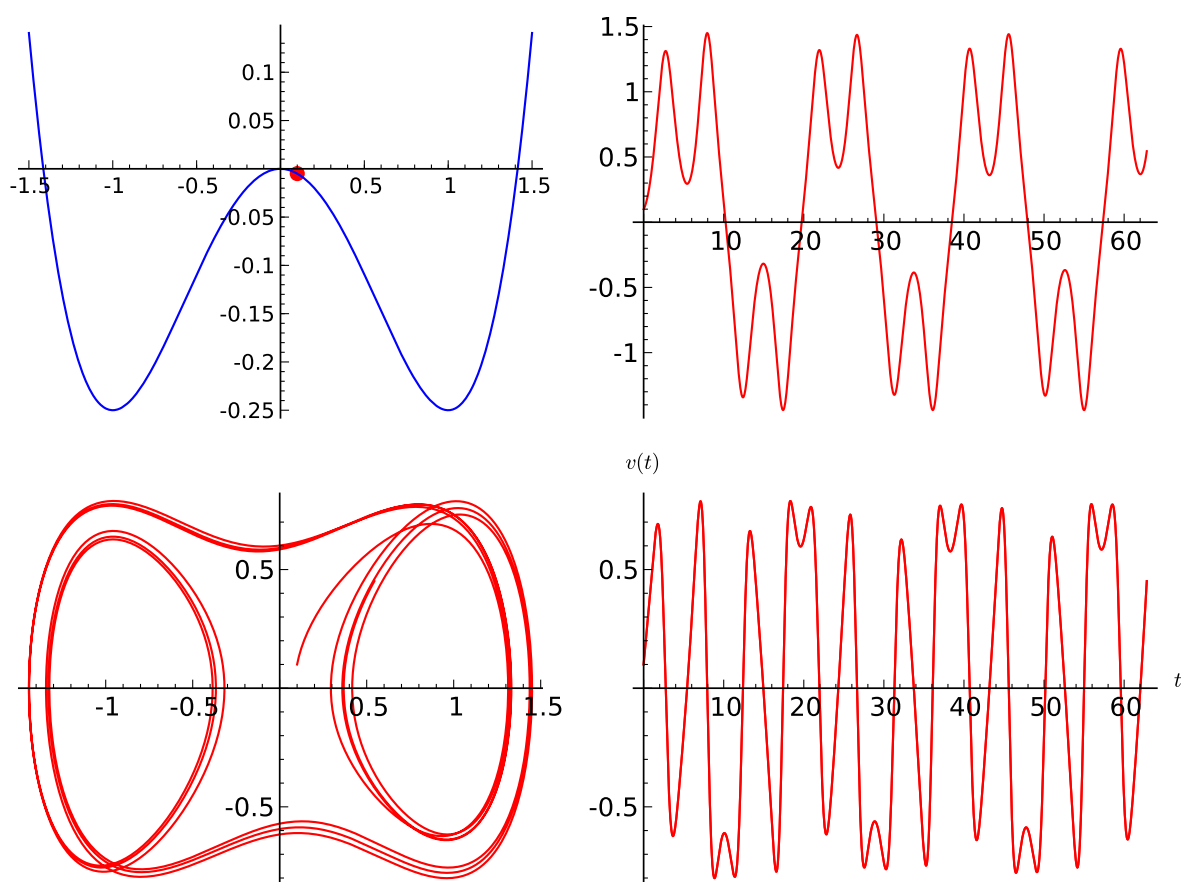
Założmy następujące wartości parametrów:

$$\gamma = 0.25, \quad A = 0.3, \quad \omega_0 = 1 \quad (2.22)$$

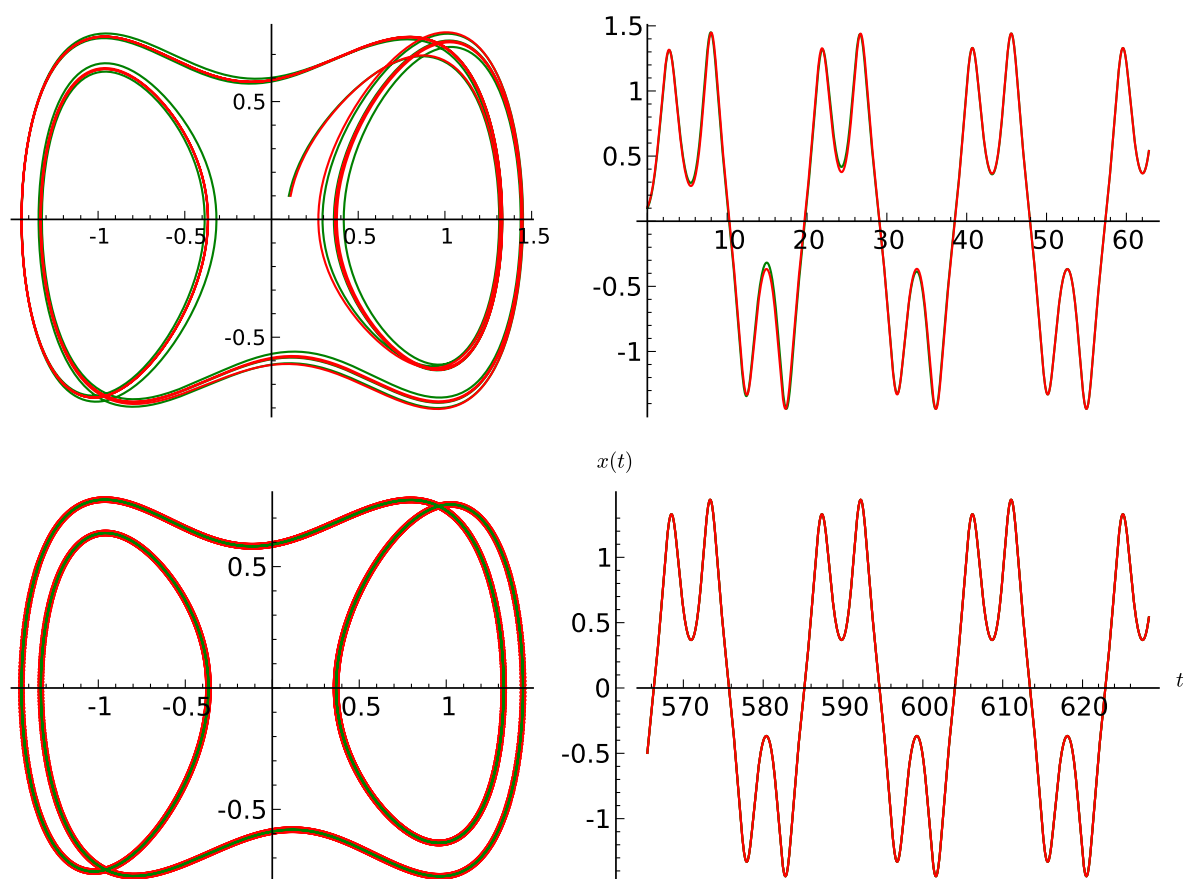
W tym przypadku obserwujemy ruch, który wydaje się być wyjątkowo nieregularny, chaotyczny.

Zobaczymy, jak tym razem ewoluują rozwiązania o 2 bliskich warunkach początkowych.

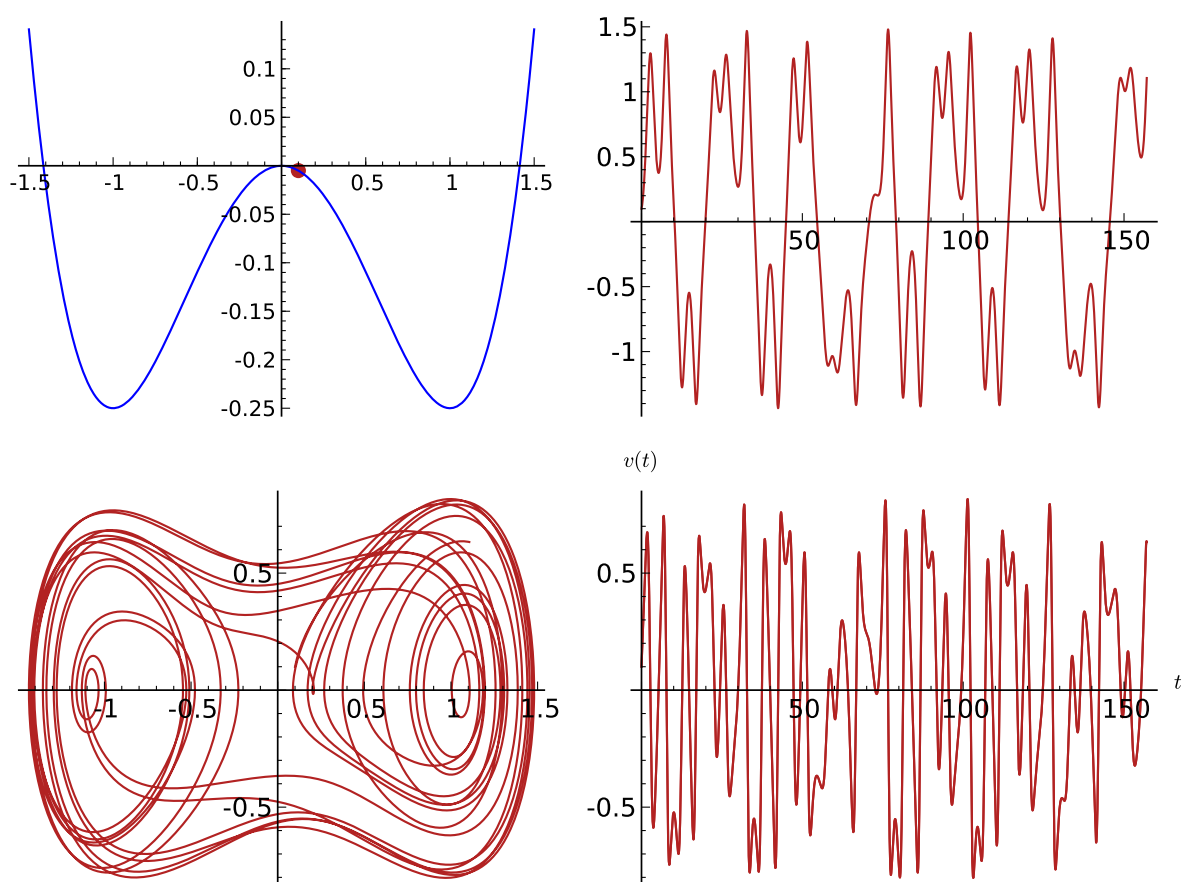
Początkowa ewolucja dwóch rozwiązań jest nierozróżnialna (ponieważ 2 warunki początkowe są bardzo blisko siebie). Po pewnym charakterystycznym czasie, zwanym czasem Lapunowa,



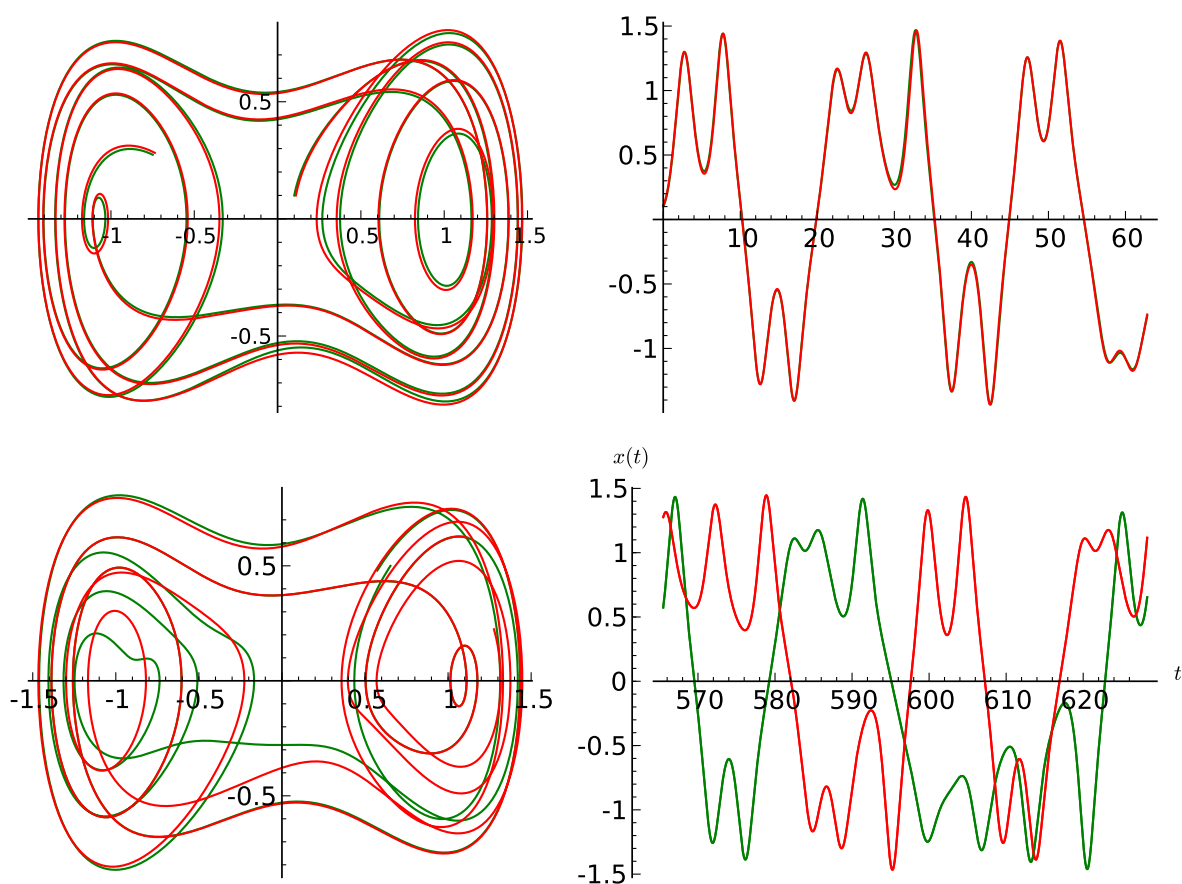
Rysunek 2.10: Potencjał (górny lewy panel), krzywe fazowe (lewy dolny panel), trajektorie (prawy górny panel) oraz prędkość chwilowa (prawy dolny panel) dla układu o okresowości 3.



Rysunek 2.11: W lewej kolumnie zaprezentowano krzywe fazowe. W prawej kolumnie wi-
dać trajektorie. Górny rząd dopowiada ewolucji początkowej a dolny asymptotycznej układu
Duffinga o okresowości 3.

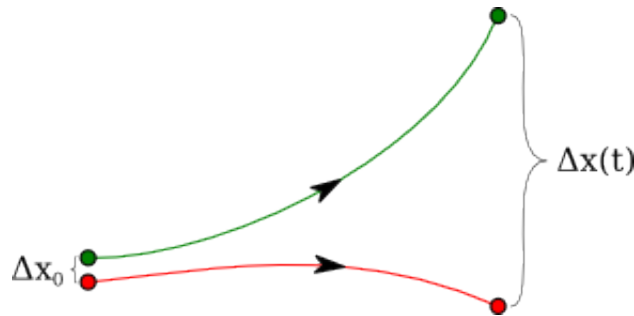


Rysunek 2.12: Potencjał (górny lewy panel), krzywe fazowe (lewy dolny panel), trajektorie (prawy górny panel) oraz prędkość chwilowa (prawy dolny panel) dla układu chaotycznego.



Rysunek 2.13: W lewej kolumnie zaprezentowano krzywe fazowe. W prawej kolumnie wi-
dać trajektorie. Górny rząd dopowiada ewolucji początkowej a dolny asymptotycznej układu
Duffinga o pracującego w reżimie chaotycznym.

trajektorie zaczynają różnić się coraz bardziej, zaczynają rozbiegać się: patrz trajektoria czerwona i zielona na dolnym prawym rysunku.



Rysunek 2.14: Schematyczne trajektorie w reżimie chaotycznym.

W reżimie chaotycznym, te dwie trajektorie oddalają się od siebie w eksponencjalnie szybkim tempie określonym przez zależność:

$$|x_1(t) - x_2(t)| = |x_1(0) - x_2(0)|e^{\lambda t}, \quad \lambda > 0 \quad (2.23)$$

lub

$$|\Delta x(t)| = |\Delta x_0|e^{\lambda t}, \quad \lambda > 0 \quad (2.24)$$

gdzie λ nazywa się wykładnikiem Lapunowa.

Różnice w ewolucji stają się zbyt duże i pojawia się dylemat: która trajektoria jest właściwa, skoro nasza aparatura nie rozróżnia bliskich warunków początkowych. Determinizm staje się złudnym. Nie możemy przewidywać właściwej ewolucji układu.

Przedstawiony powyżej reżim chaotyczny nie jest jedyny. W układzie istnieje wiele takich wartości parametrów (γ, A, ω) , dla których pojawia się ruch chaotyczny. Należy nadmienić, że dla długich czasów wiele trajektorii generowanych przez różne warunki początkowe zachowuje się bardzo podobnie, wiele trajektorii jest przyciąganych. Tu także istnieje atraktor i jego basen przyciągania. Jednakże ten atraktor jest dziwny: jego wymiar nie jest liczbą całkowitą i atraktor jest fraktalem. Dlatego nazywa się dziwnym atraktorem. Brzeg basenu przyciągania tego atraktora też ma dziwną strukturę i jego wymiar jest fraktalny.

Zadania

- Niech $\gamma = 0.1$, $\omega_0 = 1.4$, $(x_0, y_0, z_0) = (-0.5, -0.2, 0)$. Zmieniaj parametr $A = 0.1, 0.32, 0.338, 0.35$.

Obserwuj scenariusz podwojenia okresu:

- pojawia się atraktor periodyczny o okresie 1.
 - pojawia się atraktor periodyczny o okresie 2.
 - pojawia się atraktor periodyczny o okresie 4.
 - pojawia się atraktor periodyczny o okresie 8 (trudno trafić).
 - pojawia się ruch nieregularny, chaotyczny.
- Zbadaj zachowanie się układu dla następujących wartości parametrów: $\gamma = 1.35 - 1.38$, $A = 1$, $\omega_0 = 1$, $(x_0, y_0, z_0) = (0.0, 0.5, 0)$.
 - To samo dla wartości $\gamma = 0.5$, $A = 0.34875$, $\omega_0 = 1$, $(x_0, y_0, z_0) = (0, 0, 0)$

2.3 Scenariusz przejścia do chaosu

Zmieniając parametry układu oraz warunki początkowe, możemy sterować własnościami ewolucji czasowej. Widzieliśmy, że istnieją rozwiązania periodyczne. Może to być prosty ruch periodyczny charakteryzujący się jednym charakterystycznym okresem T (lub częstością). Mogą to być ruchy periodyczne bardziej skomplikowane: o okresie 2, 3, 4, itd. Zauważmy, że ruch periodyczny o okresie 3 powtarza się po czasie 3 razy dłuższym niż ruch o okresie 1. Dlatego też regularność ruchu można zaobserwować po czasie 3 razy dłuższym. Ruch periodyczny o okresie 20 powtarza się po czasie 20 razy dłuższym niż ruch o okresie 1. Dlatego też regularność ruchu jest obserwowana po czasie 20 razy dłuższym. Ruch periodyczny o okresie 2000 powtarza się po czasie 2000 razy dłuższym niż ruch o okresie 1. Dlatego też regularność ruchu może być rozpoznana po czasie 200 razy dłuższym. Zwiększając periodyczność ruchu aż do nieskończoności, zauważamy że regularność ruchu powtarza się po nieskończonym czasie, czyli ruch staje się nieregularny dla obserwatora. Trajektoria wygląda tak, jakby to był ruch przypadkowy, losowy, chaotyczny. Ruch jest ciągle deterministyczny, ale skomplikowany, niepowtarzalny, nieregularny. W niektórych przypadkach układ jest wyjątkowo wrażliwy na warunki początkowe: dla dwóch różnych, ale bardzo mało różniących się warunków początkowych, odpowiadające im trajektorie z czasem zaczynają się różnić i odbiegać od siebie. Jeżeli zmniejszymy odległość między warunkami początkowymi, to czas po jakim można rozróżnić 2 trajektorie wydłuża się, ale prędzej czy później, trajektorie zaczynają się rozbiegać. Z praktycznego punktu widzenia, warunki początkowe można zadawać ze skończoną dokładnością, ale nie z zerową dokładnością, tak jak to się zakłada w twierdzeniach matematycznych. Dlatego też w reżimie, w którym układ jest czuły na warunki początkowe, w praktyce niepewność warunków początkowych powoduje niepewność ewolucji czasowej. Można to sprecyzować w matematycznym sensie w następujący sposób:

Niech $x(t)$ będzie trajektorią z warunkiem początkowym $x(0)$, a $X(t)$ będzie trajektorią z warunkiem początkowym $X(0)$. Niech dwa warunki początkowe różnią się o małą wielkość:

$$|X(0) - x(0)| = \epsilon_0 \quad (2.25)$$

gdzie $|\dots|$ oznacza odległość przy zadanej metryce. Jeżeli różnią się o taką wielkość lub mniejszą, wówczas są one dla nas nierozróżnialne. Traktujemy je jako takie same w ramach błędu pomiarowego. Pytamy, jaka jest różnica

$$|X(t) - x(t)| = \epsilon(t) \quad (2.26)$$

po czasie t spowodowana niepewnością warunków początkowych ϵ_0 . Wrażliwość na warunki początkowe oznacza, że 2 trajektorie oddalają się od siebie w bardzo szybkim, eksponencjalnym tempie:

$$\epsilon(t) = \epsilon_0 e^{\lambda t} \quad (2.27)$$

gdzie λ nazywa się wykładnikiem Lapunowa. Jeżeli $\lambda < 0$ to dwie bliskie sobie trajektorie nie oddalają się od siebie. Z kolei jeżeli $\lambda > 0$ to trajektorie rozbiegają się.

Niech δ będzie dokładnością naszej aparatury do rozróżniania trajektorii. Jeżeli $\epsilon(t) > \delta$, czyli gdy dwie trajektorie stają się dla nas rozróżnialne, to nie możemy przewidzieć dalszej ewolucji układu: która trajektoria jest właściwa, $x(t)$ czy $X(t)$? Po jakim czasie nasze przewidywania tracą sens? Po takim czasie, że

$$\epsilon(t) = \epsilon_0 e^{\lambda t} > \delta \quad (2.28)$$

czyli po czasie

$$t > \frac{1}{\lambda} \ln \left[\frac{\delta}{\epsilon_0} \right] \quad (2.29)$$

Założmy, że stan początkowy możemy zadać z dokładnością $\epsilon_0 = 10^{-8}$, a tolerancja $\delta = 10^{-5}$ jest dla nas satysfakcjonująca (choć jest to 3 rzędy wielkości gorzej niż dla warunku początkowego). Jak długo nasze przewidywania są akceptowalne:

$$t_1 \approx \frac{1}{\lambda} \ln \left[\frac{\delta}{\epsilon_0} \right] = \frac{1}{\lambda} \ln \left[\frac{10^{-5}}{10^{-8}} \right] = \frac{1}{\lambda} \ln [10^3] = \frac{3}{\lambda} \ln 10 \quad (2.30)$$

Założmy, że ktoś jest w stanie przygotować stan początkowy ze znacznie lepszą dokładnością, a mianowicie 1000 razy lepiej, tzn. $\epsilon_0 = 10^{-11}$. O ile dłużej możemy przewidywać ewolucję układu:

$$t_2 \approx \frac{1}{\lambda} \ln \left[\frac{10^{-5}}{10^{-11}} \right] = \frac{1}{\lambda} \ln [10^6] = \frac{6}{\lambda} \ln 10 = 2t_1 \quad (2.31)$$

To jest zaledwie 2 razy dłuższy czas!! Widać, że gdy układ jest w reżimie chaotycznym, przewidywalność czasowa jest bardzo ograniczona. Zwiększanie dokładności wyznaczania warunków początkowych 1000-krotnie powoduje wydłużenie czasu przewidywalności zaledwie 2 razy. To jest właśnie problem z prognozą pogody. Możemy zwiększać sieć punktów pomiarowych, a i tak przewidywania pogody są rozsądne zaledwie na kilka dni do przodu.

Problem, czy układ wykazuje własności chaotyczne czy nie, nie jest łatwy do stwierdzenia. Ponieważ układ równań różniczkowych zwykle nie można analitycznie rozwiązać, trzeba bazować na metodach komputerowych. Z jednej strony układ jest czuły na warunki początkowe, z drugiej strony sama metoda numeryczna i obliczenia komputerowe obciążone są błędami, których nie można wyeliminować. Może zdarzyć się, że to nie własność układu a artefakty komputerowe wytwarzają złudzenie chaosu. Trzeba na to być czułym. Obecnie istnieją dobre programy komputerowe uwzględniające niedoskonałości o których mowa. Ponadto istnieją kilka charakterystyk, które mają specyficzne własności dla układów chaotycznych. Oto te charakterystyki:

1. Wykładniki Lapunowa λ_i
2. Widmo (spektrum) mocy $P(\omega)$
3. Funkcja korelacyjna $C(\tau)$
4. Cięcie Poincarego
5. Entropia Kołmogorowa K

Badanie wszystkich charakterystyk jest uciążliwe i czasochłonne, ale eliminuje możliwość pomyłki w stwierdzeniu chaotyczności. Przedstawimy główne cechy tych wielkości jakie występują w reżimie chaotycznym i niechaotycznym.

2.3.1 Scenariusz podwojenia okresu

Przedstawimy teraz standardowy scenariusz przejścia do chaosu, który nazywa się przejściem do chaosu poprzez podwojenie okresu. Jest uniwersalny scenariusz, występujący zarówno w

układach z ciągłym czasem jaki i w układach dyskretnych. Został potwierdzony w wielu eksperymentach na różnorodnych układach fizycznych.

```
var('x y z')
x0, y0, z0 = -0.5, -0.1, 0
kolor = ['blue', 'red', 'green', 'black']

#model
F = x-x^3
g, w = 0.5, 1

#miejsca podwajania okresu
Akeys = ['$a_1$', '$a_2$', '$a_3$', '$a_4$']
Aval = [0.325, 0.354, 0.357, 0.358]
A = dict(zip(Akeys, Aval))

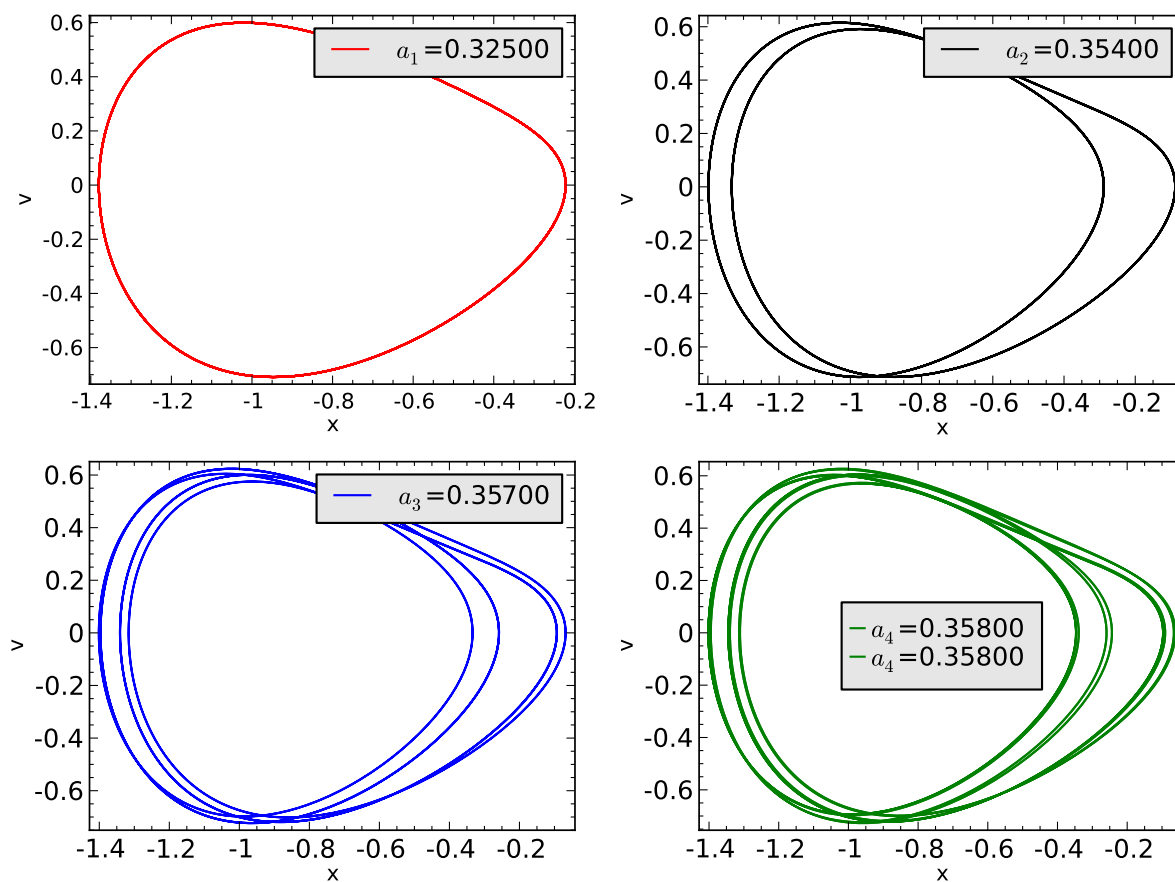
p = A
j=0
for a in A.keys():
    dx, dy, dz = y, F - g*y + A[a]*cos(z), w
    T = srange(0, 100*pi, 0.01)
    num = desolve_odeint(vector([dx, dy, dz]), [x0, y0, z0], T, [x, y, z])
    start, stop = int(len(num[:, 0])*0.8), len(num[:, 0])
    p[a] = list_plot(zip(num[:, 0][start:stop], num[:, 1][start:stop]))
    j+=1
```

Wystarczy teraz tylko narysować wykresy zmagazynowane w liście p.

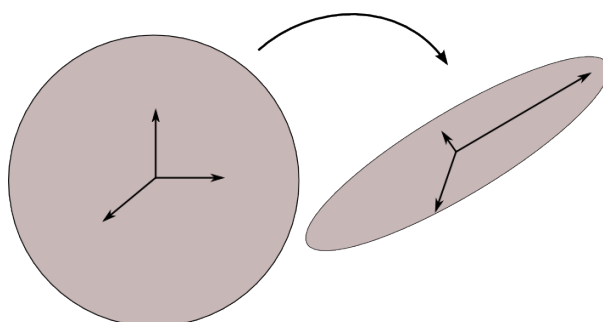
```
bif_p = [0.34357, 0.35506, 0.35785, 0.35846]
i = 2
delta_2 = (bif_p[i-1] - bif_p[i-2]) / (bif_p[i] - bif_p[i-1])
i = 3
delta_3 = (bif_p[i-1] - bif_p[i-2]) / (bif_p[i] - bif_p[i-1])
```

2.3.2 Wykładniki Lapunowa

Dla rozpatrywanego układu oscylatora Duffinga przestrzeń fazowa jest 3-wymiarowa. Dlatego też w rzeczywistości są 3 wykładniki Lapunowa, a nie 1 jak powiedzieliśmy powyżej. Aby wyjaśnić ten problem, musimy rozważyć zbiór warunków początkowych, które tworzą kulę K w badanej przestrzeni fazowej. Jeżeli będziemy iterować równania dla $x(t), y(t), z(t)$ startując z wszystkich warunków początkowych w kuli K , to zbiór punktów zawartych początkowo w kuli zmieni swój kształt. Kula już nie będzie kulą. Prędkość z jaką kula ulega deformacji we wszystkich 3 kierunkach (x, y, z) w przestrzeni fazowej jest określona przez 3 wykładniki Lapunowa $\lambda_1, \lambda_2, \lambda_3$. Jeżeli badany układ jest chaotyczny, to zazwyczaj kula powiększa się w jednym kierunku, a maleje w dwóch pozostałych przyjmując kształt elipsoidy. W takim wypadku możemy zdefiniować trzy wykładniki Lapunowa mierzące deformacje elipsoidy w trzech wzajemnie prostopadłych kierunkach. Ilość wykładników Lapunowa jest więc zależna od wymiaru układu. Są one jednym z kryteriów chaotyczności ruchu. Jeżeli elipsoida w jednym kierunku rozciąga się, wielkość jej osi w tym kierunku rośnie i wykładnik Lapunowa jest dodatnie. W kierunkach, w których osie elipsoidy maleją, wykładniki Lapunowa są ujemne.



Rysunek 2.15: Podwojenie okresu.



Rysunek 2.16: Schematyczna reprezentacja przestrzeni fazowej.

Dwie trajektorie leżące początkowo blisko siebie propagują w czasie w odległości $l(t) \propto e^{\lambda_1 t}$, powierzchnia S zmienia się w tempie $S(t) \propto e^{(\lambda_1 + \lambda_2)t}$, a objętość M zmienia się w tempie $M(t) \propto e^{(\lambda_1 + \lambda_2 + \lambda_3)t}$. W reżimie chaotycznym co najmniej jeden z wykładników Lapunowa jest dodatni. Oznacza to, że w przestrzeni fazowej trajektorie rozbiegają się w jednym kierunku. Jeżeli wszystkie 3 wykładniki są ujemne, układ jest w reżimie regularnym (periodycznym, quasi-periodycznym). Nie ma metod analitycznych pozwalających obliczyć wykładniki Lapunowa. Metody numeryczne też nie są proste. W literaturze można znaleźć algorytmy służące do wyznaczania $\lambda_1, \lambda_2, \lambda_3$.

W przypadku oscylatora Duffinga można otrzymać cząstkowe informacje o wykładnikach Lapunowa.

1. Trzecie równanie dla pomocniczej zmiennej z można rozwiązać otrzymując funkcję

$$z(t) = \omega t + c \quad (2.32)$$

Z pewnością dwie bliskie sobie trajektorie $z_1(t) = \omega t + c_1$ oraz $z_2(t) = \omega t + c_2$ dla chwili $t = 0$ nie rozbiegają się eksponencjalnie ponieważ

$$|z_1(t) - z_2(t)| = |c_1 - c_2| \quad (2.33)$$

Dlatego też jeden z wykładników wynosi zero, np.

$$\lambda_2 = 0 \quad (2.34)$$

2. Przypomnijmy w tym miejscu, że oscylator Duffinga jest opisany przez układ równań

$$\begin{aligned} \dot{x} &= F_1 = y, & x(0) &= x_0, \\ \dot{y} &= F_2 = x - x^3 - \gamma y + A \cos z, & y(0) &= y_0, \\ \dot{z} &= F_3 = \omega, & z(0) &= 0. \end{aligned} \quad (2.35)$$

Zbadajmy, jak zmienia się w czasie objętość fazowa układu. W tym celu musimy obliczyć dywergencję pola wektorowego

$$\operatorname{div} \vec{F} = \frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} + \frac{\partial F_3}{\partial z} = -\gamma < 0 \quad (2.36)$$

Oznacza to, że objętość fazowa w przestrzeni 3-wymiarowej maleje w tempie (zobacz paragraf o układach dyssypatywnych)

$$M(t) \propto e^{-\gamma t} \quad (2.37)$$

Z drugiej strony, jak powiedzieliśmy powyżej,

$$M(t) \propto e^{(\lambda_1 + \lambda_2 + \lambda_3)t} \quad (2.38)$$

Wynika stąd, że suma wszystkich wykładników jest stała i wynosi

$$\lambda_1 + \lambda_2 + \lambda_3 = -\gamma < 0 \quad (2.39)$$

czyli tylko stała tłumienia γ określa tempo malenia objętości fazowej. Ponieważ $\lambda_2 = 0$, otrzymujemy interesujący związek pomiędzy dwoma pozostałymi wykładnikami:

$$\lambda_1 + \lambda_3 = -\gamma \quad (2.40)$$

W reżimie chaotycznym jeden z wykładników jest dodatni, np. $\lambda_1 > 0$ oraz drugi wykładnik musi być ujemny, np. $\lambda_3 < 0$. Mamy obecnie następujące informacje o wykładnikach Lapunowa dla oscylatora Duffinga:

$$\lambda_1 > \lambda_2 > \lambda_3, \quad \lambda_1 > 0, \quad \lambda_2 = 0, \quad \lambda_3 < 0, \quad M(t) = M(0)e^{(\lambda_1 + \lambda_2 + \lambda_3)t} = M(0)e^{-\gamma t} \quad (2.41)$$

Zwracamy uwagę na to, że elipsoida w 3-wymiarowej przestrzeni fazowej rozciąga się w jednym kierunku, kurczy się w drugim kierunku i nie zmienia się w trzecim kierunku oraz objętość elipsoidy cały czas maleje. Tak to wygląda w reżimie chaotycznym. W reżimie niechaotycznym: elipsoida kurczy się w jednym kierunku, kurczy się w drugim kierunku i nie zmienia się w trzecim kierunku oraz objętość elipsoidy cały czas maleje. Atraktory, które pokazywaliśmy poprzednio, istnieją w 3-wymiarowej przestrzeni fazowej, ale ponieważ objętość fazowa cały czas maleje, wymiar atraktorów musi być mniejszy od 3. W reżimie niechaotycznym, n-okresowe atraktory (krzywe) mają wymiar 1. Atraktory w reżimie chaotycznym mają wymiar większy niż 1, ale mniejszy niż 3. Kaplana i Yorke (1979) postawili hipotezę, że istnieje związek pomiędzy wymiarem fraktalnym atraktora D_A a wykładnikami Lapunowa. Relacja ta ma postać:

$$D_A = 2 + \frac{\lambda_1}{|\lambda_3|} > 2 \quad (2.42)$$

Jeżeli analizujemy wymiar atraktora w odwzorowaniu Poincarego (na płaszczyźnie), to wymiar ten jest o 1 mniejszy:

$$d_A = D_A - 1 \quad (2.43)$$

Do dzisiaj jest to tylko hipoteza, choć w wielu przypadkach potwierdzona przez eksperymenty numeryczne.

2.3.3 Widmo mocy

Jest to kolejna wielkość, która może być indykatorem chaotycznego zachowania się układu deterministycznego. Pojęcie widma mocy jest dobrze ugruntowane w teorii sygnałów, traktowanych jako nośnik informacji. W ogólności sygnały mogą być deterministyczne (jak w naszym przypadku) i losowe (stochastyczne). W sensie inżynierskim, sygnał to dowolna funkcja czasu. Jako modele sygnałów wprowadza się również wielkości nazywane dystrybucjami (funkcjami uogólnionymi). Tylko nieliczne proste sygnały można opisać formułami matematycznymi. Większość sygnałów, z jakimi spotykamy się w praktyce, ma przebieg na tyle złożony i nieregularny, że ich bezpośredni opis jako funkcji czasu jest kłopotliwy. Dlatego też należy posługiwać się ich różnego rodzaju reprezentacjami. Reprezentacja sygnału stanowi pewien rodzaj jego symbolicznego opisu, niekiedy o znacznym stopniu abstrakcji. Jej istotą jest to, że zawiera ona pełną informację o sygnale, choć zwykle wyrażoną w innym języku, niż bezpośredni język w terminach funkcji czasu. Oznacza to, że znając sygnał, możemy jednoznacznie wyznaczyć jego reprezentację, znając zaś tę reprezentację – odtworzyć jednoznacznie

sygnał. Istnieje wiele sposobów reprezentacji sygnałów. Jednym z nich jest analiza furierowska za pomocą transformat Fouriera lub szeregów Fouriera.

Przypomnijmy pojęcie transformacji Fouriera funkcji lub dystrybucji. W najprostszym ujęciu transformatą Fouriera $\hat{f}(\omega)$ funkcji $f(t)$ nazywamy całkę

$$\hat{f}(\omega) = \int_{-\infty}^{\infty} e^{i\omega t} f(t) dt \quad (2.44)$$

gdzie ω jest dowolną liczbą rzeczywistą.

Ponieważ nas interesuje ewolucja sygnału czasowego $f(t) = (x(t), y(t), z(t), \dots)$ dla czasów $t > 0$, zdefiniujemy nieco inaczej transformatę Fouriera jako graniczną wartość całki:

$$\hat{f}(\omega) = \lim_{T \rightarrow \infty} \int_0^T e^{i\omega t} f(t) dt \quad (2.45)$$

W praktyce obliczeń komputerowych nigdy nie wykonujemy dokładnej granicy $T \rightarrow \infty$, ale rozpatrujemy dostatecznie długi czas, gdy pojawia się stan ustalony i efekty przejściowe w ewolucji zanikają. Ze względu na występowanie funkcji podcałkowej $e^{i\omega t}$, transformata Fouriera jest funkcją zespoloną. Dlatego też bada się funkcję rzeczywistą w postaci

$$P(\omega) = |\hat{f}(\omega)|^2 \quad (2.46)$$

Nazywa się ona widmem mocy sygnału czasowego $f(t)$. W pewnych przypadkach, faktycznie jest to wielkość fizyczna mająca interpretację mocy, a liczba ω jest częstotliwością, która jest wielkością dodatnią, $\omega > 0$. W dalszym ciągu przyjmiemy to założenie o dodatniości “częstotliwości”. W ogólności, jej związek z mocą (w sensie fizycznym) jest luźny. To widmo mocy jest zdefiniowane inaczej niż w teorii stacjonarnych procesów stochastycznych: tam jest to transformacja Fouriera funkcji korelacyjnej $C(t)$ procesu stochastycznego.

Aby wyrobić sobie intuicję o własnościach transformaty Fouriera i widma mocy, wystarczy rozpatrzyć kilka przypadków funkcji $f(t)$.

Przypadek 1 Jedna harmonika (fala monochromatyczna)

$$f_1(t) = A \cos(\Omega t), \quad \hat{f}_1(\omega) = A \int_0^{\infty} e^{i\omega t} \cos(\Omega t) dt = \frac{\pi}{2} A \delta(\omega - \Omega) \quad (2.47)$$

Transformatą Fouriera jest delta Diraca δ , czyli w widmie mocy pojawia się jeden pik (który w praktyce jest zawsze skończony).

Przypadek 2 Kilka harmonik

$$f_2(t) = \sum_{k=1}^n A_k \cos(\Omega_k t), \quad \hat{f}_2(\omega) = \sum_{k=1}^n A_k \int_0^{\infty} e^{i\omega t} \cos(\Omega_k t) dt = \frac{\pi}{2} \sum_{k=1}^n A_k \delta(\omega - \Omega_k) \quad (2.48)$$

Transformatą Fouriera jest suma przesuniętych delt Diraca δ , czyli w widmie mocy pojawia się szereg pików (które w praktyce są zawsze skończone).

Zauważmy, że dla tak zdefiniowanych transformacji Fouriera nie istnieje widmo mocy, ponieważ w ścisłym sensie matematycznym nie istnieje $\delta^2(\omega - \Omega)$ dla delty Diraca. Jednak nie chodzi o precyzję matematyczną, ale o to że pojawia się pik, który nigdy nie jest nieskończony

jak w delcie Diraca. My jednak potrzebujemy praktycznej metody sprawdzania chaotyczności procesu i zwykle sygnał próbujemy dla dyskretnych wartości czasu t . Dlatego też należy wykorzystać aparat Dyskretnej Transformacji Fouriera, która skończony ciąg sygnału

$$\{x_0, x_1, x_2, \dots, x_{N-1}\} \quad (2.49)$$

przekształca w skończony ciąg amplitud

$$\{A_0, A_1, A_2, \dots, A_{N-1}\} \quad (2.50)$$

odpowiednich harmonik poprzez relacje:

$$A_k = \sum_{n=0}^{N-1} x_n e^{-2\pi i k n / N}, \quad x_n = \frac{1}{N} \sum_{k=0}^{N-1} A_k e^{2\pi i k n / N} \quad (2.51)$$

Dla odpowiednio dużej liczby N (w praktyce rzędu 100), zgodność pomiędzy transformatą Fouriera a Dyskretną Transformatą Fouriera jest zadziwiająco dobra.

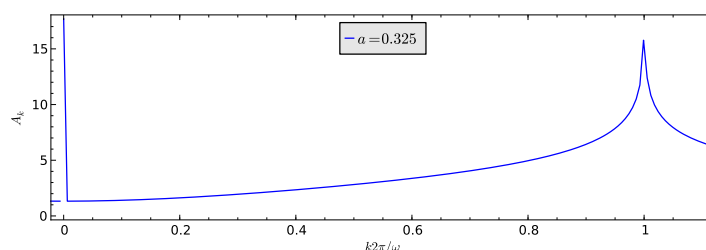
```
var('x y z')
g, w0 = 0.5, 1
x0, y0, z0 = 0.1, 0.1, 0
Aval = [0.325, 0.354, 0.357, 0.358, 0.4]
kolor = ['blue', 'red', 'green', 'black', 'orange']
p = []
j = 0
for a in Aval:
    dx, dy, dz = y, x - x**3 - g*y + a*cos(z), w0
    h, T = 0.1, 1100
    skip = 100
    iskip = int(skip/h)

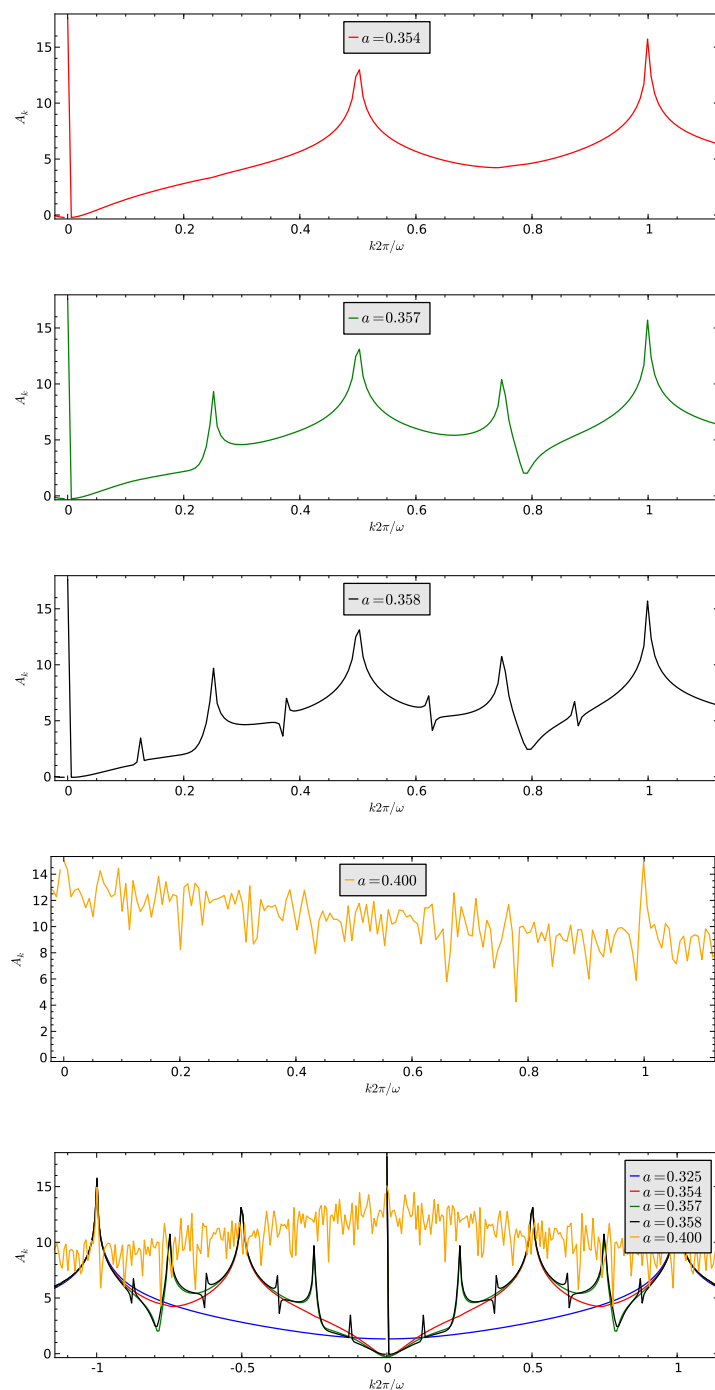
    listT = srange(0, T, h)
    num = desolve_odeint(vector([dx, dy, dz]),
                        [x0, y0, z0], listT, [x, y, z])

    iks = num[:, 0].tolist()[iskip:]
    li = len(iks)
    freq = [i/(T-skip) for i in range(li/2)] + \
           [-li/(T-skip) + i/(T-skip) for i in range(li/2, li)]
    freq = [f*2.*n(pi)/w0 for f in freq]

    vx = vector(iks)
    A = vx.fft().apply_map(lambda x: x.abs2())
    p.append(list_plot(zip(freq, A.apply_map(lambda x: x.log()))))

    j += 1
```





Rysunek 2.17: Podwojenie okresu - widmo mocy.

2.3.4 Funkcja korelacyjna

Jeżeli badamy deterministyczny proces, nie zawsze jest sens mówić o wartości średniej, w takim sensie jak w teorii procesów stochastycznych lub na wykładach z fizyki statystycznej: uśrednianie po realizacjach lub po zespole statystycznym. Ale jeżeli proces deterministyczny jest ergodyczny (trudne pojęcie!), to średnia wartość jest dobrze określona i średnia po zespole

jest równoważna średniej po czasie. Jeżeli dodatkowo proces jest stacjonarny, to można zdefiniować funkcję korelacyjną $C(\tau)$ dla procesu deterministycznego. W naszym przypadku: dla położenia lub prędkości, jest ona zdefiniowana przez relacje:

$$C(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T [x(t+\tau) - \langle x(t+\tau) \rangle][x(t) - \langle x(t) \rangle] dt, \quad (2.52)$$

$$\langle x(t) \rangle = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T x(t) dt.$$

Jeżeli mamy rozwiązanie równania ruchu $x(t)$, to w zależności od postaci tego rozwiązania również SAGE poradzi sobie z rozwiązaniem całki. Jeżeli analityczny wzór będzie poza możliwościami obliczeń symbolicznych, zawsze możemy wygenerować sobie szereg czasowy $x = \{x_1, x_2, \dots\}$. Realizacja funkcji korelacyjnej w SAGE nie będzie stanowić problemu numerycznego. Możemy pokusić się o samodzielne sformułowanie problemu, lub skorzystać z metod pakietu `finance`.

```
def korelator(dane, tau=0):
    ret = None
    if tau == 0:
        return 1
    else:
        tau = abs(tau)
        m = mean(dane)
        dane = [dane[i] - m for i in xrange(len(dane))]
        v = vector(dane)
        sigma = v.dot_product(v)
        if tau < len(dane):
            ret = v[:-tau].dot_product(v[tau:])
        ret /= sigma
    return ret
```

Teraz obliczymy sobie ową funkcję korelacji dla oscylatora Duffinga.

```
var('x y z')
a, g, w0 = 0.3, 0.26, 1
x0, y0, z0 = 0.1, 0.1, 0

dx = y
dy = x - x**3 - g*y + a*cos(z)
dz = w0

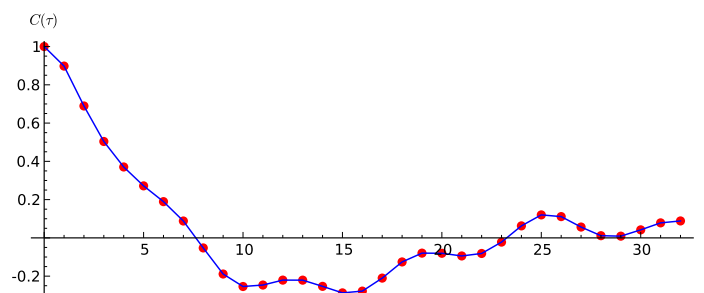
h = 0.1
T = 1000
lT = srange(0, T, float(h), include_endpoint=True)
num = desolve_odeint(vector([dx, dy, dz]), [x0, y0, z0], lT, [x, y, z])
```

Skorzystamy zarówno z naszej funkcji jak i z wbudowanego w SAGE pakietu `finance`, obliczając funkcję (auto)korelacji dla położenia i dla prędkości.

```
#x
dane = num[:, 0].tolist()
```

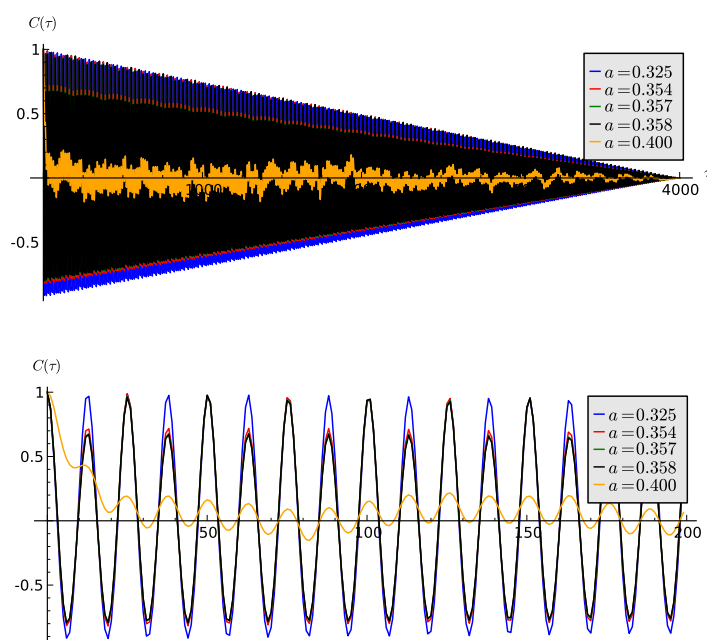
```
# nasz korelator
my_acorr = [korelator(dane,i*10) for i in range(33)]

# funkcja SAGE
v = finance.TimeSeries(dane)
sage_acorr = [v.autocorrelation(i*10) for i in range(33)]
```



Rysunek 2.18: Funkcja korelacyjna.

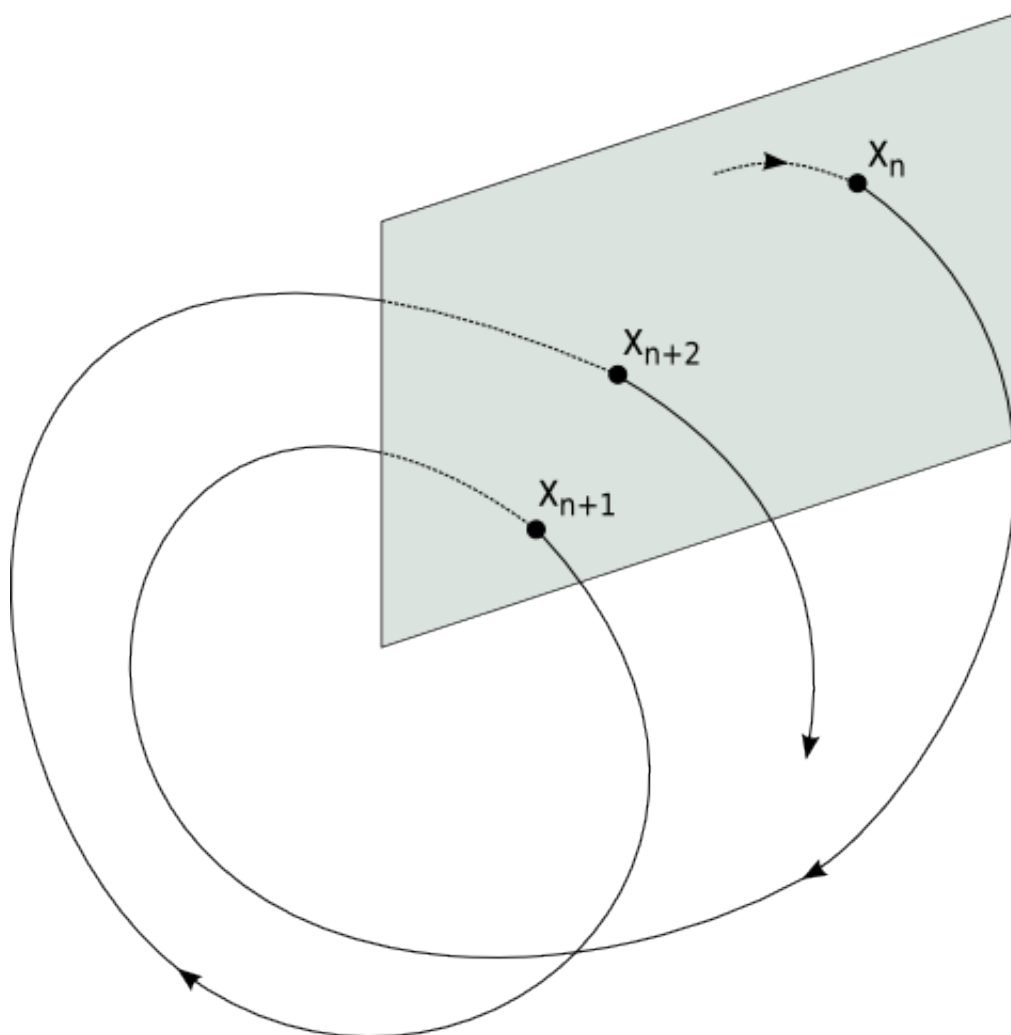
Powyższe rachunki możemy powtórzyć dla wszystkich punktów o których była mowa przy omawianiu bifurkacji.



Rysunek 2.19: Funkcja autokorelacyjna.

2.3.5 Odwzorowanie (cięcie) Poincarego

Odwzorowanie Poincarego jest innym przedstawieniem dynamiki układu. Najprościej jest to wytłumaczyć na przykładzie oscylatora Duffinga. Jego przestrzeń fazowa jest 3-wymiarowa. Ruch w trzecim wymiarze jest jednostajny, $z(t) = \omega_0 t$. Rzut orbity na płaszczyznę (x, y) jest przedstawiony w postaci krzywych fazowych w poprzednich częściach książki.



Rysunek 2.20: Konstrukcja cięcia Poincarego.

Jak widać, we wszystkich przykładach krzywe fazowe na płaszczyźnie są ograniczone na pewnym obszarze (x, y) . We wszystkich rozpatrywanych przypadkach ruch wydaje się być prawie-periodyczny: układ ciągle powraca w te same obszary. Można zbudować następujące przedstawienie tego ruchu.

Okres siły periodycznej wynosi

$$T = \frac{2\pi}{\omega_0} \quad (2.53)$$

Wprowadzamy dyskretny czas

$$t_n = nT, \quad n = 1, 2, 3, \dots \quad (2.54)$$

Zapisujemy położenie i prędkość cząstki w dyskretnych chwilach czasu:

$$x_n = x(t_n), \quad y_n = y(t_n), \quad x(0) = x_0, \quad y(0) = y_0 \quad (2.55)$$

Współrzędne tych punktów nanosimy na płaszczyznę. Otrzymujemy odwzorowanie które nazywamy odwzorowaniem Poincarego. Obrazowo mówiąc można w 3-wymiarowej przestrzeni fazowej wprowadzić płaszczyznę, tak aby nigdzie nie była styczna do trajektorii i była transversalna do trajektorii (ściślej mówiąc do potoku fazowego), czyli aby trajektoria przecinała płaszczyznę, a nie była równoległa do niej (nie omijała jej).

Odwzorowanie Poincarego to przyporządkowanie:

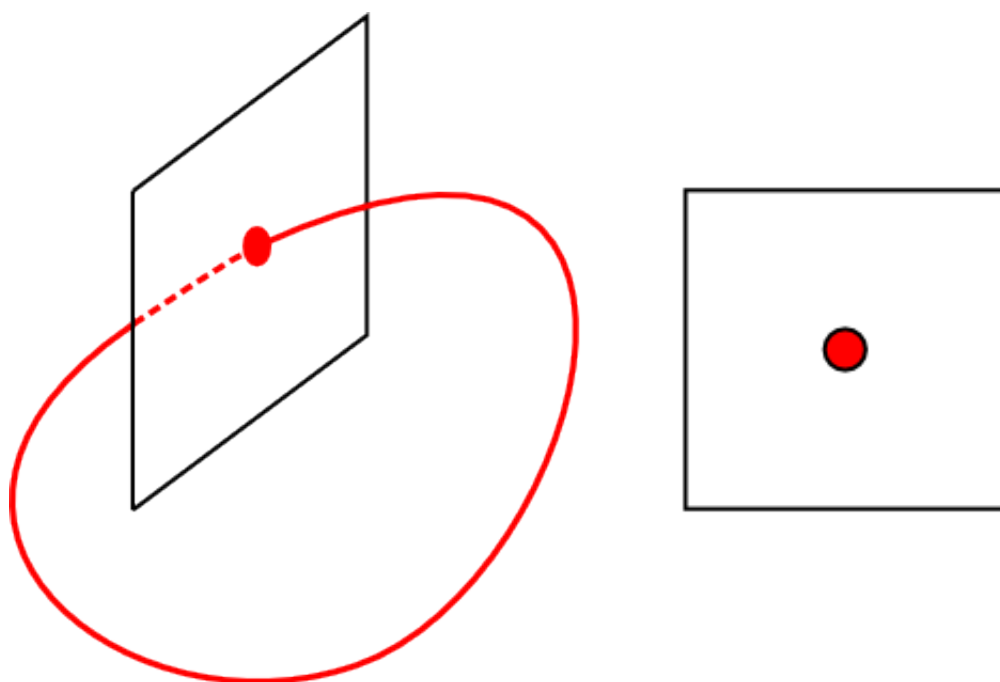
$$x_{n+1} = \mathcal{G}(x_n) \quad (2.56)$$

Jawna konstrukcja tego odwzorowania z wyjściowego układu równań różniczkowych jest możliwa tylko w bardzo specjalnych przypadkach. W przypadku oscylatora Duffinga, nie można otrzymać jawnej postaci tego odwzorowania. Jedynie użycie komputera pozwala na graficzne przedstawienie funkcji \mathcal{G} .

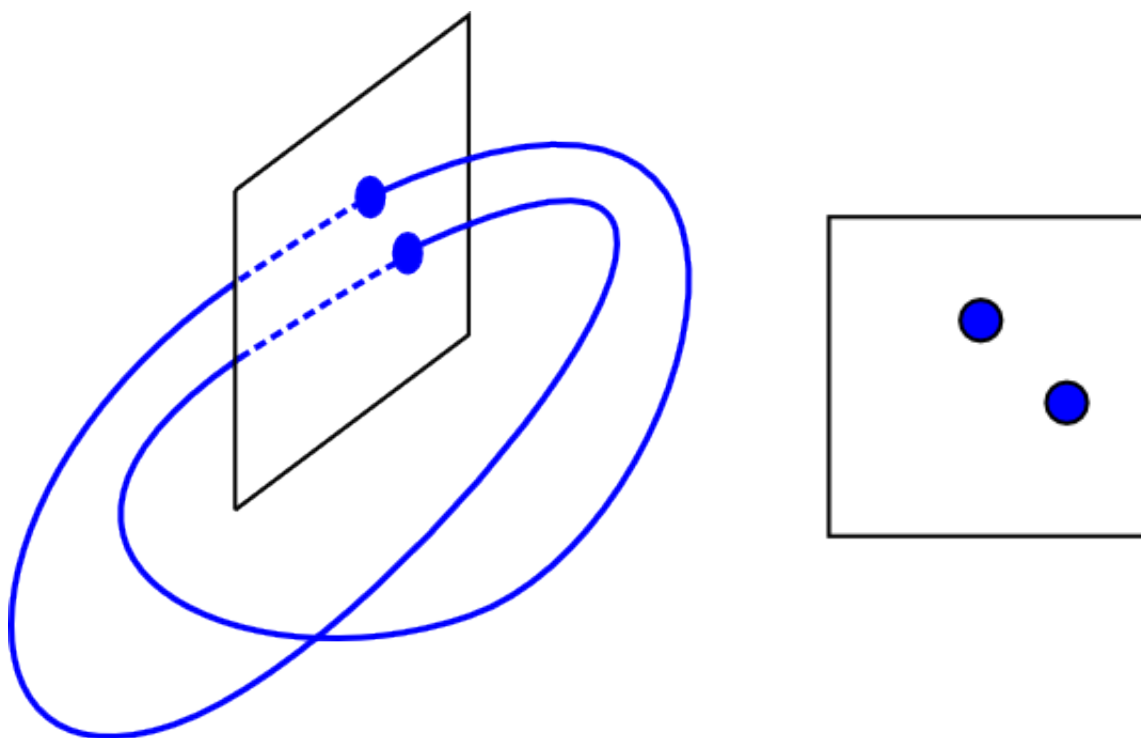
Jakie wnioski płyną z takiego przedstawienia.

1. Gdyby trajektoria była krzywą zamkniętą w kształcie elipsy (atraktor o okresie 1) to na cięciu Poincarego otrzymalibyśmy 1 punkt:
2. Gdyby trajektoria była atraktorem o okresie 2 to na cięciu Poincarego otrzymalibyśmy 2 punkty:
3. Gdyby trajektoria była chaotyczna, to za każdym razem przebiega przez inne punkty płaszczyzny i tworzy zbiór składający się z nieskończenie wielu punktów. Poniżej pokazano takie odwzorowanie dla oscylatora Duffinga.

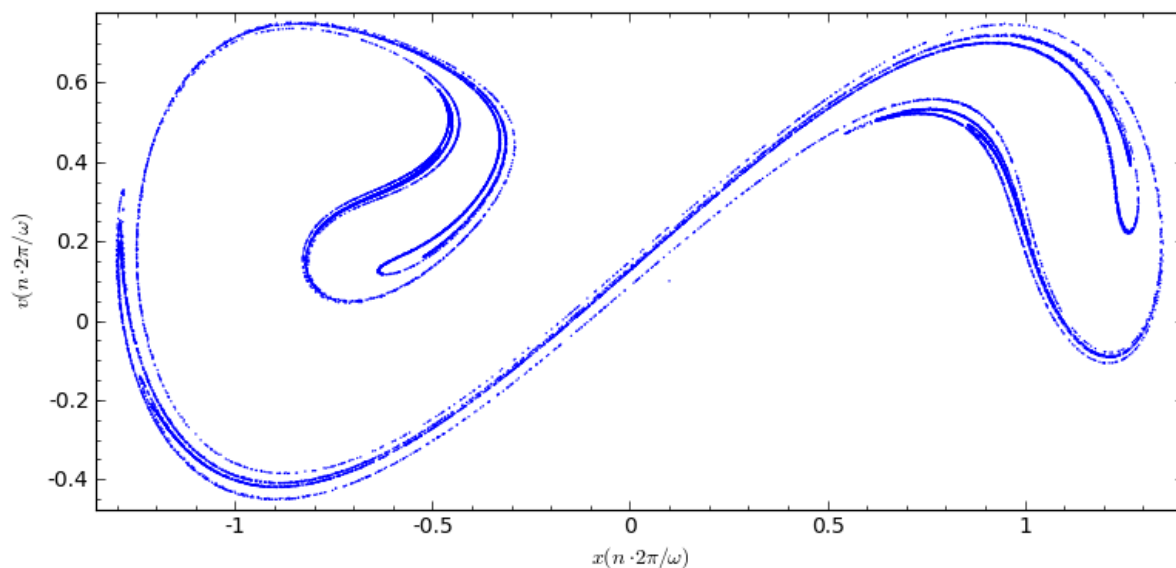
Jeżeli jesteśmy w stanie zbudować graficznie przedstawienie Poincarego danego układu dynamicznego z ciągłym czasem, wówczas możemy rozpoznać takie reżimy które są “podejrzane” o własności chaotyczne. Numerycznie nie powinno nastręczać to większych problemów. Jeżeli znamy ω_0 bądź okres powrotu do obliczenia cięcia to wystarczy wykorzystać poniższy kod Sage. Zwracamy jedynie uwagę na to, że odpowiednio “gęsty” obraz uzyskamy dla bardzo długich przebiegów (dużych T).



Rysunek 2.21: Atraktor o okresie 1.



Rysunek 2.22: Atraktor o okresie 2.



Rysunek 2.23: Dziwny atraktor.

```
# parametry układu równań różniczkowych
a, g = 0.3, 0.26

# częstotliwość (do obliczania cięcia Poincarego)
w0 = 1

# wartości początkowe
x0, y0, z0 = 0.1, 0.1, 0

#układ równań różniczkowych
dx = y
dy = x - x**3 - g*y + a*cos(z)
dz = w0

#krok co jaki wypełniać się ma nasza lista
#rozwiązań ustawiamy równy okresowi
h = 2.0*pi/w0

###
#symulacje
###
T = 10000
lT = srange(0,T,float(h), include_endpoint=True)
sol = desolve_odeint(vector([dx,dy,dz]), [x0,y0,z0],lT,[x,y,z])
```

2.3.6 Przykłady chaosu w Naturze

Należy odróżnić procesy chaotyczne od procesów losowych. Procesy chaotyczne są deterministyczne, a procesy stochastyczne są procesami losowymi. Procesy chaotyczne są badane

przez matematyków, fizyków, chemików, biologów, socjologów, meteorologów, astrofizyków, w teorii informacji i neurobiologii. We wszystkich tych gałęziach nauki, występują deterministyczne modele wykazujące własności chaotyczne. Od lat 60-tych XX wieku opublikowano tysiące prac na temat układów chaotycznych. Matematycy mówią, że prawie wszystkie układy dynamiczne są chaotyczne, a tylko nieliczne układy nie wykazują tej własności. Matematycy dowodzą, że przestrzeń fazowa układu modelowanego przez autonomiczny układ równań różniczkowych musi być co najmniej 3-wymiarowa, aby istniał chaos. Dla układów dyskretnych nie ma takich ograniczeń: jedno równanie rekurencyjne $x_{n+1} = f(x_n)$ także wykazuje własności chaotyczne.

Poniżej podajemy kilka przykładów rzeczywistych zjawisk wykazujących własności chaotyczne.

1. Dynamika cieczy i turbulencja
2. Lasery
3. Układy elektroniczne
4. Plazma
5. Reakcje chemiczne

Na stronie internetowej Wikipedii z hasłem *Chaos Theory* można znaleźć dalsze przykłady oraz podstawowe prace na ten temat. Na zakończenie tej części książki musimy wspomnieć o człowieku, który to wszystko zapoczątkował w 1961 roku. Był to Edward Lorenz, matematyk i meteorolog amerykański, który analizował jeden z najprostszych modeli pozwalających przewidywać pogodę. To z jego nazwiskiem związany jest “efekt motyla” obrazujący niezwykłą czułość dynamiki na zaburzenia warunków początkowych: czy ruch motyla w Brazylii może spowodować tornado w Teksasie (ściśle rzecz ujmując to Philip Merilees zasugerował Lorenzowi taki tytuł wykładu podczas posiedzenia American Association for the Advancement of Science w 1972 roku). W tym obrazowym powiedzeniu zawarta jest istota chaosu: Motyl poprzez swój lot zaburza lokalnie ruch powietrza. Ten zaburzony ruch powietrza narasta i powoduje coraz to większe zmiany pogodowe, zmienia radykalnie “trajektorię” doprowadzając do tornada, które pojawi się nad Teksasem. Czy faktycznie motyl może być taki groźny?

Dynamika stochastyczna

3.1 Równania stochastyczne i ich interpretacja

Dotychczas rozważaliśmy modele deterministyczne bazujące na równaniach różniczkowych zwyczajnych. Twierdzenia matematyczne dają nam pewność że przy odpowiednich założeniach rozwiązania równań różniczkowych są jednoznaczne przy zadanych warunkach początkowych. To jest kluczowe dla szerokiej klasy zjawisk w przyrodzie. Ten determinizm pozwala przewidywać ewolucję układów, pozwala konstruować i używać urządzenia, które pracują zgodnie z jego planowanymi funkcjami. To zapewniają twierdzenia matematyczne. Ale też pokazaliśmy, że matematyczny determinizm może być w praktyce złudny ponieważ dokładne w sensie matematycznym przygotowanie układu w określonych stanach początkowych jest niemożliwe. Niedokładności warunków początkowych mogą w trakcie ewolucji narastać powodując utratę przewidywalności ukrytą w równaniach różniczkowych. Dlatego ten fenomen nazywa się deterministycznym chaosem, czyli czymś co jest nieprzewidywalne ale jednocześnie nie jest losowe. Teraz rozpoczynamy wędrówkę po innej klasie zjawisk i procesów dynamicznych, a mianowicie po krainie procesów losowych. Ta klasa zjawisk bazuje na teorii procesów stochastycznych. To jest teoria matematyczna, której fundamenty oparte są o teorię prawdopodobieństwa. Niewątpliwie procesy stochastyczne nie są procesami deterministycznymi. Ich opis używa takich pojęć jak wartość średnia, fluktuacje czyli niespodziewane odchylenia od wartości średniej, korelacja w różnych momentach czasu, charakterystyki spektralne ważne z eksperymentalnego punktu. Opis procesów stochastycznych jest do pewnego stopnia podobny do przedstawionego wcześniej opisu: bazuje on też na równaniach różniczkowych w których pojawiają się wyrażenia mogące przyjmować w sposób losowy różne wartości. Aby to wyjaśnić, posłużmy się przykładem z klasycznej mechaniki Newtona. Ba, możemy przywołać tu równanie Newtona dla cząstki o jednym stopniu swobody w potencjale $V(x)$. Jest ono postaci:

$$m\ddot{x} + \gamma\dot{x} = -V'(x) \quad (3.1)$$

W równaniu tym pojawia się wyraz $\gamma\dot{x}$, który uwzględnia tarcie, tłumienie. Skąd pojawia się tarcie (tłumienie) w układzie? Wywodzi się ono z oddziaływania układu z otoczeniem. Rzeczywiste układy zawsze są w jakimś ośrodku: w powietrzu, wodzie, kryształach, komórce biologicznej. Owo oddziaływanie z otoczeniem jest źródłem tarcia, ale też jest jednym z czynników powodujących losowość. Wpływ otoczenia nie da się przewidzieć ponieważ rzeczywiste

otoczenie składa się w bardzo dużej ilości cząstek. Przypomnijmy że w szkance wypełnionej wodą mamy około 10^{23} cząstek. Nie znamy ani położenia tych cząstek, ani prędkości tych cząstek. Ale gdybyśmy nawet znali te wielkości, to co nam z takiej wiedzy. Gdybyśmy chcieli modelować oddziaływanie układu z otoczeniem jako oddziaływanie opisane przez równania Newtona, to liczba równań Newtona rzędu 10^{23} jest przerażająco wielka. Nie bylibyśmy w stanie je zapisać, nie mówiąc o ich rozwiązywaniu. W tej beznadziejnej sytuacji powstaje fizyka statystyczna i teoria procesów stochastycznych. W tych teoriach nie ma determinizmu. Musi nam wystarczyć tylko prawdopodobieństwo tego, że cząstka w chwili $t = 17$ znajduje się w jakimś obszarze A , że średnia wartość prędkości w chwili $t = 32s$ wynosi $36m/s$ oraz że cząstka średnio oddaliła się o $x = 2m$ od położenia początkowego.

W powyższym równaniu Newtona uwzględniliśmy wyraz opisujący tarcie (czyli oddziaływanie z otoczeniem), ale nie uwzględniliśmy jeszcze jednej siły, siły jaką cząstki otoczenia wywierają na układ. Poprawniejsze równanie miałoby postać:

$$m\ddot{x} + \gamma\dot{x} = -V'(x) + \Gamma(t) \quad (3.2)$$

Siła $\Gamma(t)$ opisuje proces oddziaływania (zderzeń) cząsteczek otoczenia z układem. Dla przykładu, możemy rozpatrywać kulkę polimerową o promieniu mikronów (10^{-6} metra) zanurzoną w wodzie. Taka kulka bez przerwy zderza się z cząsteczkami wody i porusza się w nieprzewidywalny sposób. Co możemy powiedzieć o sile $\Gamma(t)$, która z pewnością jest losowa.

1. Po pierwsze, żaden kierunek nie jest uprzywilejowany: kulka jest atakowana ze wszystkich stron i średnia siła jest zerowa:

$$\langle \Gamma(t) \rangle = 0 \quad (3.3)$$

2. Po drugie, siła w chwili t_1 nie zależy od siły w chwili t_2 , czyli nie jest ona skorelowana w różnych momentach czasu:

$$\langle \Gamma(t_1)\Gamma(t_2) \rangle = \langle \Gamma(t_1) \rangle \langle \Gamma(t_2) \rangle = 0 \quad (3.4)$$

3. Po trzecie, $\Gamma(t)$ powinna być wielkością losową o rozkładzie normalnym, co wynika z centralnego twierdzenia granicznego. Mówiąc prosto, jeżeli jakaś wielkość jest wynikiem bardzo wielu drobnych losowych czynników, to niezależnie od rozkładu każdego z tych czynników, jej rozkład będzie zbliżony do normalnego. W rzeczywistości taki rozkład nie jest ściśle realizowany, ale jest to przybliżenie dobrze oddające charakter oddziaływania otoczenia na układ. Pamiętajmy, że centralne twierdzenie graniczne jest jednym z najważniejszych twierdzeń rachunku prawdopodobieństwa, które uzasadnia powszechne występowanie w przyrodzie rozkładów zbliżonych do rozkładu normalnego.

Modelem siły losowej może być tzw. biały szum gaussowski. Jest to losowy proces, zupełnie taki jak ten opisany powyżej, różni się tylko tym, że funkcja korelacyjna ma postać:

$$\langle \Gamma(t_1)\Gamma(t_2) \rangle = 2D\delta(t_1 - t_2) \quad (3.5)$$

gdzie D jest natężeniem szumu, które zależy od temperatury T : $D = D(T)$. W wyższej temperaturze energia kinetyczna cząstek wody jest większa i z większą siłą cząsteczki wody uderzają w kulkę. Z własności dystrybucji delta Diraca wynika własność (3.4). Jedną własność

może wydawać się nieuzasadniona, a mianowicie dla tych samych chwil czasu $t_1 = t_2$ delta Diraca jest nieskończona, co oznacza że drugi moment statystyczny nie istnieje (jak mówią fizycy, jest nieskończony). Ale to nie jest aż takie kłopotliwe. Okazuje się bowiem że równanie Newtona z taką siłą losową, która jest białym szumem gaussowskim, jest w zgodzie z fizyką statystyczną, w szczególności stan stacjonarny $p(x, \dot{x})$ opisany równaniem (3.2) jest stanem równowagi termodynamicznej określonym przez rozkład kanoniczny Gibbsa:

$$p(x, \dot{x}) = N_0 \exp \left[-\frac{1}{k_B T} \left(\frac{m \dot{x}^2}{2} + V(x) \right) \right] \quad (3.6)$$

gdzie N_0 jest stałą normalizacyjną. Otrzymanie powyższej gęstości prawdopodobieństwa z równania (3.2) nie jest zadaniem łatwym. W dalszych częściach postaramy się pokazać, dla-czego rów. (3.6) wynika z rów. (3.2).

Rozpatrzmy teraz tzw. reżim przetłumiony, czyli przypadek silnego tłumienia. Jeżeli tarcie jest duże, to trudno jest (prawie niemożliwe) doświadczalnie wyznaczyć przyspieszenie cząstki. Innymi słowy wyraz $m\ddot{x}$ można w równaniu (3.2) zaniedbać. W takim reżimie równanie (3.2) redukuje się do postaci:

$$\gamma \dot{x} = -V'(x) + \Gamma(t) \quad (3.7)$$

lub

$$\dot{x} = -\tilde{V}'(x) + \tilde{\Gamma}(t) \quad (3.8)$$

gdzie przeskalowaliśmy potencjał i siłę losową: podzieliliśmy obustronnie przez stałą γ i zdefiniowaliśmy nowe funkcje

$$\tilde{V}(x) = \frac{1}{\gamma} V(x), \quad \tilde{\Gamma}(t) = \frac{1}{\gamma} \Gamma(t) \quad (3.9)$$

Równanie (3.8) jest wyjściowym równaniem do dalszych rozważań i uogólnień. Będziemy badali nieco uogólniejszą postać tego równania, a mianowicie

$$\frac{dX}{dt} = F(X) + G(X)\Gamma(t) \quad (3.10)$$

gdzie $X = X(t)$ jest procesem zależnym od czasu, funkcje $F(X)$ i $G(X)$ są (raczej) dowolnymi funkcjami oraz $\Gamma(t)$ jest białym szumem gaussowskim określonym przez związki (3.3) i (3.5). Fizycy nazywają rów. (3.10) równaniem Langevina. Matematycy preferują inny zapis tego równania, a mianowicie

$$dX = F(X)dt + G(X)dW(t), \quad dW(t) = \Gamma(t)dt \quad (3.11)$$

gdzie proces losowy $W(t)$ nazywa się procesem Wienera. Można by zapisać relację

$$\Gamma(t) = \frac{dW(t)}{dt} \quad (3.12)$$

choć matematycy dowodzą, że pochodna nie istnieje w żadnym rozsądnym sensie, co nie przeszkadza fizykom wykorzystywanie tej relacji, głównie w celach rachunkowych. Po raz kolejny uogólnimy równanie (3.11) do postaci

$$dX = F(X)dt + G(X)d\xi(t) \quad (3.13)$$

gdzie $\xi(t)$ jest jakimś dopuszczalnym procesem losowym nazywanym szumem, losowymi fluktuacjami, zaburzeniem przypadkowym lub procesem stochastycznym. Wszystkie te nazwy będziemy używali zamiennie. Równanie (3.11) nazywa się równaniem Ito. Równanie (3.13) też będziemy nazywali równaniem Ito, a równanie

$$\frac{dX}{dt} = F(X) + G(X)\eta(t) \quad (3.14)$$

nazywać będziemy równaniem Langevina. W równaniu tym $\eta(t)$ jest jakimś możliwym procesem losowym. Okazuje się, że równanie (3.13) lub (3.14) nie jest jednoznacznie zdefiniowane jeżeli funkcja $G(X)$ zależy od X . Gdy $G(X)$ jest funkcją stałą, równanie jest dobrze określone.

Dlaczego pojawia się niejednoznaczność w interpretacji tych równań? Przyczyną tego jest proces losowy występujący w ostatnim wyrazie. Niewinnie wyglądająca różniczką $dW(t)$ lub $d\xi(t)$ to przyrost procesu losowego:

$$dW(t) = W(t + dt) - W(t), \quad d\xi(t) = \xi(t + dt) - \xi(t) \quad (3.15)$$

co oznacza, że powinniśmy znać własności procesu W oraz ξ w różnych chwilach czasu. Ponadto z punktu widzenia matematyki, powyższe równania różniczkowe są umownym zapisem całkowej wersji tych równań:

$$X(t) - X(t_0) = \int_{t_0}^t F(X(s), s)ds + \int_{t_0}^t G(X(s), s)d\xi(s) \quad (3.16)$$

które powstaje przez obustronne całkowanie w przedziale czasu $[t_0, t]$. Otrzymujemy równanie całkowe dla procesu $X(t)$. W równaniu tym pojawiają się dwa typy całek: “tradycyjna” całka Riemanna-Stieltjesa

$$I_1 = \int_{t_0}^t F(X(s), s)ds \quad (3.17)$$

oraz całka, w której występuje proces $\xi(t)$:

$$I_2 = \int_{t_0}^t G(X(s), s)d\xi(s) \quad (3.18)$$

Powinniśmy zawsze pamiętać o tym, że całka jest graniczną wartością odpowiedniej sumy. I tak pierwsza całka

$$I_1 = \int_{t_0}^t F(X(s), s)ds = \lim_{n \rightarrow \infty} \sum_{i=0}^{n-1} F(X(\tilde{s}_i), \tilde{s}_i)[s_{i+1} - s_i] \quad (3.19)$$

gdzie granicę należy rozumieć w sensie średniokwadratowym oraz $\tilde{s}_i \in [s_i, s_{i+1}]$ jest dowolną liczbą z danego przedziału $[s_i, s_{i+1}]$. W kursie analizy matematycznej wykazuje się, że graniczna wartość sumy (czyli wartość całki) nie zależy od tego gdzie wybieramy wartość $X(\tilde{s}_i)$ dla \tilde{s}_i w przedziale $[s_i, s_{i+1}]$. Może ona leżeć w lewym końcu przedziału, w prawym końcu przedziału, w środku lub każdym innym punkcie tego przedziału. Okazuje się, że w ogólności tej własności nie ma drugi typ całki!! W takim razie w jakim punkcie przedziału należy wybrać wartość $\xi(\tilde{s}_i)$ w całce, w której pojawia się proces $\xi(t)$? Ogólnej recepty na to nie ma. W literaturze istnieją 2 przepisy, gdzie ma leżeć $\xi(\tilde{s}_i)$.

3.1.1 Całka Ito

W tej definicji (preferowanej przez matematyków) wybiera się wartość $W(s_i)$ z lewej strony przedziału z czysto praktycznej przyczyny (ułatwia to rachunki). Aby wyjaśnić dlaczego tak się postępuje, rozpatrzmy nieco inną całkę z procesem Wienera, a mianowicie

$$I_3 = \int_{t_0}^t H(W(s), s) dW(s) = \lim_{n \rightarrow \infty} \sum_{i=0}^{n-1} H(W(s_i), \tilde{s}_i) [W(s_{i+1}) - W(s_i)] \quad (3.20)$$

Tak określona całka nazywa się całką Ito i ma “przyjazne” własności z tego powodu, że wartości średnie typu

$$\langle H(W(s_i), \tilde{s}_i) [W(s_{i+1}) - W(s_i)]^k \rangle = \langle H(W(s_i), \tilde{s}_i) \rangle \cdot \langle [W(s_{i+1}) - W(s_i)]^k \rangle \quad (3.21)$$

rozbijają się na iloczynny wartości średnich ponieważ proces Wienera jest procesem o niezależnych przyrostach na nieprzekrywających się przedziałach, a wartość średnia iloczynu niezależnych zmiennych losowych jest równa iloczynowi wartości średnich tych zmiennych. Jest to główna przyczyna takiej definicji całek Ito. Należy podkreślić, że dla rzeczywistych procesów losowych taki wybór nie zawsze jest poprawny.

Teraz możemy zdefiniować całkę

$$I_2 = \int_{t_0}^t G(X(s), s) \xi(s) = \lim_{n \rightarrow \infty} \sum_{i=0}^{n-1} G(X(s_i), \tilde{s}_i) [\xi(s_{i+1}) - \xi(s_i)] \quad (3.22)$$

Całki, w definicji których wartości procesu $X(t)$ lub $\xi(t)$ należy brać z lewej strony przedziałów $[s_i, s_{i+1}]$, tzn. dla $G(X(s_i), \tilde{s}_i)$, nazywamy *całkami Ito* lub całkami w interpretacji Ito. Ponieważ jak na razie z czysto matematycznego punktu widzenia wybór punktu z lewej strony przedziału jest arbitralny, każdy inny punkt jest równo uprawniony. Ale należy bezwzględnie pamiętać, że zmiana położenia punktu \tilde{s}_i w przedziale $[s_i, s_{i+1}]$ dla $X(\tilde{s}_i)$ czy dla $\xi(\tilde{s}_i)$ oznacza zmianę wartości całki. To odróżnia całki stochastyczne od “tradycyjnych” całek Riemanna. W związku z tym pojawia się poważny problem, gdy chcemy stosować równania stochastyczne do modelowania realnych zjawisk i procesów. Czy istnieją jakieś racjonalne kryteria na wybór punktu pośredniego \tilde{s}_i ? Dylemat ten przez pewien okres czasu był przedmiotem dyskusji i polemik w literaturze naukowej.

3.1.2 Całka Stratonowicza

Istnieją także inne definicje całek stochastycznych. Druga, konkurencyjna definicja jest następująca:

$$I_o = \int_{t_0}^t G(X(s), s) \circ d\xi(s) = \lim_{n \rightarrow \infty} \sum_{i=0}^{n-1} G\left(\frac{X(s_{i+1}) + X(s_i)}{2}, \tilde{s}_i\right) [\xi(s_{i+1}) - \xi(s_i)] \quad (3.23)$$

gdzie oznaczenie \circ w całce ma informować o tym, że wartość funkcji $G(X(t), t)$ na przedziale $[s_i, s_{i+1}]$ jest brana dla średniej arytmetycznej $[X(s_{i+1}) + X(s_i)]/2$. Tak określona całka nazywa się *całką Stratonowicza* lub całką w sensie Stratonowicza.

Czytelnik łatwo zauważy, że obie całki są szczególnymi przypadkami takiej oto całki:

$$I_{\bullet} = \int_{t_0}^t G(X(s), s) \bullet d\xi(s) = \lim_{n \rightarrow \infty} \sum_{i=0}^{n-1} G(\lambda X(s_{i+1}) + (1 - \lambda)X(s_i), \tilde{s}_i) [\xi(s_{i+1}) - \xi(s_i)] \quad (3.24)$$

gdzie $\lambda \in [0, 1]$ i może przyjmować dowolną wartość z tego przedziału. Szczególne przypadki to: $\lambda = 0$ (definicja Ito); $\lambda = 1/2$ (definicja Stratonowicza).

Uwaga:

1. Istnieją twierdzenia mówiące o tym, że jeżeli proces $x_i(t)$ jest skorelowany, to obie definicje są równoważne. Problem pojawia się tylko wówczas gdy $\xi(t)$ jest procesem stochastycznym o niezależnych przyrostach. Takimi procesami są podstawowe modelowe procesy stochastyczne: proces Wienera, proces Poissona i proces Levy'ego.
2. Jeżeli te trzy procesy są przybliżeniami odpowiednich procesów skorelowanych, to właściwa definicja jest definicją Stratonowicza. Innymi słowy, wyjściowe całki ze skorelowanymi procesami nie zależą od definicji, ale w granicy gdy czas korelacji dąży do zera, wartości całek są takie jak w definicji Stratonowicza.
3. Istnieje związek między całkami Ito i Stratonowicza: z całki Ito można otrzymać całkę Stratonowicza i odwrotnie: z całki Stratonowicza można otrzymać całkę Ito. Więc generalnie nie należy się przejmować interpretacją tak długo jak prowadzimy formalne obliczenia, ale w odpowiednim momencie trzeba wybrać odpowiednią interpretację całki, ponieważ końcowe wyniki zależą od tej interpretacji.

3.1.3 Przypadek wielowymiarowy

W przypadku układu n -równań różniczkowych rozpatrujemy uogólnienie równania (3.13) w postaci

$$dX_i = F_i(X_1, X_2, X_3, \dots, X_n)dt + \sum_{j=1}^n G_{ij}(X_1, X_2, X_3, \dots, X_n)d\xi_j(t), \quad i, j = 1, 2, 3, \dots, n \quad (3.25)$$

lub odpowiednik równania (3.14) ma postać

$$\frac{dX_i}{dt} = F_i(X_1, X_2, X_3, \dots, X_n) + \sum_{j=1}^n G_{ij}(X_1, X_2, X_3, \dots, X_n)\eta_j(t), \quad i, j = 1, 2, 3, \dots, n \quad (3.26)$$

gdzie wielkości losowe $\xi_i(t)$ oraz $\eta_i(t)$ są niezależnymi między sobą procesami stochastycznymi.

Mogą być takie przypadki, gdy wielkości losowe pojawiają się w nieliniowy sposób, np. w postaci

$$\frac{dX}{dt} = F(X, \eta(t)) \quad (3.27)$$

Czytelnik sam może napisać odpowiednik wielowymiarowy równania (3.27). Z punktu widzenia zastosowań ważne jest jakie istnieją modele matematyczne zaburzeń losowych $\{\eta_i(t)\}$ czy $\{\xi_i(t)\}$. Mogą to być procesy stacjonarne, procesy Markowa (markowskie) lub też procesy niemarkowskie. Mogą to być procesy skorelowane lub nieskorelowane. W następnej części podamy przykłady najczęściej stosowanych modeli szumu.

3.2 Podstawowe procesy stochastyczne

W wielkim skrócie przedstawimy 3 podstawowe rodzaje procesów stochastycznych. Ich transformacje służą do definicji szumów, fluktuacji i zaburzeń losowych. Czytelnik znajdzie szczegóły i więcej informacji w podręczniku “Procesy i zjawiska losowe”. W sensie matematycznym, większość podstawowych procesów losowych wywodzi się z prób Bernoulliego i uogólnionych prób Bernoulliego. Można by powiedzieć, że dwie próby Bernoulliego odgrywają kluczową rolę:

1. wielokrotny rzut monetą
2. losowe wybieranie punktów z pewnego przedziału liczbowego

Pierwszy rodzaj prób Bernoulliego w granicy prowadzi do procesu Wienera, drugi - w granicy prowadzi do procesu Poissona. Obie te granice są różne i odpowiadają dwóm klasom procesów stochastycznych. Oba te procesy są szczególnym przypadkiem procesu, który nazywa się procesem Levy’ego. Byłoby trudno wyprowadzić ogólny proces Levy’ego z prób Bernoulliego.

Jako model zaburzeń losowych, w ogólności możemy wybierać dowolny proces stochastyczny. Czasami nakładane są różne ograniczenia jak np. aby był to proces stacjonarny, albo proces o skończonej wartości średniej, albo proces o zerowej wartości średniej, albo proces Markowa, albo proces nieskorelowany, itd. Często zakłada się, że zaburzenie losowe jest o zerowej wartości średniej:

$$\langle \eta(t) \rangle = 0 \quad (3.28)$$

Jeżeli jest to proces stacjonarny to jego funkcja korelacyjna

$$\langle \eta(t)\eta(s) \rangle = C(t-s) \quad (3.29)$$

zależy tylko od różnicy czasów t oraz s . Jeżeli jest to proces nieskorelowany to jego funkcja korelacyjna

$$C(t-s) = 2D_0\delta(t-s) \quad (3.30)$$

modelowana jest przez deltę Diraca $\delta(t)$, gdzie wielkość D_0 nazywa się natężeniem lub intensywnością procesy stochastycznego. Dlaczego jest to proces nieskorelowany? Ponieważ dla dwóch różnych chwil czasu $t \neq s$ delta Diraca $\delta(t-s) = 0$. Oczywiście są inne, bardziej ogólne definicje procesu nieskorelowanego, ale dla naszych potrzeb powyższa definicja wykorzystująca deltę Diraca jest wystarczająca.

Z kolei przez szum rozumiemy zwykle stacjonarny proces stochastyczny. Dla procesu stacjonarnego możemy wprowadzić pojęcie widma mocy szumu (gęstość spektralna szumu) $S(\omega)$ jako transformata Fouriera jego funkcji korelacyjnej, tzn.

$$S(\omega) = \int_{-\infty}^{\infty} e^{i\omega\tau} C(\tau) d\tau \quad (3.31)$$

gdzie wprowadziliśmy oznaczenie $\tau = t - s$.

Gdy funkcja korelacyjna jest deltą Diraca to widmo mocy

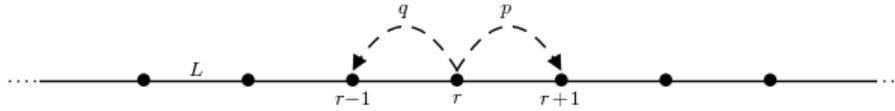
$$S(\omega) = \int_{-\infty}^{\infty} e^{i\omega\tau} 2D_0 \delta(\tau) d\tau = 2D_0 \quad (3.32)$$

nie zależy od częstości ω . W optyce widmo mocy niezależne od częstości jest charakterystyczne dla światła białego (jest to mieszanka wszystkich barw, czyli wszystkich częstości). Dlatego taki szum nazywa się białym szumem. Jeżeli szum jest skorelowany, to niektórzy autorzy nazywają taki szum szumem kolorowym, aby go odróżnić od szumu białego. Często taki kolorowy szum jest scharakteryzowany przez czas korelacji i jego widmo mocy zależy od częstości ω . Należy jednak powiedzieć, że szum skorelowany i szum kolorowy są synonimami tego samego pojęcia.

W Naturze biały szum nie występuje. Jest on kolejną idealizacją opisu rzeczywistych procesów. Jest to dobra idealizacja gdy czas korelacji szumu jest najmniejszym charakterystycznym czasem w badanym układzie. Poniżej prezentujemy kilka najbardziej popularnych modeli procesów stochastycznych oraz białego i kolorowego szumu losowego.

3.2.1 Proces Wienera

Dyskretny błądzenie przypadkowe jest jedną z realizacji wielokrotnego rzutu monetą: w losowy sposób cząstka błądząca wykonuje krok o długości L w prawo z prawdopodobieństwem p lub krok o długości L w lewo z prawdopodobieństwem q , przy czym $p + q = 1$. Cząstka wykonuje kroki w odstępach czasu T .



Rysunek 3.1: Błądzenie przypadkowe.

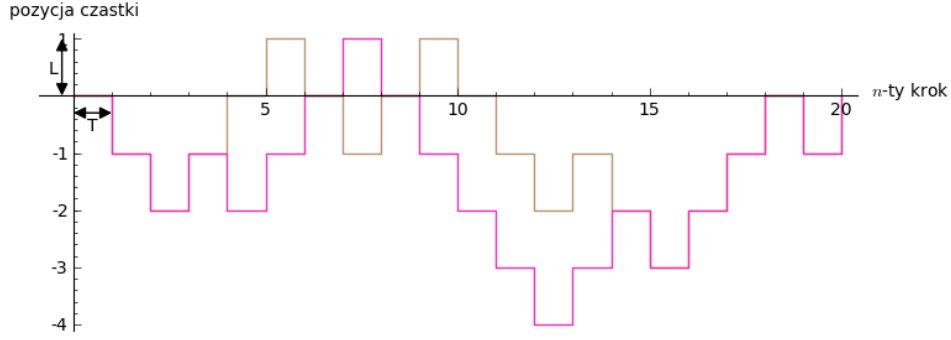
Pytamy, jakie jest prawdopodobieństwo tego, że po wykonaniu n kroków, cząstka będzie w położeniu $r = \{-n, -n + 1, -n + 2, \dots, 0, 1, 2, \dots, n - 1, n\}$; dokładniej mówiąc, po czasie $t = nT$ będzie w położeniu $x = rL$. W granicy

$$L \rightarrow 0, \quad T \rightarrow 0, \quad \text{w taki sposób, aby} \quad \frac{L^2}{T} = \text{const.} = 2D \quad (3.33)$$

gdzie stałą D nazywa się współczynnikiem dyfuzji lub natężeniem (intensywnością) procesu stochastycznego.

Innymi słowy, cząstka wykonuje coraz to mniejsze kroki z coraz to większą częstotliwością. W granicy otrzymujemy ciągły proces błądzenia przypadkowego. Jeżeli prawdopodobieństwa $p = q = 1/2$, błądzenie jest symetryczne i nazywa się procesem Wienera. Ma on następujące własności:

1. Proces stochastyczny Wienera $W(t)$ jest procesem rzeczywistym.



Rysunek 3.2: Dwie realizacje błędzenia przypadkowego.

2. $W(0) = 0$ (proces startuje z zera, ale to nie jest istotne; jest to wygodne).
3. Proces $W(t)$ ma stacjonarne i niezależne przyrosty na nieprzekrywających się przedziałach.

Oznacza to, że 2 przyrosty procesu Wienera $W(t_4) - W(t_3)$ oraz $W(t_2) - W(t_1)$ są niezależne dla dowolnych chwil czasu takich że $t_1 \leq t_2 \leq t_3 \leq t_4$. Innymi słowy wartość średnia

$$\langle [W(t_4) - W(t_3)][W(t_2) - W(t_1)] \rangle = \langle W(t_4) - W(t_3) \rangle \langle W(t_2) - W(t_1) \rangle \quad (3.34)$$

4. $W(t)$ jest procesem Gaussa o zerowej wartości średniej

$$\langle W(t_2) - W(t_1) \rangle = 0 \quad (3.35)$$

i wariancji przyrostów

$$\langle [W(t_2) - W(t_1)]^2 \rangle = 2D(t_2 - t_1), \quad t_2 > t_1 \quad (3.36)$$

Korzystając z własności 3 można obliczyć funkcję korelacyjną procesy Wienera:

$$\langle W(t_2)W(t_1) \rangle = 2D\min(t_2, t_1) = 2D[t_1\theta(t_2 - t_1) + t_2\theta(t_1 - t_2)] \quad (3.37)$$

gdzie funkcja $\min(t, s)$ oznacza mniejszą z 2 wartości t i s . Funkcję tę można wyrazić w matematycznym zapisie korzystając z funkcji skokowej Heaviside'a $\theta(x)$. Przyrost $W(t_2) - W(t_1)$ jest zmienna losową gaussowską o zerowej wartości średniej i wariancji $\sigma^2 = 2D(t_2 - t_1)$. Więc jego rozkład prawdopodobieństwa ma postać

$$f_{W(t_2)-W(t_1)}(x) = \frac{1}{\sqrt{4\pi D(t_2 - t_1)}} \exp \left[-\frac{x^2}{4D(t_2 - t_1)} \right] \quad (3.38)$$

Przyjmując $t_1 = 0$ oraz $t_2 = t$ otrzymamy gęstość prawdopodobieństwa w postaci

$$f_{W(t)}(x) = f(x, t) = \frac{1}{\sqrt{4\pi Dt}} \exp \left[-\frac{x^2}{4Dt} \right] \quad (3.39)$$

Funkcja charakterystyczna $C(\omega, t)$ procesu Wienera ma postać:

$$C(\omega, t) = \langle e^{i\omega W(t)} \rangle = \int_{-\infty}^{\infty} e^{i\omega x} f(x, t) dx = e^{-Dt\omega^2} \quad (3.40)$$

Prawdopodobieństwo tego, że w chwili t cząstka jest w przedziale $[a, b]$ dane jest przez wzór

$$Pr\{W(t) \in (a, b)\} = \int_a^b f(x, t) dx = \frac{1}{\sqrt{4\pi Dt}} \int_a^b \exp\left[-\frac{x^2}{4Dt}\right] dx \quad (3.41)$$

Czytelnik zauważy, że niekonsekwentnie piszemy czasami przedział domknięty $[a, b]$, a czasami przedział otwarty (a, b) . W tym przypadku jest to bez różnicy ponieważ

$$Pr\{W(t) \in (a, b)\} = Pr\{W(t) \in [a, b]\} = Pr\{W(t) \in [a, b)\} = Pr\{W(t) \in (a, b]\} \quad (3.42)$$

Proszę zwrócić uwagę na 4 możliwe przedziały w tych wyrażeniach. Proces Wienera jest granicznym przypadkiem błędzenia losowego: kroki są coraz mniejsze i coraz częstsze. Rozpatrzmy realizację błędzenia przypadkowego w określonym przedziale czasu $[0, t]$. W przedziale tym wybrana realizacja posiada określoną ilość skoków w których funkcja ta jest nieróżniczkowalna. Przy skalowaniu skoki są coraz mniejsze, ale jest ich znacznie więcej. Więc w przedziale czasu $[0, t]$ realizacja posiada znacznie więcej punktów, w których jest nieróżniczkowalna. W granicy, wielkość skoków dąży do zera, ale ich ilość dąży do nieskończoności. Oznacza to, że realizacja staje się funkcją ciągłą (wysokość skoków dąży do zera), ale jednocześnie nigdzie nie jest różniczkowalna (liczba skoków dąży do nieskończoności). Jest to przykład wyjątkowo dziwnej funkcji. Takiej funkcji nie możemy narysować, ale to co opisano powyżej powinno wyrobić w nas intuicję o własnościach realizacji procesu Wienera. Matematycy (jak zwykle) dowodzą to ściśle, a fizycy to czują i wiedzą dlaczego tak jest. Należy także pamiętać, że taki graniczny proces nie istnieje w rzeczywistości. Rzeczywiste procesy błędzenia przypadkowego mają różne długości (ale nie nieskończenie małe) oraz odbywają się z niezerową częstotliwością (T nie jest nieskończenie małe). Jednak gdy T jest najmniejszą skalą czasową w badanym układzie, a każde inne czasy charakterystyczne są znacznie większe, to przybliżenie otrzymane po operacji dokonania granic jest rozsądne. To jest przykład kolejnej idealizacji, jaką stosują fizycy do opisu rzeczywistości.

Proces Wienera jest procesem gaussowskim (normalnym), co jest konsekwencją centralnego twierdzenia granicznego: proces jest wynikiem bardzo wielu niezależnych losowych zdarzeń, dlatego niezależnie od rozkładu prawdopodobieństwa każdego z tych zdarzeń, jego rozkład będzie zbliżony do normalnego.

Poniżej pokazujemy jedną z możliwych realizacji (trajektorii) procesu Wienera. Czytelnik może sam wygenerować inne realizacje, zmieniając parametry pojawiające się w programie.

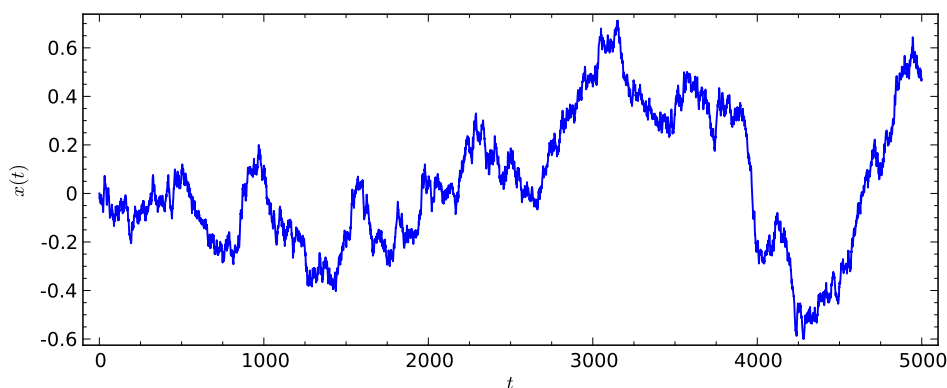
```
# parametry symulacji
h = 0.01 #krok
N = 5000 #ilosc krokow

# parametry SDE
x0 = 0 #wartosc poczatkowa
D = 0.01 #wspolczynnik dyfuzji
```

```

x = [x0]
for i in xrange(1,N):
    n01 = normalvariate(0,1)
    x.append(x[i-1] + sqrt(2*h*D) * n01)

```



Rysunek 3.3: Relizacja procesu Wienera.

3.2.2 Proces Poissona

Rozważamy przedział liczbowy $[0, T]$. Z przedziału tego wybieramy losowo jeden punkt, jedną liczbę. Ponieważ wszystkie liczby są “równo rozłożone”, więc prawdopodobieństwo tego, że punkt ten jest w przedziale $(t_1, t_2) \subset [0, T]$ wynosi

$$P(A) = p = \frac{t_2 - t_1}{T} \quad (3.43)$$

Jeżeli wylosowany punkt jest w tym przedziale, uważam to za sukces. Wybieramy teraz losowo n punktów z przedziału $[0, T]$. Prawdopodobieństwo tego, że k z tych wszystkich n -punktów będzie w przedziale $(t_1, t_2) \subset [0, T]$ jest określone przez rozkład dwumianowy $p_n(k)$. Jeżeli przedział $[0, T]$ będzie przedziałem czasowym, to w przedziale tym mamy n losowych chwil czasu oraz $k \leq n$ losowych chwil czasu w przedziale $(t_1, t_2) \subset [0, T]$. Teraz wykonujemy przejście graniczne:

$$n \rightarrow \infty, \quad p \rightarrow 0 \quad \text{ale} \quad n \cdot p = \frac{n(t_2 - t_1)}{T} = \lambda \quad (3.44)$$

co da się osiągnąć gdy założymy że $T \rightarrow \infty$ przy czym

$$\mu = \frac{n}{T} \quad (3.45)$$

jest stałą wielkością i oznacza ilość losowych punktów w jednostkowym przedziale, czyli gęstość losowo wybranych punktów na osi czasu. Przy takim skalowaniu otrzymamy wyrażenie na parametr

$$\lambda = \mu(t_2 - t_1) \quad (3.46)$$

Przypiszmy teraz każdej losowej chwili czasu akt narodzin dziecka. Prawdopodobieństwo tego że w przedziale (t_1, t_2) jest k -punktów odpowiada prawdopodobieństwu tego że w przedziale

czasu (t_1, t_2) urodziło się k dzieci. W ten sposób otrzymujemy losowy proces urodzin. Elementem losowym jest czas narodzin. Wynikiem jest liczba dzieci $N(t)$ w pewnej chwili czasu t . Rozkład prawdopodobieństwa jest dany przez rozkład Poissona, opisujący prawdopodobieństwo tego, że k dzieci urodzi się w przedziale (t_1, t_2) :

$$Pr\{k, (t_1, t_2)\} = Pr\{N(t_2) - N(t_1) = k\} = e^{-\mu(t_2-t_1)} \frac{[\mu(t_2-t_1)]^k}{k!} \quad (3.47)$$

Przyjmijmy teraz, że $t_1 = 0$ i $t_2 = t$. Wówczas $t_2 - t_1 = t - 0 = t$ oraz

$$p_k(t) = Pr\{N(t) = k\} = Pr\{k, (0, t)\} = e^{-\mu t} \frac{(\mu t)^k}{k!} \quad (3.48)$$

jest prawdopodobieństwem tego, że w przedziale czasu $(0, t)$ urodziło się k dzieci. Jeżeli liczbę urodzonych dzieci oznaczmy przez $N(t)$ to otrzymamy proces stochastyczny nazywany procesem urodzin. Założymy, że $N(0) = 0$. Oczywiście nie jest to bezwzględny wymóg, ale konwencja. Równie dobrze mógłbyśmy badać dowolny inny stan początkowy $N(0)$. Proces urodzin tak jak każda zmienna losowa o rozkładzie Poissona przyjmuje wartości całkowite nieujemne:

$$N(t) = \{0, 1, 2, \dots\} \quad (3.49)$$

Jeżeli $N(t) = k$ to mówimy, że układ jest w stanie k . Takie sformułowanie jest często stosowane przez fizyków, zwłaszcza w kontekście fizyki kwantowej, gdzie stanami mogą być poziomy energetyczne układów kwantowych. Proces urodzin odpowiadałby przejściu z jednego poziomu energetycznego do najbliższego wyższego poziomu energetycznego. Przejście do stanu o niższej energii nie jest opisane procesem urodzin, ale procesem urodzin i śmierci.

Można udowodnić, że przyrosty procesu urodzin $N(t_2) - N(t_1)$ oraz $N(t_4) - N(t_3)$ na *nieprzekrywających się przedziałach* (t_1, t_2) oraz (t_3, t_4) są zmiennymi losowymi niezależnymi. Jest to bardzo ważna własność procesu urodzin o poissonowskiej statystyce losowych chwil urodzin t_i . Przyrost procesu urodzin $N(t_2) - N(t_1)$ jest stacjonarny, to znaczy, że prawdopodobieństwu tego że w przedziale czasu (t_1, t_2) urodziło się k dzieci zależy od różnicy czasów $t_2 - t_1$, a nie zależy od tego gdzie te chwile czasu t_1 i t_2 są na osi czasu. Jeżeli tak jest, to w ogólnym przypadku mówimy, że *proces stochastyczny ma przyrosty stacjonarne*. Uwaga: sam proces nie jest stacjonarny ponieważ rozkład prawdopodobieństwa zależy od czasu ale proces ma stacjonarne przyrosty! Realizacje procesu są *nieciągłymi* i *niemalejącymi* funkcjami czasu i są to funkcje schodkowe o skoku 1, przy czym skok następuje w losowych chwilach czasu t_i (zakładamy, że tylko jeden osobnik rodzi się w danym momencie czasu t_i i w rzeczywistości jest to prawda abstrahując od kwestii spornych przy cesarskim cięciu w przypadku wielodzietnych ciąży). Funkcje schodkowe można zapisać przy pomocy funkcji theta Heaviside'a, więc proces urodzin $N(t)$ można przedstawić w eleganckiej matematycznej postaci:

$$N(t) = \sum_{i=1}^{\infty} \theta(t - t_i) \quad (3.50)$$

Aby otrzymać jedną realizację procesu urodzin w przedziale czasu $[0, T]$, należy losowo wybrać w tym przedziale punkty t_i które są momentami urodzenia. Mówimy, że punkty te mają statystykę Poissona, mając na myśli ich jednorodne (ale losowe) rozłożenie na osi czasu. Mimo że realizacje są funkcjami nieciągłymi (punktami nieciągłości są skoki), proces ten *jest procesem ciągłym* w sensie średnio-kwadratowym! Co to oznacza? Poglądowo oznacza to tyle,

że jeżeli wybiorę jakiś punkt t na osi czasu, to realizacje będą ciągłe w tym punkcie. Inaczej mówiąc, prawdopodobieństwo tego, że punkt t “trafi” w nieciągłość (w skok schodka) jest zerowe! Podsumowując, możemy to wszystko zebrać w matematyczną definicję procesu stochastycznego Poissona (lub poissonowskiego procesu urodzin).

Definicja Procesem Poissona $N(t)$ nazywamy proces stochastyczny o następujących własnościach:

1. Przestrzenią stanów jest zbiór liczb całkowitych nieujemnych, $X = \{k\}_0^\infty = \{0, 1, 2, \dots\}$
2. $N(0) = 0$ (proces startujący z zera)
3. $N(t_2) - N(t_1)$ jest liczbą punktów w przedziale (t_1, t_2)
4. $N(t)$ ma stacjonarne i niezależne przyrosty na nieprzekrywających się przedziałach o rozkładzie prawdopodobieństwa

$$Pr\{N(t_2) - N(t_1) = k\} = e^{-\mu(t_2 - t_1)} \frac{[\mu(t_2 - t_1)]^k}{k!} \quad (3.51)$$

Własności procesu Poissona

1. Wartość średnia

$$m(t) = \langle N(t) \rangle = \mu t \quad (3.52)$$

Z tej relacji mamy interpretację parametru

$$\mu = \frac{\langle N(t) \rangle}{t} \quad (3.53)$$

Parametr μ to średnia liczba punktów w jednostce czasu lub średnia liczba urodzeń w jednostce czasu. Ten parametr może być oszacowany z danych statystycznych czy też z danych doświadczalnych.

2. Moment statystyczny drugiego rzędu

$$\langle N^2(t) \rangle = \mu^2 t^2 + \mu t \quad (3.54)$$

3. Stąd mamy relację dla fluktuacji procesu

$$\sigma^2(t) = \langle N^2(t) \rangle - \langle N(t) \rangle^2 = \mu t \quad (3.55)$$

4. Funkcja charakterystyczna $C(\omega, t)$ procesu Poissona ma postać:

$$\begin{aligned}
 C(\omega, t) &= \langle e^{i\omega N(t)} \rangle = \sum_{k=0}^{\infty} e^{i\omega k} p_k(t) = \sum_{k=0}^{\infty} e^{i\omega k} e^{-\mu t} \frac{(\mu t)^k}{k!} = \\
 &= e^{-\mu t} \sum_{k=0}^{\infty} e^{i\omega k} \frac{(\mu t)^k}{k!} = \exp [\mu t (e^{i\omega} - 1)]
 \end{aligned} \tag{3.56}$$

5. Z p. 3 wynika, że średnio-kwadratowe odchylenie

$$\sigma(t) = \sqrt{\mu t} \tag{3.57}$$

Ważna jest interpretacja tego wyniku: Fluktuacje procesu narastają liniowo w czasie. To oznacza, że im dłuższy jest czas tym odchylenia od wartości średniej mogą być coraz większe. Można przeprowadzić następujące dywagacje:

Badamy proces urodzin i przeprowadzamy szereg symulacji komputerowych tego procesu. Otrzymujemy wiele realizacji procesu urodzin. Z analizy tych realizacji możemy metodami statystyki otrzymać wartość średnią procesu. Oczekujemy (coś to ma wspólnego z wartością oczekiwaną czyli wartością średnią), że wiele realizacji będzie przebiegać w pobliżu wartości średniej. Owszem to prawda, ale jednocześnie pojawiają się realizacje które daleko odbiegają od wartości średniej. Im dłuższy jest ten czas, tym większe odchylenia od wartości średniej mogą zaobserwować. Jest to własność, która w deterministycznym świecie nie występuje (deterministyczny świat nie jest światem rzeczywistym, jest jego idealizacją, bardziej lub mniej adekwatną).

Oczywiście powyższe relacje można podać dla przyrostów procesu Poissona:

1. Wartość średnia przyrostów procesu Poissona

$$\langle N(t_2) - N(t_1) \rangle = \mu(t_2 - t_1) \tag{3.58}$$

2. Moment statystyczny drugiego rzędu dla przyrostów procesu Poissona

$$\langle [N(t_2) - N(t_1)]^2 \rangle = \mu^2(t_2 - t_1)^2 + \mu(t_2 - t_1) \tag{3.59}$$

3. Funkcja korelacyjna procesu Poissona

$$\begin{aligned}
 R(t_2, t_1) &= \langle N(t_2)N(t_1) \rangle = \mu^2 t_2 t_1 + \mu \min(t_2, t_1) = \\
 &= \langle N(t_2) \rangle \langle N(t_1) \rangle + \mu \min(t_2, t_1)
 \end{aligned} \tag{3.60}$$

gdzie funkcja dwóch zmiennych $\min(x, y)$ oznacza wartość mniejszej liczby z dwóch liczb x i y :

$$\min(x, y) = \begin{cases} x & \text{if } x < y \\ y & \text{if } y < x \end{cases} \tag{3.61}$$

Z powyższych relacji wynika, że *proces urodzin jest procesem skorelowanym*.

Pokażemy, w jaki sposób można wyznaczyć postać funkcji korelacyjnej. Tą samą metodę stosuje się do obliczenia funkcji korelacyjnej procesu Wienera i procesu Levy'ego. Dlatego w tych przypadkach czytelnik sam powtórzy wszystkie kroki obliczeń. Rozpatrzmy dwa przypadki pamiętając że $N(t_0) = N(0) = 0$:

1. Niech $t_2 > t_1 > t_0 = 0$. Przyrosty $N(t_2) - N(t_1)$ oraz $N(t_1) - N(t_0)$ są zmiennymi losowymi niezależnymi dla których

$$\begin{aligned} \langle [N(t_2) - N(t_1)][N(t_1) - N(t_0)] \rangle &= \\ = \langle N(t_2) - N(t_1) \rangle \cdot \langle N(t_1) - N(t_0) \rangle &= \mu(t_2 - t_1) \mu t_1. \end{aligned} \quad (3.62)$$

Skorzystaliśmy tu z tego, że wartość średnia iloczynu zmiennych losowych niezależnych jest iloczynem wartości średnich zmiennych losowych niezależnych. Z drugiej strony, wymnożymy wyrażenia w nawiasach pamiętając, że $N(t_0) = N(0) = 0$ (proces Poissona startuje z zera). Wówczas otrzymamy

$$\langle N(t_2)N(t_1) - N^2(t_1) \rangle = \langle N(t_2)N(t_1) \rangle - \langle N^2(t_1) \rangle = \mu(t_2 - t_1) \mu t_1 \quad (3.63)$$

Stąd wynika, że

$$\begin{aligned} \langle N(t_2)N(t_1) \rangle &= \mu(t_2 - t_1) \mu t_1 + \langle N^2(t_1) \rangle = \\ &= \mu(t_2 - t_1) \mu t_1 + \mu^2 t_1^2 + \mu t_1 \end{aligned} \quad (3.64)$$

$$\mu^2 t_2 t_1 + \mu t_1 \quad \text{dla } t_2 > t_1 \quad (3.65)$$

2. Niech $t_1 > t_2 > t_0 = 0$. Przyrosty $N(t_1) - N(t_2)$ oraz $N(t_2) - N(t_0)$ są zmiennymi losowymi niezależnymi. Możemy powtórzyć trzy kroki analogiczne do tych w powyższych trzech równaniach otrzymując

$$\langle N(t_1)N(t_2) \rangle = \mu(t_1 - t_2) \mu t_2 + \langle N^2(t_2) \rangle = \mu^2 t_1 t_2 + \mu t_2, \quad \text{dla } t_1 > t_2 \quad (3.66)$$

Ponieważ

$$\langle N(t_2)N(t_1) \rangle = \langle N(t_1)N(t_2) \rangle \quad (3.67)$$

to z powyższych równań otrzymujemy tezę.

Rozkład prawdopodobieństwa Poissona

$$p_k(t) = \text{Pr}\{N(t) = k\} = e^{-\mu t} \frac{(\mu t)^k}{k!} \quad (3.68)$$

spełnia następujący układ równań ewolucji (ang. master equations)

$$\frac{dp_0(t)}{dt} = -\mu p_0(t), \quad p_0(0) = 1 \quad (3.69)$$

$$\frac{dp_k(t)}{dt} = \mu p_{k-1}(t) - \mu p_k(t), \quad p_k(0) = 0, \quad k = 1, 2, 3, \dots \quad (3.70)$$

Słuszność tego układu równań można sprawdzić bezpośrednim rachunkiem, z jednej strony różniczkując wyrażenie dla $p_k(t)$, z drugiej strony wstawiając wyrażenia na $p_k(t)$ oraz $p_{k-1}(t)$. Równania te mają też jasną interpretację. W tym celu spojrzymy nieco inaczej na nasz proces. Mówimy, że układ jest w stanie k jeżeli

$$N(t) = k \quad (3.71)$$

Stan układu jest określony przez możliwe wartości populacji w danej chwili czasu, $k = 0, 1, 2, \dots$. Zmiana stanu k układu w danej chwili czasu, opisywana przez tempo zmian czyli pochodną $dp_k(t)/dt$, może zachodzić z dwóch powodów: albo stan k pojawia się ze stanu $(k-1)$ ponieważ następuje akt narodzin, albo stan k znika i układ przechodzi do stanu $k+1$ (ponieważ następuje akt narodzin). W wyrażeniu powyższym wyraz $\mu p_{k-1}(t)$ wchodzi ze znakiem dodatnim ponieważ stan k pojawia się ze stanu $k-1$; wyraz $\mu p_k(t)$ wchodzi ze znakiem ujemnym ponieważ stan k znika i układ przechodzi w stan $k+1$ (ponieważ ktoś się urodził). Jest to nic innego jak równanie bilansu.

Proces Poissona można uogólnić tak, aby skok nie był tylko w górę o 1. Można modelować skoki w górę i w dół o dowolne wielkości. Mogą to być skoki deterministyczne lub losowe. Oto jedno z możliwych uogólnień:

$$N(t) = \sum_i z_i \theta(t - t_i), \quad (3.72)$$

gdzie $\theta(x)$ jest funkcją schodkową Heaviside'a oraz $\{t_i\}$ jest zbiorem losowych chwil skoków o średniej gęstości μ . Amplitudy skoków $\{z_i\}$ są niezależnymi zmiennymi losowymi o tym samym rozkładzie prawdopodobieństwa $\rho(z)$ i są niezależne od t_i . Realizacjami takiego procesu są funkcje schodkowe o skokach w losowych chwilach czasu t_i i o losowych wielkościach skoku z_i . Wartość średnia takiego procesu Poissona wynosi

$$\langle N(t) \rangle = \mu \langle z_i \rangle t \quad (3.73)$$

gdzie

$$\langle z_i \rangle = \int_{-\infty}^{\infty} z \rho(z) dz \quad (3.74)$$

Funkcja korelacyjna tego procesu Poissona ma postać:

$$\begin{aligned} R(t_2, t_1) &= \langle N(t_2) N(t_1) \rangle = \mu^2 \langle z_i \rangle^2 t_2 t_1 + \mu \langle z_i \rangle \min(t_2, t_1) = \\ &= \langle N(t_2) \rangle \langle N(t_1) \rangle + \mu \langle z_i \rangle \min(t_2, t_1) \end{aligned} \quad (3.75)$$

Pojedynczą Realizację procesu Poissona można uzyskać poprzez wygenerowanie N niezależnych punktów losowo rozłożonych na osi czasu na odcinku $[0, T]$. Ilość punktów generujemy z rozkładu Poissona a ich położenie na osi czasu zgodnie z rozkładem jednorodnym $U(0, N)$. Najprościej będzie posłużyć się pakietem `scipy`, aby wygenerować ilość punktów korzystając z rozkładu Poissona. Punkty czasowe, w których nastąpi skok uzyskamy wykorzystując podstawową funkcję `random()`.

```

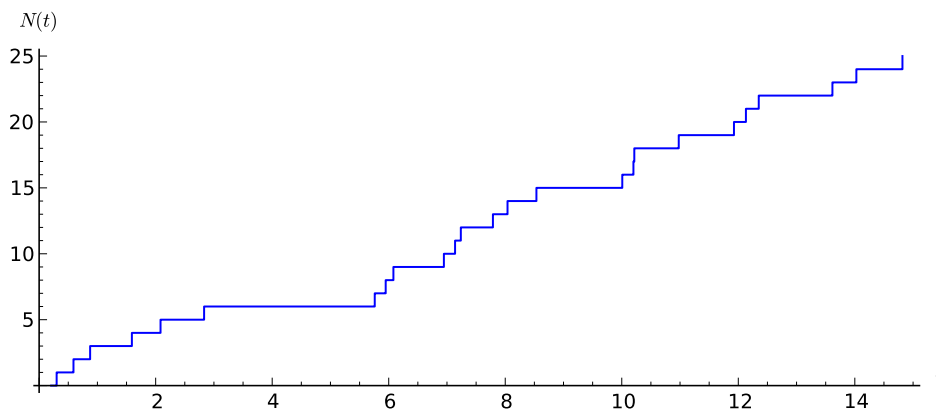
from scipy import stats
import matplotlib.pyplot as plt

T = 15
mu = 1.3
N = stats.poisson.rvs(T*mu)
steps = range(N+1)

points = sorted([random()*T for i in steps])

plot_step_function(zip(points, steps))

```



Rysunek 3.4: Relizacja procesu Poissona.

3.2.3 Proces Levy-ego

Uogólnimy dwa powyższe procesy: Wienera i Poissona, które w potocznym sensie są jednocześnie najbardziej losowe. Są też procesami Markowa. Podkreślamy, że są to podstawowe procesy stochastyczne, ale nie powinno się je nazywać szumem ponieważ nie są to procesy stacjonarne. Są one najbardziej losowe, ponieważ następujące po sobie przyrosty tych procesów są niezależne.

Powyższe dwa procesy są szczególnymi przypadkami procesu który nazywa się procesem Levy'ego. Wiele innych procesów są albo funkcjami procesu Levy'ego albo funkcjami (trochę ogólniejszymi zależnościami) tego procesu.

Definicja procesu Levy'ego $L(t)$ jest relatywnie prosta: jest to rzeczywisty proces stochastyczny, prawostronnie ciągły o skończonych granicach lewostronnych oraz

1. $L(t)$ startuje z zera, tzn. $L(0) = 0$
2. $L(t)$ ma niezależne przyrosty na nieprzekrywających się przedziałach, tzn. przyrosty $L(t_4) - L(t_3)$ oraz $L(t_2) - L(t_1)$ są niezależnymi zmiennymi losowymi na nieprzekrywających się przedziałach $[t_1, t_2]$ and $[t_3, t_4]$ dla dowolnych chwil czasu takich że $0 \leq t_1 \leq t_2 \leq t_3 \leq t_4$
3. $L(t)$ ma stacjonarne przyrosty, tzn. rozkład prawdopodobieństwa zmiennych losowych $L(t_2) - L(t_1)$ zależy od różnicy czasów $t_2 - t_1$ dla dowolnych $0 \leq t_1 \leq t_2$.

4. $L(t)$ jest stochastycznie ciągły, tzn. dla dowolnych $t \geq 0$ oraz $\epsilon > 0$:

$$\lim_{s \rightarrow t} P(|L(t) - L(s)| > \epsilon) = 0 \quad (3.76)$$

Własności te są takie same jak dla procesu Wienera i procesu Poissona. Ale istnieją jeszcze inne procesy, które mają te same własności. Dlatego włączenie tych innych procesów do procesu Wienera i Poissona daje w rezultacie proces Levy'ego.

Z własności (1) i (2) wynika, że funkcja korelacyjna procesu Levy'ego ma postać (patrz wykłady: Procesy i zjawiska losowe)

$$\langle L(t)L(s) \rangle = 2D_0 \min(t, s) \equiv 2D_0 [t\theta(s-t) + s\theta(t-s)], \quad (3.77)$$

gdzie $D_0 > 0$ jest intensywnością (natężeniem) procesu Levy'ego.

Własności probabilistyczne tego procesu można wyznaczyć np. z jego funkcji charakterystycznej, która zgodnie z twierdzeniem Levy'ego-Chinczyna ma postać

$$C(\omega, t) = \langle e^{i\omega L(t)} \rangle = e^{t\psi(\omega)}, \quad (3.78)$$

gdzie eksponenta Levy'ego ma postać

$$\psi(\omega) = ia_0\omega - \frac{1}{2}b\omega^2 + \int_{-\infty}^{\infty} [e^{i\omega y} - 1 - i\omega y I_{(-1,1)}(y)] \nu(dy), \quad (3.79)$$

Stałe $a_0 \in R, b \geq 0$. Oznaczenie $I_A(y)$ jest tzw. indykatorem zbioru A - jest to funkcja o własności:

$$I_A(y) = 1 \quad \text{gdy} \quad y \in A \quad \text{oraz} \quad I_A(y) = 0 \quad \text{w pozostałych przypadkach} \quad (3.80)$$

$\nu(dy)$ jest tzw. miarą Levy'ego o własności

$$\nu(R - [-1, 1]) < \infty, \quad \int_{-1}^1 y^2 \nu(dy) < \infty. \quad (3.81)$$

Dla niewtajemniczonych i nie-ekspertów matematycznych: zamiast tajemniczego zapisu $\nu(dy)$ można używać zapisu $\nu(dy) = h(y)dy$ gdzie $h(y)$ jest nieujemną funkcją. Z powyższego przedstawienia funkcji charakterystycznej wynika, że proces Levy'ego jest określony przez 3 liczby, tzw. tryplet Levy'ego-Chinczyna (a_0, b, ν) , gdzie a_0 charakteryzuje dryf, b to parametr procesu Wienera oraz ν charakteryzuje skoki procesu. Tryplet $(0, b, 0)$ to rozkład Gaussa dla procesu Wienera. Tryplet $(0, 0, \mu\delta(y-1))$ to rozkład Poissona z parametrem μ który opisuje proces Poissona o skokach 1 w chwilach o równomiernym rozkładzie. Tzw. złożony proces Poissona to proces o losowych skokach o różnej wielkości z rozkładem prawdopodobieństwa ν z $\nu(R) < \infty$. Eksponenta Levy'ego przyjmuje wówczas postać

$$\psi(\omega) = \mu \int_{-\infty}^{\infty} [e^{i\omega y} - 1] \nu(dy). \quad (3.82)$$

Jeżeli $\nu(R) = \infty$ wówczas $L(t)$ jest skokowym procesem o nieskończonej liczbie małych skoków w dowolnie małym przedziale czasu. W rzeczywistości taki proces nie istnieje, ale

jeżeli pojawia się bardzo duża ilość skoków w małych przedziałach czasu, to przybliżenie takie dobrze modeluje układ.

Z twierdzenia Levy'ego-Ito wynika, że proces Levy'ego $L(t)$ składa się z 4 niezależnych procesów elementarnych:

$$L(t) = L_1(t) + L_2(t) + L_3(t) + L_4(t), \quad (3.83)$$

gdzie $L_1(t)$ to dryf, $L_2(t)$ to proces Wienera, $L_3(t)$ to złożony proces Poissona oraz $L_4(t)$ to czysto skokowy proces zwany martingalem (a pure jump martingale). Mówiąc w prosty sposób, martingał to taki proces stochastyczny, że warunkowa wartość średnia

$$\langle L(t_{n+1}) | L(t_1)L(t_2)...L(t_n) \rangle = L(t_n) \quad (3.84)$$

Innymi słowy, to taki proces w którym warunkowa wartość średnia procesu w momencie $t + 1$, gdy znamy jego wartości do jakiegoś wcześniejszego momentu t , jest równa wartości procesu w momencie t .

Taka dekompozycja wynika z eksponenty Levy'ego, która może być przedstawiona w postaci 4 wyrażań:

$$\psi(\omega) = \psi_1(\omega) + \psi_2(\omega) + \psi_3(\omega) + \psi_4(\omega), \quad (3.85)$$

gdzie

$$\psi_1(\omega) = ia_0\omega, \quad \psi_2(\omega) = -\frac{1}{2}b\omega^2, \quad \psi_3(\omega) = \int_{|y| \geq 1} [e^{i\omega y} - 1] \nu(dy), \quad (3.86)$$

$$\psi_4(\omega) = \int_{|y| < 1} [e^{i\omega y} - 1 - i\omega y] \nu(dy). \quad (3.87)$$

Warto podkreślić, że liniowa kombinacja niezależnych procesów Levy'ego jest także procesem Levy'ego.

Szczególny, ale bardzo ważnym przypadkiem procesu Levy'ego jest tzw. α -stabilny proces $L_\alpha(t)$ o indeksie $\alpha \in (0, 2]$. To jest przypadek trypletu $(a, 0, \nu)$ z miarą Levy'ego w postaci

$$\nu(y) = [c_1 I_{(0, \infty)}(y) + c_2 I_{(-\infty, 0)}(y)] |y|^{-\alpha-1} dy, \quad (3.88)$$

gdzie $c_1 > 0$ i $c_2 > 0$. Funkcja charakterystyczna jest postaci

$$\psi(\omega) = \begin{cases} ia\omega - c|\omega|^\alpha \left(1 - i\beta \operatorname{sgn} \omega \tan \frac{\pi\alpha}{2}\right), & \text{for } \alpha \neq 1, \\ ia\omega - c|\omega| \left(1 + i\beta \frac{2}{\pi} \operatorname{sgn} \omega \ln |k|\right), & \text{for } \alpha = 1, \end{cases} \quad (3.89)$$

gdzie

$$\alpha \in (0, 2], \quad \beta = \beta(c_1, c_2) \in [-1, 1], \quad c = c(\alpha, c_1, c_2) \in (0, \infty) \quad (3.90)$$

oraz $a = a(a_0, \alpha, c_1, c_2)$. Przypadek $c_1 = c_2$ implikuje $\beta = 0$ i wówczas proces jest symetryczny.

Tzw. charakterystyczny funkcjonal (pewien szczególny rodzaj odwzorowania, funkcji) symetrycznego α -stabilnego białego szumu Levy'ego $Y_\alpha(t)$ (gdy $a = 0, \beta = 0$ w równaniu (3.89)) jest w postaci

$$\mathcal{C}_{Y_\alpha}[f] = \langle \exp \left[i \int_0^t ds f(s) Y_\alpha(s) \right] \rangle = \exp \left[-c \int_0^t dt |(s)|^\alpha \right], \quad (3.91)$$

Ten zapis ma uzmysłwić, że funkcjonal \mathcal{C} zależy od historii na przedziale $(0, t)$. Funkcja f jest dowolną funkcją. Jeżeli testowa funkcja $f(t)$ jest stała, $f(s) = \omega$, to funkcjonal redukuje się do funkcji charakterystycznej.

Generowanie realizacji procesu Levy'ego wymaga nieco większego nakładu pracy. Metoda generowania zmiennych z rozkładu α -stabilnego opisana jest w dodatku numerycznym na końcu tego podręcznika. Sama realizacja procesu przebiega identycznie jak w przypadku procesu Wienera. Jedynymi różnicami są wykładnik stojący przy prefaktorze zmiennej losowej oraz rozkład którym posługujemy się do generowania liczb losowych.

Posłużymy się metodą rvs funkcji `levy_stable` wbudowaną w pakiet `stats` biblioteki `scipy`. Pobiera ona 5 argumentów

```
stats.levy_stable(alpha, beta, loc=0, scale=1, size=1)
```

gdzie

- `alpha, beta` - to parametry rozkładu Levy'ego
- `loc=0, scale=1` - to odpowiednio średnia i odchylenie standardowe (wraz z domyślnymi wartościami)
- `size=1` - to ilość zmiennych losowych generowanych jednorazowym wywołaniem

Więcej informacji uzyskamy w pomocy (wpisując `stats.levy_stable?`)

```
# kilka importów
import numpy as np
from scipy import stats
import warnings
warnings.simplefilter('ignore', DeprecationWarning)

# parametry symulacji
h = 0.01 #krok
N = 5000 #ilosc krokow

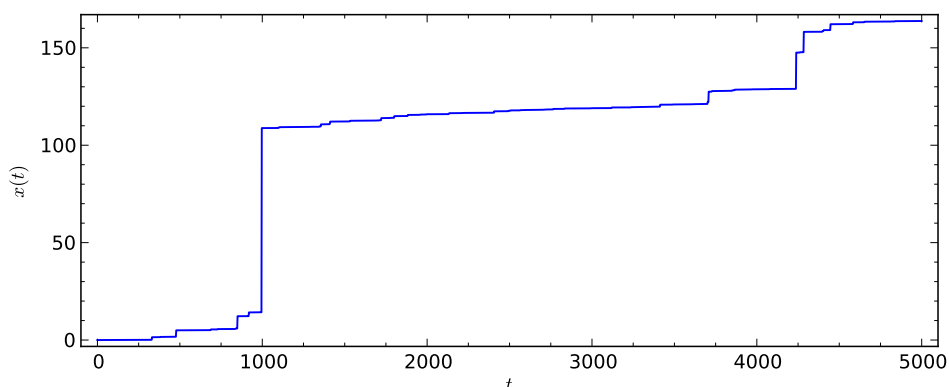
# parametry SDE
x0 = 0 #wartosc poczatkowa

#Levy - Smirnof
alpha = 1/2
beta = 1
mu = 0
sigma = 0.1

x = [x0]
for i in xrange(1, N):
```

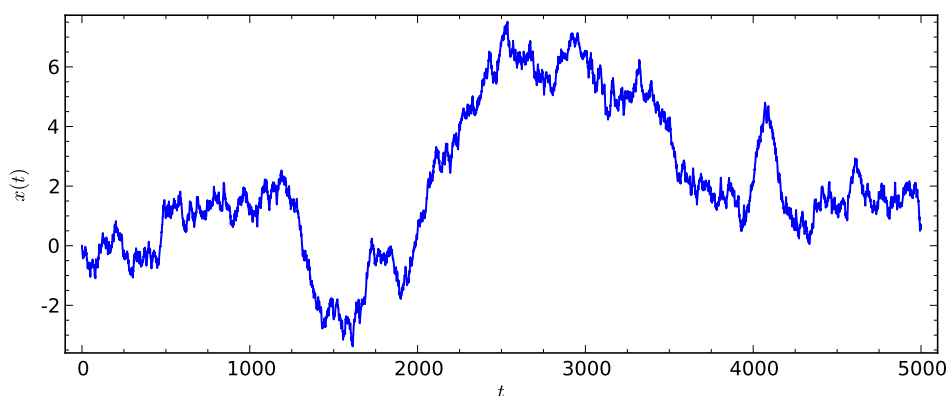


```
lab = stats.levy_stable.rvs(alpha, beta, loc=mu, scale=sigma)
x.append(x[i-1] + h**(1/alpha) * lab)
```



Rysunek 3.5: Relizacja procesu Levy-Smirnoffa.

Korzystając z rozkładu Levy'ego możemy oczywiście uzyskać zmienne normalne $N(0, 1)$ kładąc $\alpha = 2, \beta = 0, \sigma = 1/\sqrt{2}, \mu = 0$.



Rysunek 3.6: Relizacja procesu Wienera uzyskana z rozkładu Levy'ego.

3.3 Rachunki w całkach Ito - Stratonowicza

Przedstawimy użyteczne formuły, które pojawiają się w teorii Ito i Stratonowicza. Ograniczymy się do przypadku 1-wymiarowego, tzn. jednego równania stochastycznego w interpretacji Ito:

$$dX = F(X)dt + G(X)dW(t) \quad (3.92)$$

gdzie proces Wienera ma charakterystyki

$$\langle dW(t) \rangle = 0, \quad \langle [dW(t)]^2 \rangle = 2Ddt \quad (3.93)$$

Stąd, w sensie średniokwadratowym

$$dW(t) = \sqrt{2Ddt} \quad (3.94)$$

Powyższe relacje są bardzo ważne dla zrozumienia rachunku Ito, który różni się od znanego w klasycznej analizie matematycznej.

3.3.1 Rachunek różniczkowy Ito

Rozpatrzmy teraz funkcję dwóch zmiennych $g(x, t)$, która jest różniczkowalna dostateczną ilość razy. W szczególności zakładamy, że istnieją pochodne

$$g'(x, t) = g' = \frac{\partial g(x, t)}{\partial x}, \quad \dot{g}(x, t) = \dot{g} = \frac{\partial g(x, t)}{\partial t}, \quad g''(x, t) = g'' = \frac{\partial^2 g}{\partial x^2}. \quad (3.95)$$

Przyjmijmy taką konwencję, że różniczkowanie względem pierwszego argumentu (względem 'położenia' x) oznaczmy apostrofem ' ; różniczkowanie względem drugiego argumentu (względem czasu t) oznaczmy kropką .

Rozpatrzmy teraz funkcję $g(X, t)$, gdzie teraz pierwszym argumentem jest proces stochastyczny $X(t)$ określony przez powyższe równanie stochastyczne Ito. Obliczmy przyrost funkcji $g(X, t)$ rozwijając funkcję w szereg Taylora:

$$\begin{aligned} \Delta g(X, t) &= g(X + dX, t + dt) - g(x, t) = \\ &= \frac{\partial g}{\partial X} dX + \frac{\partial g}{\partial t} dt + \frac{1}{2} \frac{\partial^2 g}{\partial X^2} dX dX + \frac{\partial^2 g}{\partial X \partial t} dX dt + \frac{1}{2} \frac{\partial^2 g}{\partial t^2} dt dt + \dots \end{aligned} \quad (3.96)$$

Wstawimy teraz wyrażenie na dX z równania Ito

$$\begin{aligned} \Delta g(X, t) &= g'(X, t) [F(X, t)dt + G(X, t)dW] + \dot{g}(X, t)dt + \\ &+ \frac{1}{2} g''(X, t) [F(X, t)dt + G(X, t)dW] [F(X, t)dt + G(X, t)dW] + \\ &+ \dot{g}'(X, t)dt [F(X, t)dt + G(X, t)dW] + \frac{1}{2} \ddot{g}(X, t)dt dt + \dots \end{aligned} \quad (3.97)$$

W wyrażeniu tym pozostawimy wyrazy rzędu co najwyżej dt pamiętając że $dW(t)$ jest rzędu \sqrt{t} :

$$dg(X, t) = g'(X, t) [F(X, t)dt + G(X, t)dW] + \dot{g}(X, t)dt + \frac{1}{2} g''(X, t) G^2(X, t) dW dW \quad (3.98)$$

W wyrażeniu tym zamieniliśmy $\Delta g(X, t) = dg(X, t)$ ponieważ przyrost liniowy w dt to różniczka. Pozostałe wyrazy dają przyczynki wyższego rzędu $dt dt$, $dt dW$ i pomijamy je. Zastąpimy teraz $dW dW = dW^2(t)$ przez $2Ddt$ otrzymując słynną formułę Ito

$$dg(X, t) = [\dot{g}(X, t) + g'(X, t)F(X, t) + Dg''(X, t)G^2(X, t)]dt + g'(X, t)G(X, t)dW \quad (3.99)$$

Przykład 1

Zobaczmy, jakie nietypowe wnioski można wyciągnąć z tej formuły. W tym celu rozważmy szczególny przypadek i przyjmijmy następujące wyrażenia:

$$dX = dW, \quad \text{tzn.} \quad F(X, t) = 0, \quad G(X, t) = 1, \quad D = \frac{1}{2} \quad (3.100)$$

Niech

$$g(X, t) = X^2, \quad X(0) = 0, \quad D = \frac{1}{2} \quad (3.101)$$

Wówczas z formuły Ito otrzymamy:

$$d(X^2) = 2XdX + dt \quad (3.102)$$

Stąd wynika, że

$$XdX = \frac{1}{2}d(X^2) + \frac{1}{2}dt \quad (3.103)$$

Ponieważ w tym przykładzie $dX = dW$, czyli $X = W$, to możemy równie dobrze napisać

$$WdW = \frac{1}{2}d(W^2) + \frac{1}{2}dt \quad (3.104)$$

Obustronnie całkowanie daje taki oto wynik

$$\int_0^t WdW = \frac{1}{2} \int_0^t d(W^2) + \frac{1}{2} \int_0^t dt = \frac{1}{2}W^2 + \frac{1}{2}t \quad (3.105)$$

Widać, że w porównaniu z tradycyjnym rachunkiem różniczkowym i całkowym, tutaj pojawia się dodatkowy składnik $(1/2)t$. Przykład ten pokazuje, że metody rachunkowe, których nauczyliśmy się na kursie analizy matematycznej, w teorii procesów stochastycznych nie muszą obowiązywać. Różniczkowanie i całkowanie wielkości, w których bezpośrednio lub pośrednio pojawiają się procesy Wienera, Poissona, czy ogólniej Levy'ego należy wykonywać biorąc pod uwagę równania stochastyczne typu Ito. Tutaj reguły są nierozzerwalnie związane z równaniami stochastycznymi definiującymi proces stochastyczny. Czytelnik powinien zauważyć, że wszelkie odstępstwa od tradycyjnego rachunku różniczkowego i całkowego pojawiają się dlatego, że podstawowe procesy takie jak procesy Wienera, Poissona, czy ogólniej Levy'ego są to procesy o przyrostach niezależnych. To z kolei pociąga za sobą własność narastania w czasie fluktuacji tych procesów. Fluktuacje te rosną jak \sqrt{t} . Oto przyczyna wszelkich odstępstw. A pomyśleć, że wszystko to wzięło swój początek z rzucania monetą i rozmów telefonicznych.

Przykład 2

Rozpatrzmy równanie Ito dla geometrycznego ruchu Browna, rozważanego w jednym z poprzednich rozdziałów:

$$dX = cXdt + XdW(t), \quad F(X, t) = cX, \quad G(X, t) = X \quad (3.106)$$

Niech

$$g(X, t) = \ln X(t) = Y(t) \quad (3.107)$$

Teraz $Y(t)$ to nowy proces stochastyczny będący transformacją procesu $X(t)$. Jakie równanie spełnia proces $Y(t)$? Obliczmy różniczkę procesu $Y(t)$ korzystając z formuły Ito. Obliczamy:

$$\dot{g}(x, t) = 0, \quad g'(x, t) = \frac{1}{x}, \quad g''(x, t) = -\frac{1}{x^2}, \quad (3.108)$$

i wstawiamy do formuły Ito pamiętając że $dg(X, t) = dY(t)$. Otrzymamy końcowy wynik w postaci

$$dY(t) = (c - D)dt + dW(t) \quad (3.109)$$

Równanie to jest proste i można je scałkować:

$$Y(t) = Y(0) + (c - D)t + W(t) \quad (3.110)$$

Od razu widać, że jest to proces gaussowski i łatwo wyznaczyć jego rozkład prawdopodobieństwa obliczając jego wartość średnią i wariancję.

3.3.2 Równanie Stratonowicza a równanie Ito

Rozważmy równanie stochastyczne w sensie Ito:

$$dX(t) = F(X(t), t)dt + G(X(t), t)dW(t) \quad (3.111)$$

oraz równanie stochastyczne w sensie Stratonowicza

$$dX(t) = F(X(t), t)dt + G(X(t), t) \circ dW(t) \quad (3.112)$$

Pytamy, jaki jest związek między tymi równaniami. Rozpatrzmy ostatni wyraz w równaniu Stratonowicza

$$\begin{aligned} G(X(t), t) \circ dW(t) &= G\left(\frac{[X(t+dt) + X(t)]}{2}, t\right) dW(t) = \\ &= G\left(X(t) + \frac{1}{2}[X(t+dt) - X(t)], t\right) dW(t) = \\ &= G\left(X(t) + \frac{1}{2}dX(t), t\right) dW(t) = \\ &= \left[G(X(t), t) + \frac{1}{2}G'(X(t), t)dX(t) + \dots\right] dW(t) = \\ &= G(X(t), t)dW(t) + \frac{1}{2}G'(X(t), t)dX(t)dW(t) + \dots = \\ &= G(X(t), t)dW(t) + \frac{1}{2}G'(X(t), t)[F(X(t), t)dt + G(X(t), t)dW(t)]dW(t) + \dots = \\ &= G(X(t), t)dW(t) + \frac{1}{2}G'(X(t), t)G(X(t), t)dW(t)dW(t) + \dots = \\ &= G(X(t), t)dW(t) + \frac{1}{2}G'(X(t), t)G(X(t), t)2Ddt + \dots, \end{aligned} \quad (3.113)$$

gdzie rozwijaliśmy w szereg Taylora funkcję $G(X + h, t)$ dla $h = dX(t)/2$ oraz zatrzymaliśmy wyrazy co najwyżej rzędu dt , Skorzystaliśmy z tego, że w sensie średniokwadratowym $dW(t)dW(t) = 2Ddt$. Jeżeli teraz wstawimy to wyrażenie do równania to otrzymamy równanie w sensie Ito w postaci

$$\begin{aligned} \text{Stratonowicz} \rightarrow dX(t) &= F(X(t), t)dt + G(X(t), t) \circ dW(t) = \\ &= [F(X(t), t) + DG'(X(t), t)G(X(t), t)]dt + G(X(t), t)dW(t) \leftarrow \text{Ito} \end{aligned} \quad (3.114)$$

Równanie to ustala związek między równaniem Stratonowicza a równaniem w sensie Ito. Przy transformacji równania Stratonowicza do równania Ito pojawia się wyraz

$$\frac{1}{2}G'(X(t), t)G(X(t), t)\langle [dW(t)]^2 \rangle = DG'(X(t), t)G(X(t), t)dt \quad (3.115)$$

nazywany dryfem indukowanym przez szum, ponieważ jest to dodatkowy wyraz przy różniczkę dt .

Z powyższej formuły wynika także odwrotny związek:

$$\begin{aligned} \text{Ito} \rightarrow dX(t) &= F(X(t), t)dt + G(X(t), t)dW(t) = \\ &= [F(X(t), t) - DG'(X(t), t)G(X(t), t)]dt + G(X(t), t) \circ dW(t) \leftarrow \text{Stratonowicz} \end{aligned} \quad (3.116)$$

W wielu przypadkach warto “pracować” z równaniami w sensie Stratonowicza, ponieważ operacje na tych równaniach są takie same jak w standardowej (klasycznej) analizie matematycznej. Zademonstrujemy to przykładzie równania opisującego geometryczny ruch Browna:

$$dX = cXdt + XdW(t), \quad F(X, t) = cX, \quad G(X, t) = X \quad (3.117)$$

Jest to równanie w interpretacji Ito. W interpretacji Stratonowicza ma ono postać:

$$dX = (c - D)Xdt + X \circ dW(t) \quad (3.118)$$

Równanie to można traktować w tradycyjny sposób, Np. podzielić obustronnie równanie przez X (czego nie wolno robić w równaniu Ito):

$$\frac{dX}{X} = (c - D)dt + dW(t) \quad (3.119)$$

W tradycyjnym rachunku

$$\frac{dX}{X} = \ln X \equiv Y \quad (3.120)$$

Więc proces $Y = Y(t)$ spełnia równanie

$$dY = (c - D)dt + dW(t) \quad (3.121)$$

Jest to równanie jakie otrzymaliśmy powyżej z równania Ito przy stosowaniu rachunków Ito.

3.4 Modele szumów

3.4.1 Biały szum Levy’ego

W literaturze szumem nazywa się dowolny proces stacjonarny. Stacjonarny - oznacza to, że funkcja korelacyjna procesu $\xi(t)$ zależy tylko od różnicy czasów, tzn.

$$C(t, s) = \langle \xi(t)\xi(s) \rangle = C(t - s) \quad (3.122)$$

Innymi słowy funkcja

$$C(t) = \langle \xi(s+t)\xi(s) \rangle \quad (3.123)$$

nie zależy od czasu s , a zależy tylko od czasu t .

Podaliśmy definicję procesu Levy'ego jako najbardziej ogólnego procesu stochastycznego o niezależnych przyrostach na nieprzekrywających się przedziałach. Proces ten w szczególnych przypadkach redukuje się do procesu Wienera i procesu Poissona. Pochodna tego procesu to biały szum Levy'ego $\xi(t)$:

$$\xi(t) = \frac{dL(t)}{dt} \quad (3.124)$$

Z postaci funkcji korelacyjnej dla procesu Levy'ego wynika, że funkcja korelacyjna białego szumu Levy'ego $\xi(t)$ ma postać

$$\begin{aligned} \langle \xi(t)\xi(s) \rangle &= \frac{\partial^2}{\partial t \partial s} \langle L(t)L(s) \rangle = \frac{\partial^2}{\partial t \partial s} [2D \min(t, s)] = \\ &= 2D \frac{\partial^2}{\partial t \partial s} [t\theta(s-t) + s\theta(t-s)] = 2D\delta(t-s), \end{aligned} \quad (3.125)$$

co uzasadnia nazwę "biały szum", ponieważ widmo mocy (transformata Fouriera funkcji korelacyjnej) jest niezależna od częstości, podobnie jak dla światła białego.

Przy obliczaniu pochodnych należy pamiętać, że $\theta'(t) = \delta(t)$ oraz $t\delta(t-s) = s\delta(t-s)$ i $\delta(t-s) = \delta(s-t)$.

Szczególne przypadki tego szumu to:

1. biały szum gaussowski (pochodna procesu Wienera)
2. biały szum poissonowski, nazywany białym szumem śrutowym lub poissonowskim ciągiem δ -impulsów (pochodna procesu Poissona)

3.4.2 Szumy gaussowskie

Szumem gaussowskim nazywamy dowolny stacjonarny proces stochastyczny o gaussowskim rozkładzie prawdopodobieństwa. Jest on najczęściej wykorzystywany w modelowaniu z powodu centralnego twierdzenia granicznego. Szum gaussowski pojawia się przy opisie układu oddziałującego z otoczeniem (termostatem) ponieważ termostat jest układem o wielkiej liczbie stopni swobody (większej od liczby Avogadra 10^{23}). Dlatego szum termiczny i równowagowe fluktuacje termiczne modelowane są szumem gaussowskim. Szum taki jest całkowicie określony przed dwie wielkości: wartość średnią oraz funkcję korelacyjną. Z matematycznego punktu widzenia, proces taki możemy opisać tzw. funkcjonalem rozkładu prawdopodobieństwa, który jest uogólnieniem gęstości rozkładu prawdopodobieństwa :

$$\mathcal{D}P[\xi] = \mathcal{D}z \exp \left[-\frac{1}{2} \int dt \int ds \xi(t) K(t-s) \xi(s) \right], \quad (3.126)$$

gdzie $\mathcal{D}z$ jest pewną miarą funkcjonalną (podobnie jak iloczyn różniczek i jacobianu transformacji w całkach wielokrotnych) oraz funkcja $K(t)$ jest odwrotnością funkcji korelacyjnej $C(t)$

w takim sensie, że zachodzi relacja

$$\int ds K(t-s)C(s-u) = \delta(t-u) \quad (3.127)$$

Biały szum gaussowski odpowiada przypadkowi, gdy funkcja korelacyjna ma postać:

$$C(t) = 2D_0\delta(t), \quad (3.128)$$

gdzie D_0 jest intensywnością szumu. Dla białego szumu gaussowskiego $\xi(t)$ funkcjonal rozkładu prawdopodobieństwa ma postać:

$$\mathcal{D}P[\xi] = \mathcal{D}z \exp \left[-D_0 \int dt \xi^2(t) \right], \quad (3.129)$$

Dlaczego wprowadza się taki dziwny obiekt matematyczny jak funkcjonal $\mathcal{D}P[\xi]$. Przypomnijmy sobie, że zmienna losowa o rozkładzie Gaussa jest określona przez wartość średnią zmiennej losowej i jej wariancję

$$\sigma^2 = \langle \xi^2 \rangle - \langle \xi \rangle^2 \quad (3.130)$$

Dla białego szumu $\langle \xi^2(t) \rangle = \infty$ i dlatego nie istnieje 1-wymiarowy rozkład prawdopodobieństwa $f(x, t)$ ponieważ $\sigma^2 = \infty$. Dlatego należy uogólnić teorię tak, aby opisywała także biały szum gaussowski. To umożliwia funkcjonal $\mathcal{D}P[\xi]$.

Istnieją także innego typu szumy gaussowskie. Są to szumy skorelowane (kolorowe). Ważnym przykładem jest eksponencjalnie skorelowany szum:

$$C(t) = \frac{D_0}{\tau_c} \exp \left(-\frac{|t|}{\tau_c} \right), \quad (3.131)$$

gdzie D_0 jest intensywnością (natężeniem) szumu oraz τ_c nazywa się czasem korelacji szumu. Ten szum jest generowany przez proces Ornsteina-Uhlenbecka, który opiszemy w następnym rozdziale. Kolejnym przykładem szumu skorelowanego jest tzw. szum harmoniczny o funkcji korelacyjnej typu:

$$C(t) = a_1 e^{-a|t|} (\cos \omega t + a_2 \sin \omega t) \quad (3.132)$$

gdzie a_1, a_2 oraz $a > 0$ są stałymi. Ten szum jest opisany równaniem podobnym do równania Newtona dla oscylatora harmonicznego tłumionego. Ostatnim przykładem jest szum skorelowany algebraicznie, tzn. jego funkcja korelacyjna jest postaci:

$$C(t) = C_0 \left(1 + \frac{|t|}{\tau_c} \right)^{-\kappa} \quad (3.133)$$

gdzie liczba $\kappa > 0$.

Wszystkie trzy przykłady szumu skorelowanego dążą do białego szumu gaussowskiego w pewnych granicznych przypadkach. Aby to pokazać trzeba skorzystać z twierdzenia o ciągach delta-podobnych.

3.4.3 Szumy poissonowskie

Poissonowski biały szum to pochodna procesu Poissona. Uogólniony proces Poissona można zapisać przy pomocy funkcji theta Heaviside'a w postaci

$$N(t) = \sum_i z_i \theta(t - t_i), \quad (3.134)$$

gdzie $\theta(x)$ jest funkcją schodkową Heaviside'a oraz $\{t_i\}$ jest zbiorem losowych chwil skoków o średniej gęstości μ . Amplitudy skoków $\{z_i\}$ są niezależnymi zmiennymi losowymi o tym samym rozkładzie prawdopodobieństwa $\rho(z)$ i są niezależne od t_i . Realizacjami takiego procesu są funkcje schodkowe o skokach w losowych chwilach czasu t_i i o losowych wielkościach skoku z_i . Pochodna tego procesu

$$\xi(t) = \frac{dN(t)}{dt} = \sum_i z_i \delta(t - t_i) \quad (3.135)$$

to biały szum poissonowski. Faktycznie jest to losowy ciąg impulsów typu δ -Diraca.

Pamiętajmy, że parametr μ określa średnią liczbę delta-impulsów na jednostkę czasu. Ponieważ średnia wartość procesu Poissona wynosi $\langle N(t) \rangle = \mu \langle z_i \rangle t$ więc średnia wartość białego szumu wynosi

$$\langle \xi(t) \rangle = \frac{d\langle N(t) \rangle}{dt} = \mu \langle z_i \rangle \quad (3.136)$$

gdzie

$$\langle z_i \rangle = \int_{-\infty}^{\infty} z \rho(z) dz \quad (3.137)$$

Możemy przededefiniować tak szum, aby jego wartość wynosiła 0. Zdefiniujmy nowy szum jako

$$Y_0(t) = \sum_i z_i \delta(t - t_i) - \mu \langle z_i \rangle \quad (3.138)$$

Wartość średnia oraz funkcja korelacyjne tego szumu to:

$$\langle Y_0(t) \rangle = 0, \quad \langle Y_0(t) Y_0(u) \rangle = 2D_S \delta(t - u), \quad (3.139)$$

gdzie $D_S = (1/2)\mu \langle z_i^2 \rangle$ nazywa się intensywnością szumu.

Symetryczny biały szum Poissona to taki ciąg impulsów, że gęstość prawdopodobieństwa $\rho(z)$ dla rozkładu amplitud jest parzystą (symetryczną) funkcją. Na przykład rozkład prawdopodobieństwa wielkości skoków

$$\rho(z) = e^{-|z|/A}, \quad A > 0 \quad (3.140)$$

jest funkcją symetryczną $z \rightarrow -z$. Dlatego też skoki w górę ($z > 0$) oraz w dół ($z < 0$) są równo prawdopodobne i to odpowiada procesowi urodzin i śmierci.

Jeżeli szum jest asymetryczny, to $\rho(z) \neq \rho(-z)$. Np.

$$\rho(z) = (1/A^2) z e^{(-z/A)} \theta(z), \quad A > 0 \quad (3.141)$$

opisuje niesymetryczny biały szum Poissona. Przykładem też jest rozkład

$$\rho(z) = (1/A)e^{-z/A}\theta(z), \quad A > 0 \quad (3.142)$$

W tym przypadku możliwe są tylko skoki w górę ($z > 0$) i ich średnia wartość wynosi:

$$\langle z_i \rangle = A \quad (3.143)$$

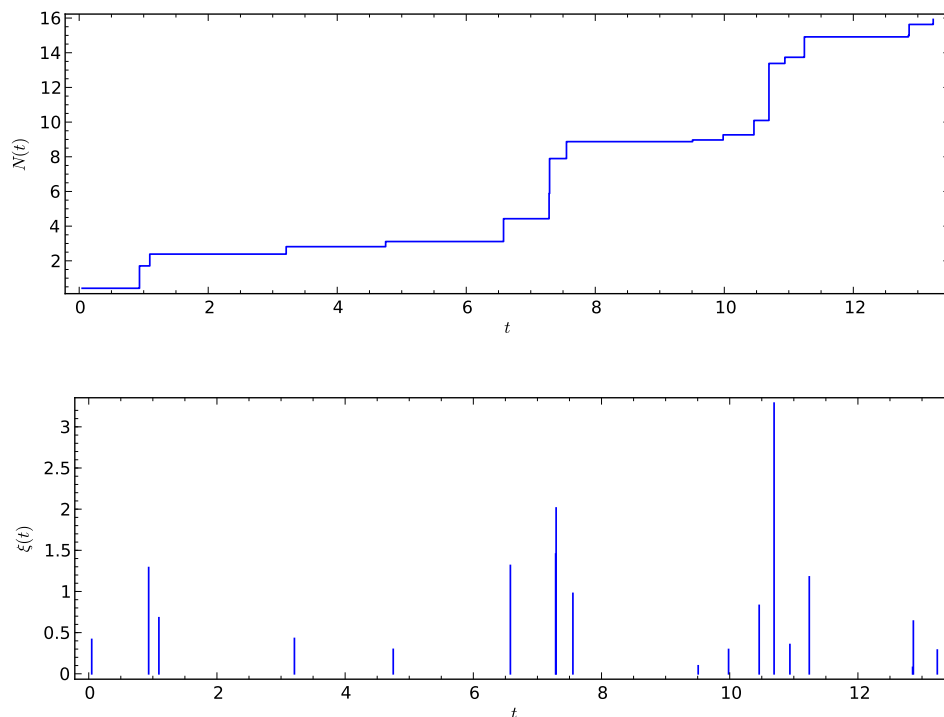
Stąd otrzymujemy interpretację parametru A w powyższym rozkładzie prawdopodobieństwa.

```
#realizacja białego szumu Poissona
reset()
from scipy import stats
from numpy import cumsum

T = 15
mu = 1.3
N = stats.poisson.rvs(T*mu)

steps = range(N+1)
pts = sorted([random()*T for i in steps])
z = stats.expon.rvs(size=N).tolist()
steps2 = cumsum(z)

plot_step_function(zip(pts, steps2))
sum([line((pts[i], 0), (pts[i], z[i])) for i in steps[:-1]])
```



Rysunek 3.7: Reliacja procesu i odpowiadającego mu szumu Poissona.

3.4.4 Szum dychotomiczny (telegraficzny)

Podamy teraz przykład niegaussowskiego szumu. Jest to 2-stanowy proces stochastyczny

$$\xi(t) = \{-a, b\}, \quad a, b > 0. \quad (3.144)$$

Szum może przyjmować dwie wartości $-a$ lub b . Przeskoki pomiędzy tymi dwoma stanami są dane przez prawdopodobieństwa przejścia w jednostce czasu

$$Pr(-a \rightarrow b) = \mu_a = 1/\tau_a, \quad Pr(b \rightarrow -a) = \mu_b = 1/\tau_b, \quad (3.145)$$

gdzie τ_a and τ_b są średnimi czasami przebywania w stanach $-a$ oraz b . Jeżeli założymy że

$$b\mu_a = a\mu_b \quad (3.146)$$

to proces jest stacjonarny o zerowej wartości średniej. Funkcja korelacyjna ma postać funkcji eksponencjalnej (podobnie jak dla szumu Ornsteina-Uhlenbecka)

$$C(t) = ab \exp\left(-\frac{|t|}{\tau_c}\right), \quad (3.147)$$

gdzie czas korelacji szumu τ_c dany jest przez formułę

$$1/\tau_c = \mu_a + \mu_b \quad (3.148)$$

Szum jest symetryczny gdy $a = b$ i to implikuje że $\mu_a = \mu_b$ gdy $\langle \xi(t) \rangle = 0$. W pewnych przypadkach szum ten może opisywać losowe przejścia pomiędzy dwoma metastabilnymi stanami w układach bistabilnych.

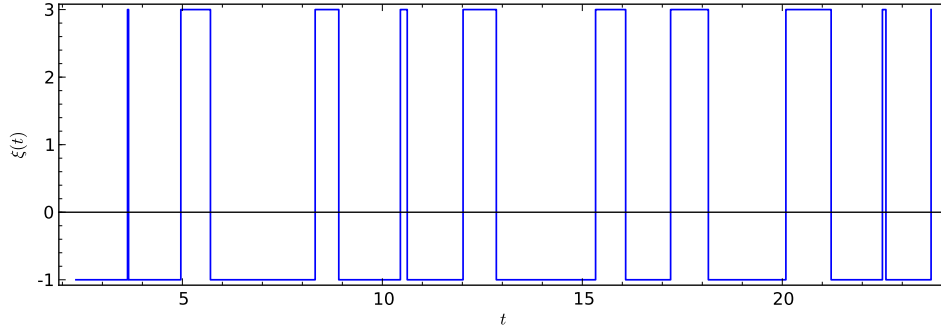
```
# definicja symetrycznego
# szumu dychotomicznego
a = -1
b = 3
stan = [a,b]
mu_a = 1
mu_b = mu_a * abs(b) / abs(a)
mu = [mu_a, mu_b]

# realizacja
N = 20
czasy = [-log(random()/mu[i%2]) for i in range(N)]
punkty = cumsum(czasy)
stany = [stan[i%2] for i in range(N)]

# wizualizacja
plot_step_function(zip(punkty, stany))
```

3.4.5 Szum kangura

Obserwacje kangurów prowadzą do ciekawych uogólnień w teorii procesów stochastycznych. Kangur od czasu do czasu skacze. Skoki kangura są o różnej długości. Pomiedzy skokami



Rysunek 3.8: Reliazcja szumu dychotomicznego.

kangur spoczywa przez jakiś czas. To zachowanie kangura prowadzi do następującego opisu procesu losowego kangura $\xi(t)$: Jest to proces nieciągły i stacjonarny w którym prawdopodobieństwo przejścia $\mathcal{W}(z|z_0)$ ze stanu z_0 do stanu z faktoryzuje się, tzn.

$$\mathcal{W}(z|z_0) = Q(z)\nu(z_0) \quad (3.149)$$

To oznacza, że kangur skacze ze stanu z_0 ze średnią częstością $\nu(z_0)$. Wielkość $\tau(z_0) = 1/\nu(z_0)$ to średni czas przebywania w stanie z_0 . Prawdopodobieństwo, że kangur skoczy do stanu z wynosi $Q(z)$ i jest unormowane do 1 w przestrzeni stanów kangura. Odpowiednie równanie ewolucji dla gęstości prawdopodobieństwa $p(z, t)$ dla procesu kangura ma postać:

$$\frac{\partial p(z, t)}{\partial t} = -\nu(z)p(z, t) + Q(z) \int_{-\infty}^{\infty} \nu(\eta)p(\eta, t)d\eta \quad (3.150)$$

W przypadku stanów stacjonarnych, gdy $p(z, t) = p(z)$, otrzymujemy z powyższego równania związek pomiędzy rozkładem stacjonarnym $p(z)$ oraz gęstością prawdopodobieństwa $Q(z)$:

$$Q(z) = \frac{\nu(z)p(z)}{\int_{-\infty}^{\infty} \nu(\eta)p(\eta, t)d\eta} = \frac{\nu(z)p(z)}{\langle \nu \rangle} \quad (3.151)$$

Dla symetrycznego procesu kangura funkcja korelacyjna jest dana przez relację (nie jest łatwym zadaniem to pokazać)

$$C(t) = 2 \int_0^{\infty} z^2 p(z) \exp(-\nu(z)|t|) dz, \quad (3.152)$$

gdzie $p(z) = p(-z)$ jest rozkładem stacjonarnym procesu $\xi(t)$ oraz $\nu(z) = \nu(-z)$. W tym przypadku wartość średnia procesu jest zero, $\langle \xi(t) \rangle = 0$.

Specjalnym przypadkiem tego procesu jest tzw. szum Kubo-Andersona gdy $\nu(z) = \nu_0$. Z rów. (3.151) wynika, że dla stałej częstości przeskoków stacjonarna gęstość $p(z) = Q(z)$. Wówczas z powyższej relacji otrzymujemy jawną postać funkcji korelacyjnej:

$$C(t) = \langle z^2 \rangle \exp\left(-\frac{|t|}{\tau_c}\right) \quad (3.153)$$

gdzie czas korelacji szumu wynosi $\tau_c = 1/\nu_0$ oraz $\langle z^2 \rangle$ jest wartością średnią obliczoną z rozkładem stacjonarnym $p(z) = Q(z)$.

Oto dwa przykłady szumów kangura:

1. proces $\xi(t)$ jest nieograniczony, określony na przedziale $(-\infty, \infty)$ i stacjonarna gęstość prawdopodobieństwa jest gaussowska:

$$p(z) = Q(z) = \frac{1}{\sqrt{2\pi}\sigma} \exp(-z^2/2\sigma^2), \quad \xi(t) \in (-\infty, \infty) \quad (3.154)$$

2. proces $\xi(t)$ jest ograniczony na przedziale $[-l, l]$ i ma jednostajną gęstość prawdopodobieństwa

$$p(z) = Q(z) = \frac{1}{2l} \theta(z+l) \theta(l-z), \quad \xi(t) \in [-l, l], \quad (3.155)$$

gdzie $\theta(x)$ jest funkcją schodkową Heaviside'a.

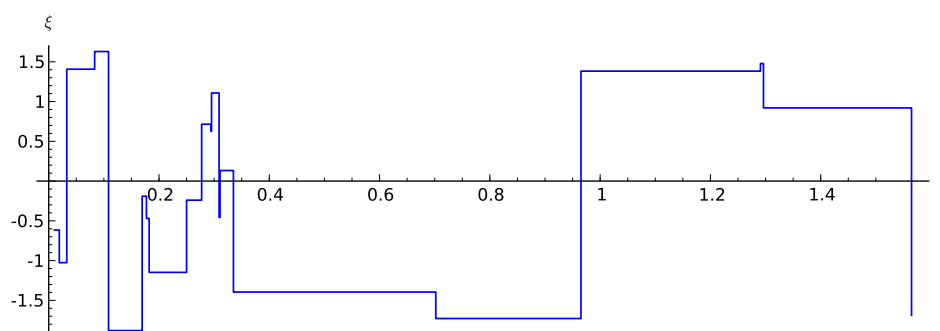
Na początek realizacja procesu ograniczonego.

```
#szum kangura
#szum Kubo - Andersona
#stała częstość \nu_0
#proces ograniczony

l = 2
N = 20
ksi = [2*1*random() - 1 for i in range(N)]

nu_0 = 2.2
czasy = stats.expon.rvs(scale=1/nu_0, size=N)

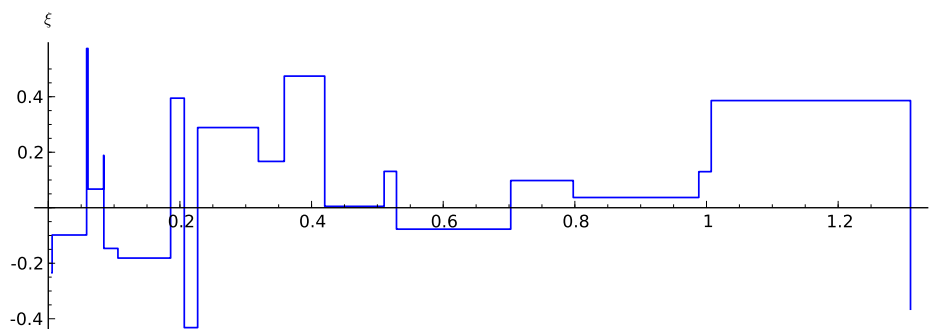
plot_step_function(zip(czasy, ksi))
```



Rysunek 3.9: Reliazcja ograniczonego szumu kangura ze stałą częstością ν_0 (szum Kubo - Andersona).

A teraz wykreślimy sobie realizację procesu kangura dla sytuacji gdy proces $\xi(t)$ jest nieograniczony. Jedyne co trzeba zmienić w powyższym kodzie to linijki definiujące ów proces jako nieograniczony.

```
sigma = 0.3
ksi = [normalvariate(0, sigma) for i in range(N)]
```



Rysunek 3.10: Reliacja nieograniczonego szumu kangura ze stałą częstością ν_0 (szum Kubo - Andersona).

3.4.6 Uwagi ogólne: proces Markowa i proces niemarkowski

Rozważmy dla prostoty równanie stochastyczne w postaci:

$$\dot{x} = F(x) + G(x)\xi(t) \quad (3.156)$$

gdzie $\xi(t)$ jest dowolnym opisanym powyżej szumem. Co możemy powiedzieć o procesie stochastycznym $x = x(t)$ generowanym przez powyższe równanie. Jeżeli szum $\xi(t)$ is białym szumem wówczas proces $x(t)$ jest procesem Markowa. To jest ogromna zaleta ponieważ proces Markowa jest całkowicie opisany gdy znamy

1. warunkową gęstość prawdopodobieństwa $p(x_2, t_2 | x_1, t_1)$
2. stan początkowy procesu $x(t)$ dany przez rozkład 1-wymiarowy $p(x, 0)$. Wówczas spełnione jest równanie Chapmana-Kołmogorowa z którego można wyprowadzić równanie Fokkera-Plancka lub Kołmogorowa-Fellera. Rozkłady wielowymiarowe wyrażają się przez powyższe 2 funkcje.

Jeżeli szum nie jest szumem białym (jest np. eksponencjalnie skorelowanym) to proces $x(t)$ jest procesem niemarkowskim. Rozkłady wielowymiarowe nie można otrzymać z powyższych 2 funkcji. Wówczas wiemy niewiele o takim procesie i jego analiza bywa bardzo złożona.

3.5 Proces Ornsteina-Uhlenbecka

Rozważmy najprostsze liniowe równanie różniczkowe

$$\dot{x} = -ax, \quad a > 0, \quad x(0) = x_0 \quad (3.157)$$

Jego rozwiązaniem jest funkcja eksponencjalna:

$$x(t) = x_0 e^{-at} \quad (3.158)$$

która zawsze dąży do zera gdy $t \rightarrow \infty$.

Zaburzymy losowo to równanie:

$$\dot{x} = -ax + a\Gamma(t), \quad \langle \Gamma(t) \rangle = 0, \quad \langle \Gamma(t)\Gamma(s) \rangle = 2D\delta(t-s) \quad (3.159)$$

Równanie to opisuje proces stochastyczny $x = x(t)$ który nazywa się procesem Ornsteina-Uhlenbecka. Proces ten jest procesem Markowa (tw. Dooba). Proces ten jest procesem Gaussa ponieważ jest to transformacja liniowa innego procesu Gaussa, a mianowicie procesu $\Gamma(t)$. Dlatego też proces ten można analizować na wiele sposobów. Po pierwsze, jest to równanie różniczkowe liniowe niejednorodne. Dlatego jego rozwiązaniem jest funkcja

$$x(t) = x_0 e^{-at} + a \int_0^t e^{-a(t-s)} \Gamma(s) ds \quad (3.160)$$

co można sprawdzić bezpośrednio, różniczkując powyższą funkcję.

Można teraz obliczyć wartość średnią tego procesu:

$$\mu(t) = \langle x(t) \rangle = \langle x_0 \rangle e^{-at} + a \int_0^t e^{-a(t-s)} \langle \Gamma(s) \rangle ds = \langle x_0 \rangle e^{-at} \quad (3.161)$$

gdzie wykorzystaliśmy własności uśredniania (średnia z sumy to suma średnich, średnia ze stałej wielkości to ta sama stała wielkość, całka to wartość graniczna sumy, itd.).

Obliczenie funkcji korelacyjnej jest bardziej pracochłonne. Trzeba z należytą ostrożnością wymnożyć $x(t)$ oraz $x(\tau)$:

$$\langle x(t)x(\tau) \rangle = \left\langle \left[x_0 e^{-at} + a \int_0^t e^{-a(t-s)} \Gamma(s) ds \right] \left[x_0 e^{-a\tau} + a \int_0^\tau e^{-a(\tau-u)} \Gamma(u) du \right] \right\rangle \quad (3.162)$$

Pamiętając, że $\langle \Gamma(t) \rangle = 0$ dla dowolnego czasu t , otrzymamy wyrażenie:

$$\langle x(t)x(\tau) \rangle = \langle x_0^2 \rangle e^{-a(t+\tau)} + a^2 e^{-a(t+\tau)} \int_0^t e^{as} \int_0^\tau e^{au} \langle \Gamma(s)\Gamma(u) \rangle ds du \quad (3.163)$$

W wyrażeniu podcałkowym występuje funkcja korelacyjna białego szumu gaussowskiego, więc

$$\langle x(t)x(\tau) \rangle = \langle x_0^2 \rangle e^{-a(t+\tau)} + 2Da^2 e^{-a(t+\tau)} \int_0^t e^{as} \int_0^\tau e^{au} \delta(s-u) ds du \quad (3.164)$$

Należy obliczyć całki z δ - Diraca. Trzeba rozważyć 2 przypadki: gdy $t > \tau$ oraz gdy $\tau > t$. W końcu otrzymamy wynik w postaci:

$$\langle x(t)x(\tau) \rangle = [\langle x_0^2 \rangle - aD] e^{-a(t+\tau)} + aD e^{-a|t-\tau|} \quad (3.165)$$

Ponieważ $x(t)$ jest procesem gaussowskim, możemy podać gęstość prawdopodobieństwa tego procesu:

$$p(x, t) = \frac{1}{\sqrt{2\pi\sigma^2(t)}} \exp \left\{ -\frac{[x - \mu(t)]^2}{2\sigma^2(t)} \right\} \quad (3.166)$$

Wartość średnia jest dana powyżej. Musimy jeszcze obliczyć wariancję:

$$\sigma^2(t) = \langle x^2(t) \rangle - \langle x(t) \rangle^2 = \sigma_0^2 e^{-2at} + aD[1 - e^{-2at}] \quad (3.167)$$

Stacjonarny rozkład prawdopodobieństwa $p_{st}(x)$ to rozkład dla $t \rightarrow \infty$. Granice taką łatwo wyznaczyć ponieważ

$$\mu(t) \rightarrow 0, \quad \sigma^2(t) \rightarrow aD \quad (3.168)$$

Więc rozkład prawdopodobieństwa $p_{st}(x)$ ma postać gaussowską:

$$p_{st}(x) = \frac{1}{\sqrt{2\pi aD}} \exp \left\{ -\frac{x^2}{2aD} \right\} \quad (3.169)$$

W ogólności $x(t)$ nie jest procesem stacjonarnym. Jeżeli tak wybrać stan początkowy aby

$$\langle x_0^2 \rangle = aD \quad (3.170)$$

to proces staje się stacjonarny i jest eksponencjalnie skorelowany:

$$C(t, \tau) = C(t - \tau) = \langle x(t)x(\tau) \rangle = aDe^{-a|t-\tau|} = \frac{D}{\tau_c} \exp \left[-\frac{|t - \tau|}{\tau_c} \right] \quad (3.171)$$

gdzie $\tau_c = 1/a$ jest czasem korelacji procesu. Ale wybór $\langle x_0^2 \rangle = aD$ oznacza, że stan początkowy jest stanem stacjonarnym, czyli $p(x, 0) = p_{st}(x)$. Innymi słowy, proces Ornsteina-Uhlenbecka zawsze staje się stacjonarny albo dla długich czasów albo dla początkowego stanu który jest stanem stacjonarnym. W tym przypadku proces ten służy jako standardowy przykład eksponencjalnie skorelowanego szumu gaussowskiego. Wyznamy teraz widmo mocy tego szumu:

$$S(\omega) = \int_{-\infty}^{\infty} e^{i\omega t} C(t) dt = aD \int_{-\infty}^{\infty} e^{i\omega t} e^{-a|t|} dt = \frac{2a^2 D}{a^2 + \omega^2} = \frac{2D}{1 + (\tau_c \omega)^2} \quad (3.172)$$

Poniżej przedstawiamy tę zależność dla kilku wartości czasu korelacji $\tau_c = 1/a$. Widmo mocy ma maksimum, podobnie jak światło o danym kolorze ma maksimum przy częstotliwości fali promieniowania elektromagnetycznego. Jeżeli czas korelacji $\tau_c \rightarrow 0$ to widmo mocy staje się coraz bardziej płaskie i dąży do stałej wartości $2D$. Z kolei funkcja korelacyjna

$$\lim_{\tau_c \rightarrow 0} \frac{D}{\tau_c} \exp \left[-\frac{|t - \tau|}{\tau_c} \right] = 2D\delta(t - \tau) \quad (3.173)$$

i szum staje się białym szumem gaussowskim.

```
var('x, t, D')
a = 1
V = a*x^2

#wartosci startowe
N = 10
x0 = [random()*4 - 2 for i in range(N)]
meanx0 = mean(x0)
sigmax0 = variance(x0)

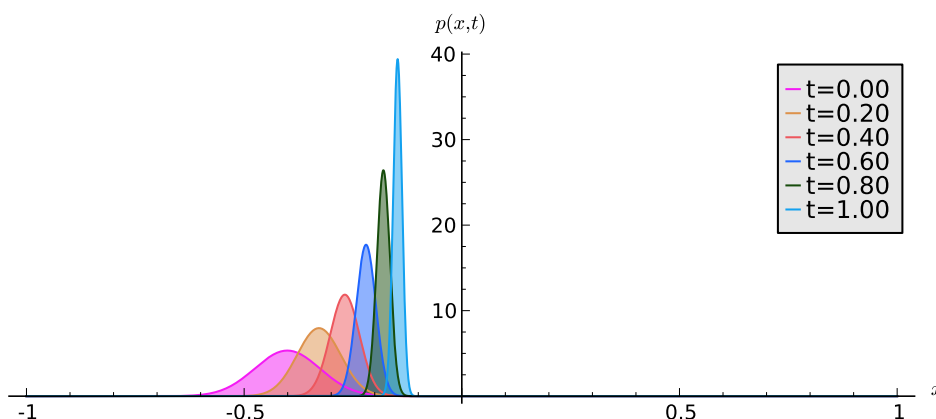
#p(x, t)
mu = mean(x0) * exp(-a*t)
sigma = (variance(x0) - a*D)*exp(-2*a*t)
p = exp(-(x-mu)^2/(2*sigma^2))/sqrt(2*pi*sigma^2)
show(p)

P = []
lT = srange(0, 1, 0.2, include_endpoint=1)
```

```

for t1 in lT:
    c = [random() for i in range(3)]
    P.append(plot(p(x=x,t=t1,D=1), (x,-1,1),
        fill=0,
        color=c,
        fillcolor=c,
        legend_label='t=%0.2f' %t1))

```



Rysunek 3.11: Gęstość prawdopodobieństwa $p(x, t)$.

3.5.1 Równanie Fokkera-Plancka dla procesu Ornsteina-Uhlenbecka

Zgodnie z regułami podanymi w poprzednim rozdziale, równanie (3.159) to równanie Langevina, dla którego dryf i funkcja dyfuzji mają postać

$$F(x, t) = -ax, \quad G(x, t) = a \quad (3.174)$$

Odpowiednie równanie Fokkera-Plancka dla warunkowej gęstości prawdopodobieństwa $p(x, t|y, s)$ ma postać:

$$\frac{\partial p(x, t|x_0, s)}{\partial t} = a \frac{\partial}{\partial x} [xp(x, t|x_0, s)] + Da^2 \frac{\partial^2}{\partial x^2} p(x, t|x_0, s), \quad (3.175)$$

$$p(x, s|x_0, s) = \delta(x - x_0)$$

Z kolei dla 1-wymiarowej gęstości prawdopodobieństwa $p(x, t)$ należy rozwiązać zagadnienie:

$$\frac{\partial p(x, t)}{\partial t} = a \frac{\partial}{\partial x} [xp(x, t)] + Da^2 \frac{\partial^2}{\partial x^2} p(x, t), \quad p(x, 0) = p(x) \quad (3.176)$$

gdzie $p(x)$ jest początkowym rozkładem prawdopodobieństwa. Warto przypomnieć, że zachodzi następująca relacja:

$$p(x, t) = \int_{-\infty}^{\infty} p(x, t|y, s) p(y, s) dy \quad (3.177)$$

czyli $p(x, t|y, s)$ jest propagatorem: propaguje rozkład w chwili wcześniejszej s do chwili późniejszej t . W szczególności

$$p(x, t) = \int_{-\infty}^{\infty} p(x, t|y, 0) p(y, 0) dy \quad (3.178)$$

propaguje rozkład początkowy do chwili t .

Oba równania Fokkera-Plancka można rozwiązać metodą funkcji charakterystycznej czyli metodą transformaty Fouriera wybierając odpowiednie warunki początkowe. Wynikiem zastosowania tej metody jest równanie różniczkowe cząstkowe pierwszego rzędu dla funkcji charakterystycznej. Takie równanie zawsze można rozwiązać tzw. metodą charakterystyk. Ponieważ typowy student nie zna tej metody, nie będziemy jej tu przedstawiać. Postąpimy inaczej. W zasadzie już rozwiązaliśmy podobny problem. Załóżmy, że dana jest wartość początkowa procesu: $x(s) = x_0$. Wybieramy jako chwilę początkową czas s . Ponieważ w równaniu Fokkera-Plancka nie występuje czas w sposób jawny, rozwiązanie $p(x, t|x_0, s)$ zależy tylko od różnicy czasów, tzn. $p(x, t|x_0, s) = p(x, t-s|x_0)$. Dlatego też w rozwiązaniu (3.167) oraz w (3.178) należy dokonać podmiany: $\langle x_0^k \rangle \rightarrow x_0^k$ oraz $t \rightarrow t-s$. W ten sposób otrzymamy warunkową gęstość rozkładu:

$$p(x, t|x_0, s) = \frac{1}{\sqrt{2\pi\sigma^2(t, s)}} \exp \left\{ -\frac{[x - x_0 e^{-a(t-s)}]^2}{2\sigma^2(t, s)} \right\} \quad (3.179)$$

gdzie teraz

$$\sigma^2(t, s) = aD[1 - e^{-2a(t-s)}] \quad (3.180)$$

Można sprawdzić, korzystając z twierdzenia o ciągach δ -podobnych, że w granicy $t \rightarrow s$ gęstość $p(x, t|x_0, s) \rightarrow \delta(x - x_0)$. Dla dowolnego warunku początkowego, korzystamy z rów. (3.178) i otrzymujemy rozkład w dowolnej chwili z dowolnym warunkiem początkowym.

```
var('x, t, s, x0, a, D')
```

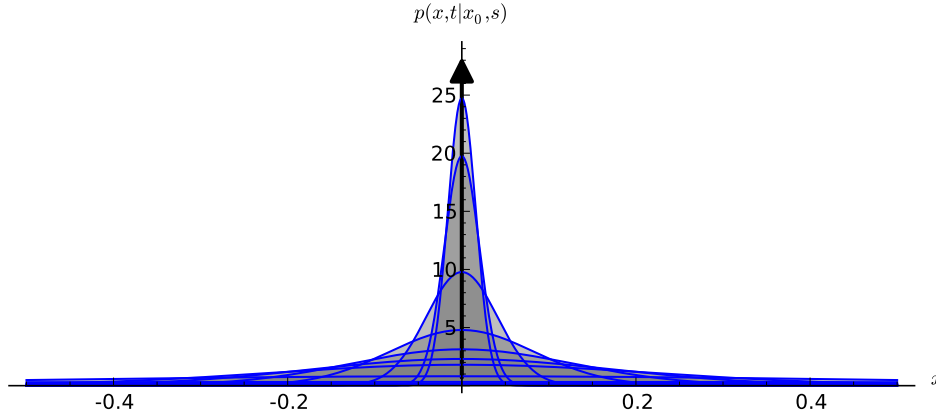
```
sigma(t,s,a,D) = a*D*(1 - exp(-2*a*(t-s)))
pc(x,t,s,x0,a,D) = exp(-(x-x0*exp(-a*(t-s)))^2/(2*sigma(t,s,a,D)^2))
pc \= sqrt(2*pi*sigma(t,s,a,D)^2)
```

```
p=[]
lT = srange(0,1,0.2) + srange(0.9,1,0.02)
for t1 in lT:
    c = [random() for i in range(3)]
    p.append(plot(pc(x=x,t=t1,s=1,x0=0,a=1,D=1), (x,-0.5,0.5), fill=0))
p.append(arrow((0,0), (0,29), color='black'))
```

3.5.2 Zastosowanie: cząstka Browna

Niech cząstka Browna o masie m porusza się w płynie. Można zbudować równanie Newtona dla takiej cząstki :

$$m\ddot{x} + \gamma\dot{x} = \Gamma(t) \quad \text{czyli} \quad m\dot{v} + \gamma v = \Gamma(t) \quad \text{gdzie} \quad v = \dot{x} \quad (3.181)$$



Rysunek 3.12: Ciąg δ -podobny.

Jest to równanie Langevina, gdzie szum $\gamma(t)$ to losowa siła działająca na cząstkę i mająca swe źródło w zderzeniach cząstek płynu z cząstką Browna. Równanie Langevina dla prędkości cząstki Browna ma taką samą postać jak równanie generujące proces Ornsteina-Uhlenbecka:

$$\dot{v} = -(\gamma/m)v + \Gamma(t)/m = -\gamma_0 v + \gamma_0 \Gamma_0(t) \quad (3.182)$$

gdzie przeskalowane wielkości mają postać

$$\gamma_0 = \frac{\gamma}{m}, \quad \Gamma_0(t) = \frac{\Gamma(t)}{\gamma} \quad (3.183)$$

Dla przeskalowanej siły losowej mamy związki:

$$\langle \Gamma_0(t) \rangle = 0, \quad \langle \Gamma_0(t) \Gamma_0(s) \rangle = \frac{2D}{\gamma^2} \delta(t-s) = 2D_0 \delta(t-s) \quad (3.184)$$

Porównując to z równaniem dla procesu Ornsteina-Uhlenbecka otrzymujemy korespondencję:

$$a \rightarrow \gamma_0 = \frac{\gamma}{m}, \quad D \rightarrow D_0 = \frac{D}{\gamma^2} \quad (3.185)$$

Posiłkując się wynikami dla procesu Ornsteina-Uhlenbecka, otrzymamy dla prędkości cząstki Browna następujące wyniki:

$$\langle v(t) \rangle = \langle v_0 \rangle e^{-\gamma t/m}, \quad \langle v(t) v(s) \rangle = \frac{D}{\gamma m} e^{-\gamma|t-s|/m} \quad (3.186)$$

przy założeniu, że jest to proces stacjonarny, czyli że w chwili początkowej

$$\langle v_0^2 \rangle = \frac{D}{\gamma m} \quad (3.187)$$

Zauważmy, że w dowolnej chwili czasu zachodzi relacja

$$\langle v^2(t) \rangle = \frac{D}{\gamma m} \quad (3.188)$$

Z twierdzenia o ekwipartycji energii (fizyka statystyczna) wynika, że

$$\langle E_k \rangle = \frac{1}{2} m \langle v^2(t) \rangle = \frac{1}{2} kT \quad \text{czyli} \quad \langle v^2(t) \rangle = \frac{kT}{m} = \frac{D}{\gamma m} \quad (3.189)$$

gdzie k jest stałą Boltzmanna i T jest temperaturą otoczenia (termostatu, płynu). Stąd otrzymujemy ważny wynik, że natężenie szumu termicznego (fluktuacji termicznych) wynosi

$$D = \gamma kT \quad (3.190)$$

Kolejny ważny wynik to czas korelacji prędkości cząstki Browna. Z relacji dla funkcji korelacyjnej wnioskujemy, że czas korelacji

$$\tau_c = \frac{m}{\gamma} \quad (3.191)$$

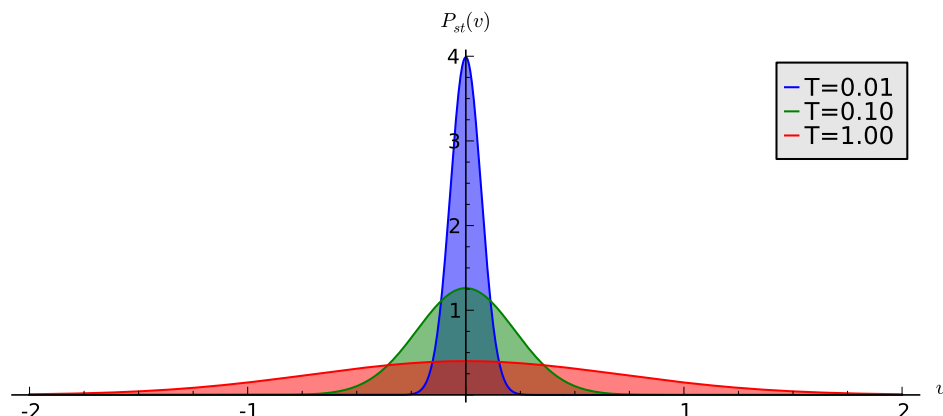
Prędkość cząstki Browna to losowy proces gaussowski, eksponencjalnie skorelowany. Stacjonarny rozkład prawdopodobieństwa $p_{st}(v)$ ma postać (patrz pierwsza część tego rozdziału)

$$p_{st}(v) = \frac{1}{\sqrt{2\pi aD}} \exp \left\{ -\frac{v^2}{2aD} \right\} = \sqrt{\frac{m}{2\pi kT}} \exp \left\{ -\frac{mv^2}{kT} \right\} \quad (3.192)$$

Jest to dobrze znany rozkład Maxwella.

```
var('v, T, m')
k = 1
pst(v,T,m) = exp(-m*v^2/k/T)*sqrt(m/2/pi/k/T)

p=[]
kolor = ['blue','green','red']
lT = [0.01,0.1,1]
for t in lT:
    color = kolor[lT.index(t)]
    p.append(plot(pst(v=x,T=t,m=1),(x,-2,2),
        color=color,
        legend_label='T=%.2f'%t,
        fill=0,fillcolor=color))
```



Rysunek 3.13: Rozkład Maxwella

3.6 Geometryczny ruch browna

Założmy, że posiadamy określoną kwotę pieniędzy i chcemy ją ulokować w banku, który w zamian oferuje odsetki w ramach proponowanych stóp procentowych. W chwili t posiadamy $X(t)$ złotych. Ile dostaniemy pieniędzy z banku po czasie $t + \Delta t$. Oznaczmy kwotę po tym czasie jako $X(t + \Delta t)$. Kwota po czasie $t + \Delta t$ to kwota jaka ulokowaliśmy plus kwota z oprocentowania:

$$X(t + \Delta t) = X(t) + \delta X(t) \quad (3.193)$$

Pierwszy składnik jest kwotą jaką lokujemy w chwili t . Drugi składnik jest kwotą jaką otrzymamy z oprocentowania lokaty. Ile wynosi ten dodatek? Ta dodatkowa kwota to

$$\delta X(t) = cX(t)\Delta t \quad (3.194)$$

Wyrażenie to ma jasną interpretację: Im więcej ulokujemy w chwili t (tzn. większe $X(t)$) tym więcej otrzymamy; im dłużej będzie trwała lokata (tzn. większe Δt) tym więcej otrzymamy. Współczynnik c zależy od stopy procentowej lokaty: im większe oprocentowanie tym większa wartość c i tym więcej otrzymamy z lokaty. Uwzględniając te dwa składniki otrzymamy równanie na przyrost naszych pieniędzy

$$X(t + \Delta t) - X(t) = cX(t)\Delta t \quad (3.195)$$

Założmy teraz, że oprocentowanie scharakteryzowane przez wielkość c nie jest ustalone, ale w każdej chwili waha się losowo, to znaczy

$$c \rightarrow c + \xi(t) \quad (3.196)$$

gdzie $\xi(t)$ opisuje losowe wahania oprocentowania. Innymi słowy jest to jakiś proces stochastyczny. Wówczas nasze równanie będzie miało postać

$$X(t + \Delta t) - X(t) = [c + \xi(t)] X(t)\Delta t \quad (3.197)$$

Z lewej strony mamy przyrost naszych pieniędzy

$$\Delta X(t) = X(t + \Delta t) - X(t) \quad (3.198)$$

Jeżeli teraz Δt jest nieskończenie małe, to nasze równanie ma postać równania stochastycznego

$$dX(t) = cX(t)dt + \xi(t) X(t)dt \quad (3.199)$$

Banki nie stosują losowych wahań oprocentowania, ale powyższy model można zastosować do cen akcji na giełdzie. Tam ceny zmieniają się w każdej chwili i w tych zmianach można odnaleźć część przewidywalnych (deterministycznych) zmian opisywanych parametrem c i część zmian losowych opisywanych funkcją losową $\xi(t)$. Jeżeli te zmiany podobne są do losowych zmian położenia cząstki Browna, to $\xi(t)$ jest białym szumem Gaussowskim

$$\xi(t) = \Gamma(t) \quad (3.200)$$

Jak wiemy biały szum Gaussowski jest pochodną procesu Wienera $W(t)$, to znaczy

$$\Gamma(t) = \frac{dW(t)}{dt} \quad (3.201)$$

lub równoważnie

$$\Gamma(t)dt = dW(t) \quad \langle dW(t) \rangle = 0, \quad \langle [dW(t)]^2 \rangle = 2Ddt \quad (3.202)$$

Stąd wynika, że równanie przyjmuje postać

$$dX(t) = cX(t)dt + X(t)dW(t) \quad (3.203)$$

gdzie c nazywa się średnią stopą zwrotu, a parametr D nazywa się zmiennością (chwiejnością) cen akcji. Równanie to ma postać równania Ito i dlatego wnioskujemy, że proces stochastyczny $X(t)$ jest procesem Markowa. Ponadto jest to proces dyfuzji opisywany równaniem Fokkera-Plancka. Równanie to zostało zaproponowane do opisu zjawisk ekonomicznych na przełomie lat pięćdziesiątych i sześćdziesiątych XX w. niezależnie przez Osborne'a (1959) i Samuelsona (1965). Równanie to opisuje proces stochastyczny, który nazywa się w literaturze *geometrycznym procesem Wienera*. Równanie to jest jednym z podstawowych elementów modelu wyceny opcji Blacka-Scholesa. Teoria ta została nagrodzona Nagrodą Nobla z ekonomii w roku 1997, a opracowana przez absolwenta fizyki i doktora matematyki Fischera Blacka oraz ekonomistę Myrona Scholesa. Teoria Blacka-Scholesa pozwala na wycenę wartości tzw. finansowych instrumentów pochodnych, czyli opcji, oraz służy do optymalizacji “bezpiecznego” portfela inwestycyjnego.

Równanie powyższe jest jedną z możliwych propozycji. Po pierwsze, należy rozstrzygnąć problem interpretacji tego równania: czy jest to wersja Ito czy wersja Stratonowicza, a może jeszcze inna. Po drugie, proces Wienera można zastąpić innymi procesami. Mogą to być różne odmiany procesu Poissona czy ogólniej Levy'ego. Wszystkie te procesy to procesy o przyrostach niezależnych. Można modelować fluktuacje cen akcji przez procesy skorelowane, procesy o zależnych przyrostach. Widać, że ten problem cen akcji na giełdzie daje niesłychanie szerokie pole do modelowania. To z kolei pozwala na daleko idące modyfikacje innych modeli, jak na przykład modelu Blacka-Scholesa.

```
h = 0.01
x0 = 1
c = 0.1
D = 0.01

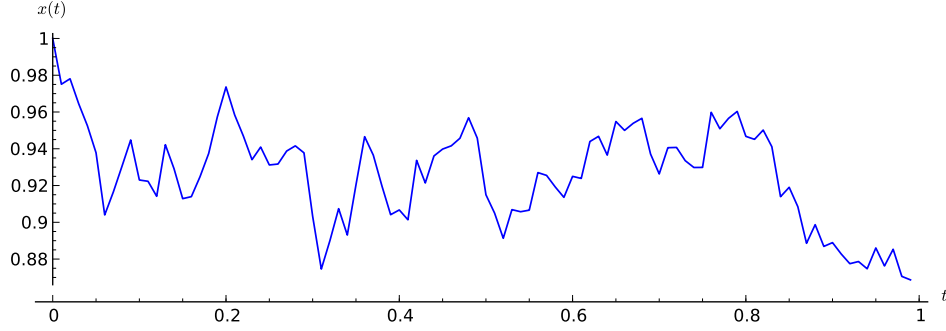
x = [x0]
N = 100
for i in xrange(N):
    x.append((1 + c*h + sqrt(2*h*D)*normalvariate(0,1))*x[i])
list_plot(zip(srange(0,N*h,h),x))
```

3.6.1 Analiza geometrycznego ruchu Browna

Założymy (wzorem większości autorów), że równanie

$$dX(t) = \mu X(t)dt + X(t)dW(t) \quad (3.204)$$

$$X(t) \geq 0, \quad \langle dW(t) \rangle = 0, \quad \langle [dW(t)]^2 \rangle = 2Ddt$$



Rysunek 3.14: Relizacja geometrycznego ruchu Browna.

jest równaniem w interpretacji Ito. Później zbadamy to równanie gdy założymy interpretację Stratonowicza. Równanie Fokkera-Plancka dla dla 1-wymiarowej gęstości prawdopodobieństwa $p(x, t)$ ma postać

$$\begin{aligned} \frac{\partial p(x, t)}{\partial t} &= -c \frac{\partial}{\partial x} [xp(x, t)] + D \frac{\partial^2}{\partial x^2} x^2 p(x, t), \\ p(x, 0) &= p(x), \quad x \geq 0 \end{aligned} \quad (3.205)$$

gdzie $p(x)$ jest początkowym rozkładem prawdopodobieństwa. Rozwiążemy to równanie metodą zamiany zmiennych. Zdefiniujemy nową zmienną y jako

$$y = \ln x, \quad x = e^y \quad (3.206)$$

Jeżeli dokonujemy zamiany zmiennych to musimy pamiętać, że gęstość prawdopodobieństwa także ulega transformacji wg zależności

$$p(x, t)dx = f(y, t)dy, \quad (3.207)$$

czyli

$$f(y, t) = p(x, t) \frac{dx}{dy} = e^y p(x, t) = xp(x, t) \quad (3.208)$$

oraz

$$p(x, t) = f(y, t) \frac{dy}{dx} = \frac{1}{x} f(y = \ln x, t). \quad (3.209)$$

Następnie musimy zamienić różniczkowanie względem x na różniczkowanie względem y :

$$\frac{\partial}{\partial x} [xp(x, t)] = \frac{\partial y}{\partial x} \frac{\partial}{\partial y} f(y, t) = \frac{1}{x} \frac{\partial}{\partial y} f(y, t) \quad (3.210)$$

oraz

$$\begin{aligned} \frac{\partial^2}{\partial x^2} [x^2 p(x, t)] &= \frac{\partial}{\partial x} \frac{\partial}{\partial x} [xf(y, t)] = \\ &= \frac{\partial}{\partial x} \left\{ \frac{\partial y}{\partial x} \frac{\partial}{\partial y} [xf(y, t)] \right\} = \frac{\partial}{\partial x} \left\{ \frac{1}{x} \frac{\partial}{\partial y} [e^y f(y, t)] \right\} = \\ &= \frac{1}{x} \frac{\partial}{\partial y} \left\{ \frac{1}{x} \left(e^y f(y, t) + e^y \frac{\partial}{\partial y} f(y, t) \right) \right\} = \frac{1}{x} \frac{\partial}{\partial y} \left\{ f(y, t) + \frac{\partial}{\partial y} f(y, t) \right\} \end{aligned} \quad (3.211)$$

W obu wyrażeniach występuje wyraz $1/x$, który przeniesiemy na lewą stronę równania Fokkera-Plancka: $x p(x, t) = f(y, t)$. Ostatecznie otrzymamy przetransformowane równanie Fokkera Plancka w postaci

$$\frac{\partial f(y, t)}{\partial t} = -[c - D] \frac{\partial}{\partial y} f(y, t) + D \frac{\partial^2}{\partial y^2} f(y, t) \quad (3.212)$$

Takie równanie spełnia proces, którego równanie Ito ma postać:

$$dY = (c - D)dt + dW(t) \quad (3.213)$$

Równanie to łatwo scałkować:

$$Y(t) = Y(0) + (c - D)t + W(t) \quad (3.214)$$

Ponieważ jest to liniowa transformacja procesu Gaussa $W(t)$ to proces $Y(t)$ jest także procesem Gaussa. Więc nie musimy rozwiązywać równani Fokkera-Plancka, aby podać jawną postać gęstości rozkładu $f(y, t)$. Wystarczy obliczyć wartość średnią oraz wariancję procesu $Y(t)$:

$$\begin{aligned} \mu_y(t) &= \langle Y(t) \rangle = \langle Y(0) + (c - D)t + W(t) \rangle = \langle Y(0) \rangle + (c - D)t, \\ \langle Y^2(t) \rangle &= \langle [Y(0) + (c - D)t + W(t)]^2 \rangle = \langle Y^2(0) \rangle + 2(c - D)t \langle Y(0) \rangle + (c - D)^2 t^2 + 2Dt \end{aligned} \quad (3.215)$$

$$\begin{aligned} \sigma_y^2(t) &= \langle Y^2(t) \rangle - \langle Y(t) \rangle^2 = \sigma_y^2(0) + 2Dt, \\ \sigma_y^2(0) &= \langle Y^2(0) \rangle - \langle Y(0) \rangle^2. \end{aligned} \quad (3.216)$$

Gęstość prawdopodobieństwa $f(y, t)$ ma postać funkcji Gaussa:

$$f(y, t) = \frac{1}{\sqrt{2\pi\sigma_y^2(t)}} \exp \left\{ -\frac{[y - \mu_y(t)]^2}{2\sigma_y^2(t)} \right\} \quad (3.217)$$

Wyjściowa gęstość prawdopodobieństwa $p(x, t)$ wyraża się wzorem:

$$p(x, t) = f(y, t) \frac{dy}{dx} = \frac{1}{x} f(y = \ln x, t) = \frac{1}{x \sqrt{2\pi\sigma_y^2(t)}} \exp \left\{ -\frac{[\ln x - \mu_y(t)]^2}{2\sigma_y^2(t)} \right\} \quad (3.218)$$

i nazywa się rozkładem log-normalnym. Poniżej znajdziecie jego wykres.

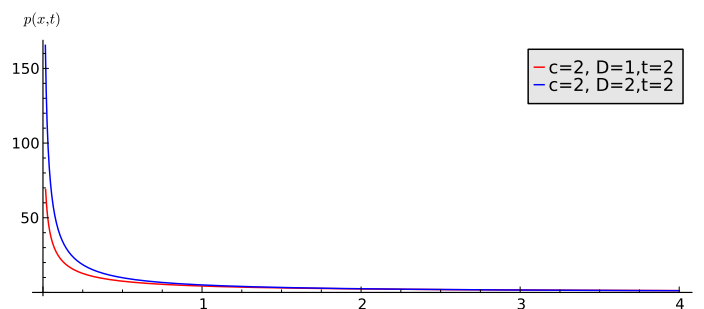
```
var('x,d,c,t')
fl(x,c,d,t) = exp(-(ln(x)-(c-d)*t)^2/(4*pi*t)) / (x*sqrt(4*pi*t))
p=plot(fl(x,2,1,2),(x,0,4),color="red",legend_label="c=2,D=1,t=2")
p+=plot(fl(x,2,2,2),(x,0,4),legend_label="c=2,D=2,t=2")
```

Jeżeli początkowa wartość procesu $X(0) = X_0$ jest dana (wiemy, ile na początku mamy pieniędzy ulokowanych w akcjach) to

$$\sigma_y^2(0) = \langle Y^2(0) \rangle - \langle Y(0) \rangle^2 = \langle [\ln X_0]^2 \rangle - \langle \ln X_0 \rangle^2 = [\ln X_0]^2 - [\ln X_0]^2 = 0 \quad (3.219)$$

W tym przypadku

$$\mu_y(t) = \ln X_0 + (c - D)t, \quad \sigma_y^2(t) = 2Dt \quad (3.220)$$



Rysunek 3.15: Rozkład log-normalny.

Momenty statystyczne wyjściowego procesu $X(t)$ dane są formułą

$$\langle X^n(t) \rangle = \int_0^\infty x^n p(x, t) dx = e^{\mu(t)n + Dtn^2} \quad (3.221)$$

i można je łatwo wyliczyć korzystając z Sage

```
var('x,m,D,n,t')
assume(t>0)
assume(D>0)
assume(n>0)
g = ((e^(n*x))/sqrt(4*pi*D*t))*exp(-(x-m)^2/(4*D*t))
integrate(g,x,-oo,oo)
```

Co powinno obliczyć daną całkę

$$e^{Dn^2t + mn} \quad (3.222)$$

W szczególności wartość średnia i wariancja (fluktuacje) procesu wynoszą

$$\begin{aligned} \mu_x(t) &= \langle X(t) \rangle = X_0 e^{ct}, \\ \sigma_x^2(t) &= \langle X^2(t) \rangle - \langle X(t) \rangle^2 = X_0^2 e^{2ct} [e^{2Dt} - 1]. \end{aligned} \quad (3.223)$$

Zauważmy, że zarówno wartość średnia jak i fluktuacje narastają eksponencjalnie w czasie. Warto też obliczyć kurtozę i skośność dla tego modelu.

Model ten w zastosowaniu do rynków finansowych posiada zarówno pozytywne elementy jak i negatywne elementy. Niewątpliwie pozytywnym elementem jest to że $X(t)$ przyjmuje tylko dodatnie wartości (nie ma ujemnych pieniędzy). Z drugiej strony, parametr D często zmienia się, tu założono jego stałość.

Dodatek numeryczny

4.1 Liczby losowe

Uwaga: W tym kursie do numerycznych realizacji używać będziemy pakietu Sage. Jest to dostępny dla każdego program typu open-source bazujący na języku Python.

4.1.1 Liczby losowe i pseudolosowe

Intuicyjnie dość dobrze rozumiemy co oznacza termin *liczba losowa*. Każdy z nas choć raz w życiu podrzucił monetę do góry po to, by “ślepy los” zdecydował za niego o jakimś wyborze (jeżeli w ten sposób zdecydowaliście o wyborze studiów, to szczerze mówiąc - gratuluję). Oczywiście na monecie nie ma żadnych liczb, ale można sobie potraktować reszkę (R) jako 0 a orła (O) jako 1 (co bardzo dobrze reprezentuje fałsz i prawdę lub niemożliwe i pewne zdarzenie w teorii prawdopodobieństwa). Teraz już możemy sobie podrzucać monetę i na kartce papieru zapisywać kolejne wylosowane (wyrzucone) przez nas liczby

0, 1, 0, 0, 0, 0, 1, 0, 1, 1, 0, 0, 1, 0, 1,

co odpowiada oczywiście wyrzuceniu kolejno

R, O, R, R, R, R, O, R, O, O, R, R, O, R, O.

W naszym przypadku zapiszemy sobie te liczby od razu do listy w notatniku Sage.

```
rzuty = [0, 1, 0, 0, 0, 0, 1, 0, 1, 1, 0, 0, 1, 0, 1]
```

Mamy teraz je dostępne pod zmienną `rzuty`. Do prostych zagadnień, gdzie potrzebne jest nam kilka, czy nawet kilkanaście takich liczb, bez problemu możemy poradzić sobie rzucając monetą. Jeżeli potrzebujemy zdecydować o wyborze pomiędzy trzema możliwościami możemy użyć sześcienną kość do gry i przykładowo wybrać wynik poprzez działanie modulo 3 ($\text{mod}3$). Tym razem dostaniemy trzy możliwe liczby 0, 1, 2

```
# rzuty kością
5, 3, 6, 5, 6, 6, 5, 3, 2, 4, 3, 1, 1, 6, 1,

# modulo 3
2, 0, 0, 2, 0, 0, 2, 0, 2, 1, 0, 1, 1, 0, 1.
```

Sytuacja robi się jednak nieco bardziej skomplikowana, gdy będziemy potrzebować tysiąc, milion czy bilion takich liczb. Jeżeli nawet grupa 10 studentów była by w stanie wyrzucić monetą tysiąc losowych zer i jedynek w pół godziny (włączając w to zapisywanie w liście Sage lub nawet na kartce papieru) to uzyskanie miliona liczb jest praktycznie nie do zrobienia w ten sposób. problem pojawia się też w momencie, gdy chcielibyśmy mieć liczby naturalne z zakresu np: 0 – 10, czy w końcu losowe liczby zmiennoprzecinkowe. Metody chałupnicze w tym momencie się kończą.

Z pomocą może przyjść nam komputer. Obecnie znakomita większość języków programowania (przynajmniej tych realnie wykorzystywanych¹) posiada w swoich standardowych bibliotekach funkcje (metody, klasy) umożliwiające wygenerowanie liczby (pseudo)losowej z przedziału $[0, 1)$ lub też $[0, \text{RAND_MAX}]$, gdzie ów `RAND_MAX` to stała zależna od architektury komputera, kompilatora i bibliotek.

W Sage liczby losowe uzyskuje się poprzez funkcję `random()`. Zwraca ona liczbę losową z przedziału $[0.0, 1.0)$. Wykorzystując proste wyrażenie listowe możemy przypisać do listy `N` liczb losowych.

```
N = 1000
lista = [random() for i in xrange(N)]
```

Inna funkcja `randint(a, b)`, zwraca liczby całkowite z przedziału $[a, b]$. Czyli symulacja rzutu monetą może być zrealizowana poprzez

```
rzut_moneta = [randint(0,1) for i in xrange(N)]
```

Zadanie 2.2.1 Zamodeluj w Sage rzut kością. Wygeneruj listę 1000 liczb odzwierciedlających 1000 rzutów symetryczną sześcienną kością do gry. Wynik zapisz w zmiennej `rzut_kostka`.

Matematycznie rzecz biorąc liczbę losową można utożsamić z wartością jaką przybiera pewna zmienna losowa ξ . Możemy napisać, że dla procesu jakim jest rzut kością zmienna losowa ξ może przybierać wartości 0 lub 1. Matematyczne konsekwencje poznaliście już na wykładzie [Procesy i zjawiska losowe](#), tutaj zajmiemy się znacznie szerzej generowaniem liczb losowych i wykorzystaniem ich właśnie do realizacji procesów losowych, ze szczególnym uwzględnieniem zastosowania dla rynków finansowych, czy w ogólności w modelach ekonomicznych.

No koniec tego rozdziału musimy sobie powiedzieć jasno: program komputerowy bazujący na deterministycznym generatorze liczb losowych może wygenerować tylko i wyłącznie liczby pseudolosowe, czyli takie, które tylko imitują prawdziwe liczby czysto losowe. Te ostatnie osiągalne są tylko procesie rzeczywistym. Możemy jednak za pomocą takich generatorów uzyskać ciąg liczb (bitów), który pod pewnymi względami będzie nierozróżnialny od ciągu uzyskanego z prawdziwie losowego źródła (np: z rzutu rzeczywistą kością).

¹ Generator liczb pseudolosowych można napisać nawet dla tak egzotycznych języków jak `Brainf*ck`.

4.1.2 Generatory liczb

Generator liczb losowych (RNG, z ang. random number generator) lub nieco bardziej ściśle *generator zdarzeń losowych* (REG, z ang. random event generator) to układ produkujący losowy ciąg elementów binarnych (bitów) najczęściej ułożony w postaci szeregu liczb losowych. Z punktu widzenia sposobu generowania liczb losowych wyróżniamy generatory sprzętowe (fizyczne, rzeczywiste) i programowe.

Generatory sprzętowe

TRNG (z ang. True RNG) - działające na zasadzie obrazowania właściwości i parametrów fizycznego procesu stochastycznego. Może to być ów rzut kością, monetą, wybieranie karty z talii kart itp. Wykorzystywać można też: efekt fotoelektryczny, szum termiczny, szum śrutowy, proces zaniku radioaktywnego...

Generatory programowe

PRNG, (z ang. Pseudo RNG) - działające na zasadzie deterministycznego obliczania ciągu liczb, które wyglądają jak liczby losowe. Algorytmy realizujące PRNG istnieją już ponad pół wieku i są obecnie zaimplementowane dla większości języków programowania. Na podstawie początkowej wartości nazywanej ziarnem czy zarodkiem (z ang. seed) oblicza kolejne wartości. Obie prezentowane funkcje Sage (`random` i `randint`) korzystają właśnie z jednego z takich algorytmów, zwanego *Mersenne Twister*. Jest to obecnie chyba najbardziej popularny algorytm opracowany w 1997 roku. Np. Matlab/GNU Octave też wykorzystuje ten algorytm. Jest on stosunkowo skomplikowany i może być trudny do realizacji, dlatego też omówimy sobie dużo prostszy, liniowy generator i omówimy jego zalety i (przede wszystkim) wady.

Programowe generowanie liczb losowych ² oparte jest na rekurencji

$$x_i = f(x_{i-1}, x_{i-2}, \dots, x_{i-k}),$$

czy w nieco bardziej zwartej formie

$$x_i = f(x_{i-1}).$$

Sekwencje te będą w oczywisty sposób deterministyczne. Problem polega na wygenerowaniu liczb których własności bardzo dobrze przypominają główne własności liczb prawdziwie losowych. Dodatkowo sekwencje liczb pseudolosowych będą powtarzały się co pewien okres, więc dość istotne jest aby generator takich liczb posiadał ów okres jak najdłuższy.

Liniowy generator kongruencyjny

LCG (linear congruential generator) wyznaczony jest przez metodę rekurencyjną

$$X_{n+1} = (aX_n + c) \mod m.$$

² Od tej chwili będziemy zawsze pisać *liczba losowa* a mieć na myśli *liczbę pseudolosową*, chyba, że napisane zostanie explicite, że mówimy o rzeczywistych liczbach losowych.

Stan początkowy to wartość ziarna (załadka). Nie jest on zbyt bezpieczny - istnieją techniki identyfikacji parametrów modelu na podstawie obserwacji wyników. Dla niektórych parametrów jest prawie losowy a dla innych dość szybko staje się okresowy. W powyższej definicji x_0 to ziarno (załadka), a mnożnik, c przesunięcie a $m \in \mathbb{Z}$ nazywamy modułem. Dwie liczby nazywamy kongruentnymi (przystającymi) modulo m jeżeli ich różnica jest podzielna przez m . Jeżeli $0 \leq a < m$ oraz $a \equiv b \pmod{m}$ wtedy a nazywamy resztą $b \pmod{m}$. Liczbę a można łatwo obliczyć z

$$a = b - \lfloor b/m \rfloor \times m$$

gdzie funkcja podłoga (z ang. floor) $\lfloor \cdot \rfloor$ oblicza największą liczbą całkowitą mniejszą od \cdot .

Jeżeli weźmiemy $c = 0$ dostaniemy multiplikatywny generator kongruencyjny. Jeżeli chodzi o moduł, to typowymi wartościami będą potęgi 2^k , a wartościami tych potęg będą typowe wielkości maszynowe dla przechowywania liczb całkowitych. Tak było przynajmniej dla wczesnych realizacji takiego generatora, co związane było z możliwością łatwej redukcji modulo poprzez wykorzystanie przepełnienia w stałopozycyjnej reprezentacji liczb w operacji mnożenia (w ciele liczb całkowitych) ax_i . W operacjach stałoprzecinkowych pierwszy bit reprezentuje znak, wobec czego w wyniku takiego mnożenia zamiast liczb z zakresu $[0, 2^{32} - 1)$ dostaniemy liczby z zakresu $[-2^{31} + 1, 2^{31} - 1]$. W ogólności wykonując operacje na liczbach większych od $2^{31} - 1$ jako wynik zachowujemy tylko bity niskiego rzędu.

Mnożnik a wybierać trzeba w taki sposób, aby LCG miał jak najdłuższy okres. Na 32-bitowych maszynach popularnymi wartościami początkowo były $m = 2^{32}$ i $a = 65539$. Jako, że dzisiejsze komputery są na tyle wydajne, by przeprowadzać redukcję modulo bardzo wydajnie, wiele ówczesnych implementacji generatora wykorzystuje operacje zmiennoprzecinkowe o zwiększonej precyzji. Inne wartości $a = 1099087573, 2396548189, 3934873077, 2304580733$ również produkują porządne sekwencje liczb losowych.

Innym dobrym wyborem dla m jest podstawienie dużej liczby pierwszej p . Wtedy okresem LCG będzie $p-1$ jeżeli tylko mnożnik ustawimy jako jego pierwiastek pierwotny. Szczególnie ważne wydają się być liczby pierwsze postaci $2^p - 1$, nazywane liczbami Mersenne'a. Na maszynach 32-bitowych popularnym wyborem bywa para $m = 2^{31} - 1$ i jej pierwiastek pierwotny $a = 7^5 = 16807$.

Implementacja LCG w Sage nie powinna nastroczać zbytnich problemów.

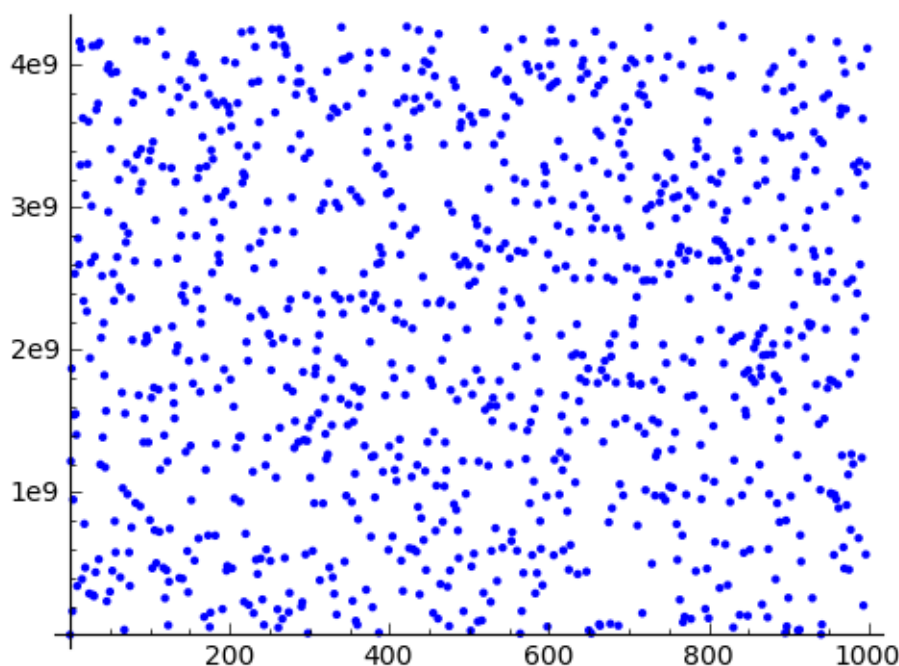
```
def myLCG(x, a=1664525, b=1013904223, m=2**32):
    return mod(a*x+b, m)
```

Możemy teraz wygenerować N liczb używając LCG i zmagazynować je w pythonowskiej liście.

```
def get_from_LCG(n=1, seed=123):
    ret = [seed]
    for i in xrange(n-1):
        ret.append(myLCG(ret[i]))
    return ret
```

```
lcg_list = get_from_LCG(N)
```

Powinniśmy dostać rysunek podobny do tego poniżej.



Rysunek 4.1: 1000 liczb losowych wygenerowanych generatorem liniowym LCG

Jak widać, program generuje liczby losowe z zakresu $[0, m)$.

W dalszej części wykładu (a raczej ćwiczeń) będziemy bazować na domyślnym generowaniu liczb losowych w Sage. Posłuży nam do tego wspomniana już funkcja `random()` zwracająca liczbę pseudolosową o rozkładzie jednorodnym na odcinku $[0, 1)$ (co często oznaczane jest poprzez $U(0, 1)$).

$$U(0, 1) = \begin{cases} 1 & 0 \leq x < 1 \\ 0 & \text{poza} \end{cases}$$

Aby uzyskać liczbę z przedziału $[0, 12.76)$ wystarczy po prostu pomnożyć liczbę zwracaną przez `random()` przez prawą granicę

```
random() * 12.76
```

a żeby uzyskać listę 123 liczb z przedziału $[-13.3, 33.1)$ należy wykonać

```
[random() * (33.1 + 13.3) - 13.3 for i in xrange(123)]
```

W ogólności do wygenerowania listy N liczb losowych z przedziału $[A, B)$ należy użyć polecenia

```
N = 100
A = -10
B = 20
[random() * (B - A) + A for i in xrange(N)]
```

Zadanie 2.2.2 Zmodyfikuj definicję `mylcg` tak, aby funkcja zwracała liczby losowe z przedziału $[0, 1)$.

Zadanie 2.2.3 LCG zdefiniowany tak jak powyżej produkuje stosunkowo dobre liczby losowe (prace naukowe nad tym stosunkowo prostym generatorem trwają do dzisiaj, dowodzone

są coraz to inne okresy bazujące na wyborze różnych zestawów parametrów a, c, m). Naszym zadaniem będzie natomiast zepsucie takiego generatora. Proszę znaleźć (numerycznie) 4 zestawy parametrów definiujących LCG takich, aby okres generatora był krótki. Wykreśl w Sage 4 rysunki LCG(N) (dla powiedzmy $N=1000$) dla owych parametrów. Powinieneś zauważyć regularność.

4.1.3 Generowanie liczb losowych o zadanym rozkładzie

Jako, że już dysponujemy generatorem liczb losowych o rozkładzie jednostajnym na odcinku $[0, 1) - U(0, 1)$ to możemy pokusić się o wygenerowanie liczb losowych o różnych innych rozkładach prawdopodobieństwa. Znałe jest kilka metod generowania takich liczb. Wszystkie przedstawione tutaj będą opierały się na tym, że umiemy generować liczby z rozkładem $U(0, 1)$. Szczególną uwagę poświęcimy generowaniu liczb z rozkładem $N(0, 1)$. Jest to standardowy zapis oznaczający rozkład normalny (Gausa) o średniej równej 0 i odchyleniu standardowym równym 1. Zanim omówimy pierwszą metodę, wcześniej zdefiniujemy sobie pojęcie *histogramu*. Będzie nam on potrzebny do wizualizacji rozkładów (czy raczej ich gęstości) z wygenerowanych liczb losowych.

Histogram

Wikipedia definiuje histogram jako jeden z graficznych sposobów przedstawiania rozkładu empirycznego cechy. Konstruuje się go jako szereg prostokątów odpowiadających liczebności elementów wpadających do określonego przedziału klasowego. Szerokości przedziałów klasowych mogą mieć stałe lub zmienne długości. W bardziej matematycznym sensie histogram to funkcja zliczająca ilość obserwacji pasujących do oddzielnych przedziałów klasowych. Jeżeli n oznacza liczbę wszystkich obserwacji, a k to liczba przedziałów, wtedy histogram m_i spełnia następujący warunek

$$n = \sum_{i=1}^k m_i$$

Ideę histogramu najlepiej zrozumieć na przykładzie. Mamy następującą listę liczb

1 = [1, -3, -5, -1, -3, 1, 5, 1, 3, -3, 4, 2, 4, -1, 4, 5, -2, 4, 3, -4]

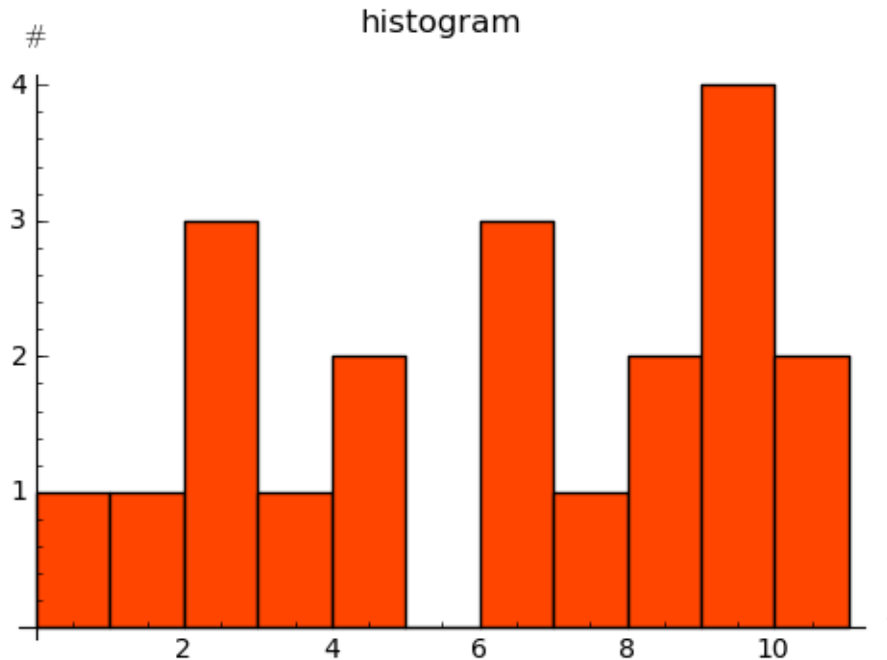
Budując histogram na początku musimy ustalić szerokość przedziału. Zaczniemy od łatwiejszej wersji: niech szerokość będzie stała. Najlepiej podzielić ową listę na przedziały zawierające liczby całkowite. W zasadzie wystarczy zliczać ile jest poszczególnych liczb całkowitych w liście 1. Zróbmy to. Widzimy, że mamy

-5	-4	-3	-2	-1	0	1	2	3	4	5
1	1	3	1	2	0	3	1	2	4	2

W zasadzie mamy już nasz histogram. Jeżeli posumujemy ilość elementów listy ($\text{len}(1)$), oraz obliczymy n zobaczymy, że dostaniemy tą samą liczbę (=20). Pozostaje narysować ów histogram. Na odciętej musimy odłożyć przedziały klasowe a na rzędnej liczebności danego przedziału. Przyjęło się rysować histogram używając słupków. Sage na chwilę obecną posiada kilka metod narysowania takiego histogramu. Jeżeli nie zależy nam na poprawnym opisanu odciętej (np: chcemy tylko zobaczyć kształt histogramu), wystarczy napisać

```
h = [l.count(i) for i in range(-5,6)]
b = bar_chart(h, width=1, color="orangered")
b.show(axes_labels=['$i$', '$\#$'], title="histogram")
```

Co pozwoli nam wygenerować taki rysunek:



Rysunek 4.2: Prosty wykres liczebności, gdzie wykorzystaliśmy funkcję `bar_chart()`.

Nie jest to prawdziwy histogram, bowiem odłożone na osi OY liczebności powinny odpowiadać rzeczywistym wartościom (przedziałom). Możemy skorzystać z pakietu `Time Series` dostępnego w Sage. Wystarczy prosta komenda aby uzyskać dostęp do wielu statystycznych funkcji typowych dla analizy szeregu czasowego.

```
v = finance.TimeSeries(l)
```

I teraz aby obliczyć histogram dla 10 równych przedziałów (od minimalnej do maksymalnej wartości występującej w liście `l`), wystarczy napisać

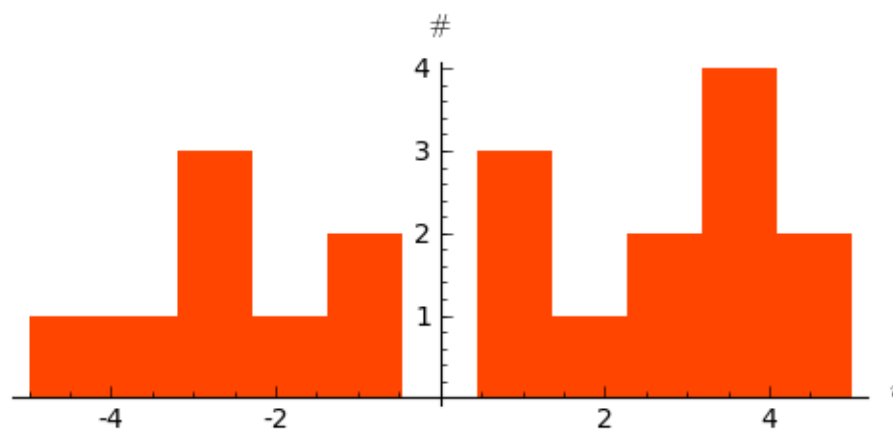
```
v.histogram(bins=11)
```

a aby narysować jego wykres

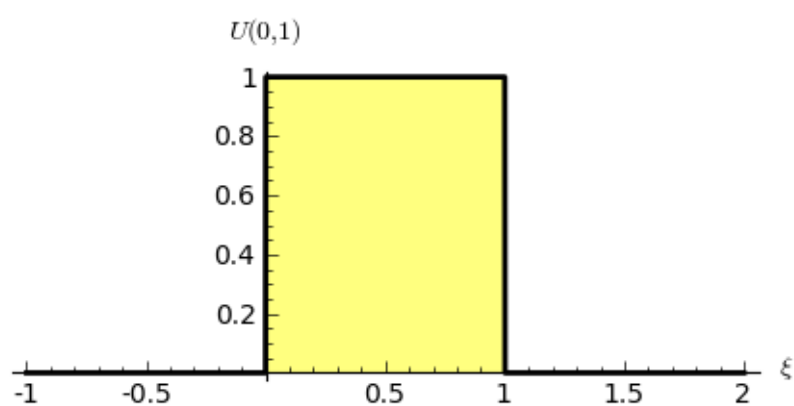
```
v.plot_histogram(bins=11, normalize=0, axes_labels=['$i$', '$\#$'])
```

Oczywiście całą procedurę można powtórzyć dla liczb zmiennoprzecinkowych (rzeczywistych, wymiernych). W tym wypadku należałoby oczywiście policzyć ile posiadanych liczb wpada do zdefiniowanych “pudełek”. Zobaczmy drugi przykład, gdzie obliczymy i narysujemy w Sage histogram dla dziesięciu tysięcy liczb z $U(0,1)$. Powinniśmy dostać

Przykład 2 Wygenerujemy 10000 liczb losowych a następnie dla nich obliczymy i narysujemy histogram.



Rysunek 4.3: Histogram dla listy 1 uzyskany z wykorzystaniem pakietu TimeSeries



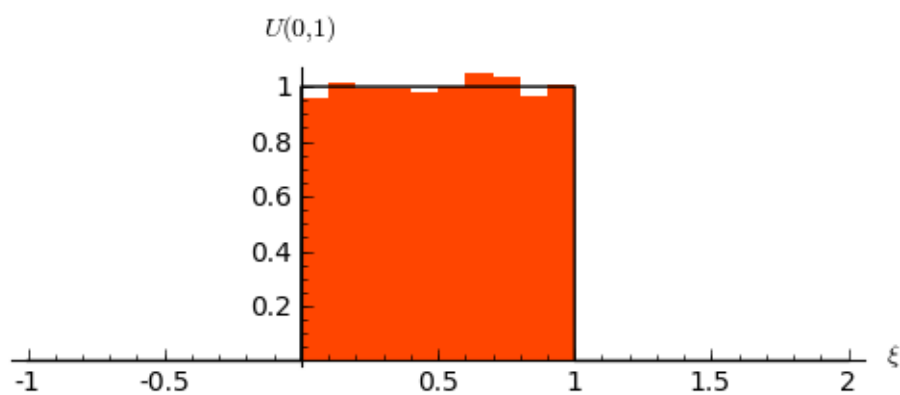
Rysunek 4.4: Wykres rozkładu $U(0,1)$


```

N = 10000
u01 = [random() for i in xrange(N)]
fu01 = lambda x: 0 if x < 0 or x > 1 else 1
v = finance.TimeSeries([random() for i in xrange(N)])
plot1 = plot(fu01, (-1,2), thickness=1, color="black")
plot2 = v.plot_histogram(bins=10, color="orangered")
(plot1 + plot2).show(axes_labels=[r'$\xi$', r'$U(0,1)$'])

```

Ostatnia linia wyrysuje nam obie funkcje na jednym wykresie. Zachęcamy czytelnika do poeksperymentowania z powyższym kodem - można zmienić liczbę prób N i łatwo zobaczyć jak histogram zaczyna oddalać się od teoretycznego rozkładu dla małych N i jak zbliża się dla dużych. Można też zobaczyć jak ilość przedziałów (parametr `bins`) wpływa na otrzymany histogram.



Rysunek 4.5: Wykres rozkładu $U(0,1)$ + histogram miliona prób.

Metoda inwersyjna

Każdy rozkład prawdopodobieństwa może być jednoznacznie scharakteryzowany poprzez pewną funkcję rzeczywistą zwaną **dystrybuantą**.

Dystrybuanta Niech \mathbb{P} będzie rozkładem prawdopodobieństwa. Funkcję $\mathbb{F} : \mathbb{R} \rightarrow \mathbb{R}$ daną wzorem

$$\mathbb{F}(\xi) = \mathbb{P}((-\infty, \xi))$$

nazywamy dystrybuantą rozkładu \mathbb{P} .

W metodzie inwersyjnej żądany rozkład o dystrybuancie \mathbb{F} uzyskuje się poprzez przekształcenie zmiennej losowej o rozkładzie $U(0,1)$ za pomocą funkcji odwrotnej do \mathbb{F} .

Twierdzenie Załóżmy, że dystrybuanta \mathbb{F} jest ściśle rosnąca. Jeśli zmienna losowa u ma rozkład $U(0,1)$ to $\mathbb{F}^{-1}(u)$ ma dystrybuantę \mathbb{F} .

Dowód TBA

Algorytm wykorzystujący powyższe twierdzenie jest bardzo prosty i wygląda następująco:

1. Generujemy liczbę $u \in U(0,1)$.

2. Przekształcamy u stosując

$$x = \mathbb{F}^{-1}(u)$$

Wynikowa liczba losowa x posiada żądany rozkład \mathbb{P} .

Oczywiście skuteczność tej metody zależy bezpośrednio od tego czy możemy łatwo obliczyć \mathbb{F}^{-1} . Jeżeli tak - jest to najprostsza znana metoda generowania liczb losowych z danym rozkładem. Do rozkładów, do których można zastosować tą metodę należą wszystkie rozkłady, których dystrybuenta znana jest jawnie *oraz* można ją łatwo odwrócić. O takich rozkładach powiemy sobie niżej.

Rozkład wykładniczy

Przejdźmy wreszcie do generowania liczb losowych z rozkładem innym niż $U(0,1)$. Na początek weźmy jeden z najbardziej powszechnych, czy popularnych rozkładów prawdopodobieństwa - **rozkład wykładniczy**. Gęstość takiego rozkładu dana jest wzorem

$$f(\xi) = \lambda e^{-\lambda \xi}$$

Jak łatwo policzyć, dystrybuenta $F(x)$ wynosi

$$F(x) = \int_{-\infty}^x f(\xi) d\xi = -e^{-\lambda \xi} \Big|_{-\infty}^x = -e^{-\lambda x} + 1,$$

a jej odwrotność

$$F^{-1}(u) = -\frac{1}{\lambda} \ln(1 - u).$$

Spróbujmy wygenerować 5000 liczb o rozkładzie wykładniczym. Następnie obliczymy sobie histogram, unormujemy go i porównamy z teoretycznym rozkładem dla kilku wartości $\lambda = 0.5, 1, 1.5$.

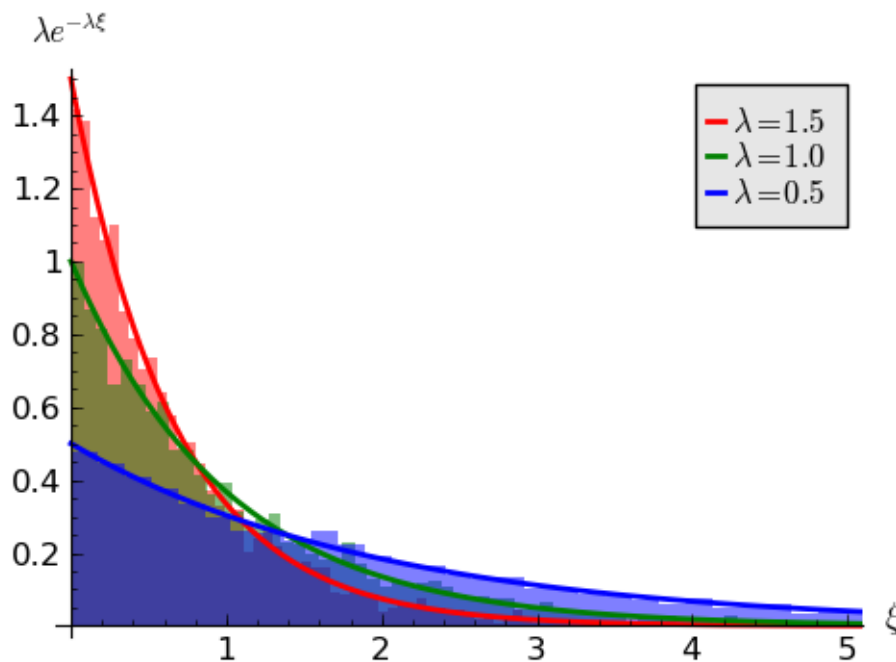
```
f(xi, a) = a * exp(-a * xi)
F(u, a) = -log(1-u)/a
N = 5000
kolor = ["red", "green", "blue"]
parlist = [1.5, 1, 0.5]
p = []
i = 0
for par in parlist:
    lista = [F(random(), par) for j in xrange(N)]
    v = finance.TimeSeries(lista)
    P = v.plot_histogram(bins=100, color=kolor[i], alpha=0.5)
    P.set_aspect_ratio("automatic")
    p.append(P)
    p.append(plot(f(xi,par), 0, max(lista),
        thickness=2, color=kolor[i],
        legend_label=r'\lambda = %.1f'%par))
    i += 1
```

```

p1 = sum(p)
p1.axes_labels(['r'$\xi$', 'r'$\lambda e^{-\lambda \xi}$'])
p1.show(xmin=0, xmax=5, figsize=5, fontsize=12)

```

Jak widać na rysunku liczby losowe przekształcone metodą inwersji w oparciu o odwrotność dystrybuanty, dość dobrze odwzorowują rozkład wykładniczy. Lepszy wynik można oczywiście uzyskać zwiększając parametry N oraz bins .



Rysunek 4.6: Wykres gęstości rozkładu wykładniczego oraz histogram z 5000 prób liczb losowych dla trzech wartości parametru λ .

Alternatywnie można wykorzystać przekształcenie bazujące na spostrzeżeniu, że liczby $1 - u$ oraz u ($u \in U(0, 1)$) posiadają ten sam rozkład $U(0, 1)$.

Rozkład Cauchy'ego

Rozkład ten dany jest wzorem

$$f(\xi) = \frac{\sigma}{\pi(\xi^2 + \sigma^2)}$$

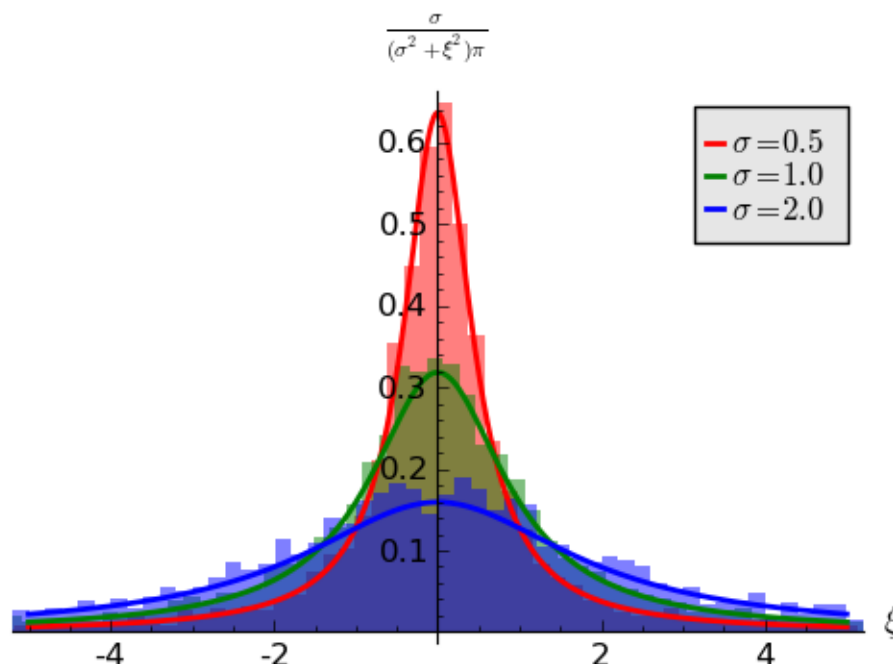
Odwrotność dystrybuanty powyższego rozkładu wynosi

$$F^{-1}(u) = \sigma \tan \left[\pi \left(u - \frac{1}{2} \right) \right].$$

Stosując proste i bardzo naturalne przekształcenie oryginalnej zmiennej $v = 1/2 - u$ dostajemy nieco prostsze wyrażenie

$$F^{-1}(v) = \sigma \tan(\pi v).$$

Stosując podobne metody jak w poprzednim rozdziale możemy sprawdzić, czy powyższe przekształcenie generuje liczby z odpowiednim rozkładem.



Rysunek 4.7: Wykres gęstości rozkładu Cauchy'ego oraz histogram z 5000 prób liczb losowych dla trzech wartości parametru $\sigma = 0.5, 1, 2$.

Rozkład logistyczny

Rozkład ten dany jest wzorem

$$f(\xi) = \frac{1}{2 + e^{\xi} + e^{-\xi}}$$

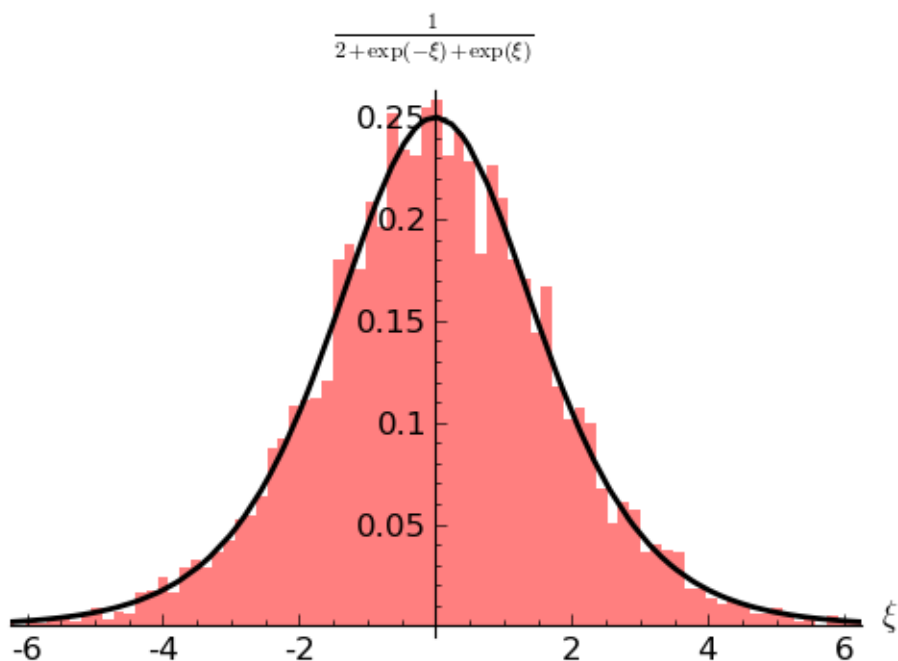
Odwrotność dystrybuanty powyższego rozkładu wynosi

$$F^{-1}(u) = \ln \frac{u}{1-u}.$$

Stosując podobne metody jak w poprzednim rozdziale możemy sprawdzić, czy powyższe przekształcenie generuje liczby z odpowiednim rozkładem.

Zadanie 2.3.1 Wygeneruj 200000 liczb losowych z rozkładem Pareto. Narysuj ich unormowany histogram oraz funkcję gęstości. Porównaj obie funkcje zmieniając wszystkie parametry rozkładu.

Zadanie 2.3.2. Wygeneruj 1000 liczb losowych z rozkładem trójkątnym. Narysuj ich unormowany histogram oraz funkcję gęstości. Porównaj obie funkcje zmieniając wszystkie parametry rozkładu.



Rysunek 4.8: Wykres gęstości rozkładu logistycznego oraz histogram z 200000 prób liczb losowych.

4.2 Równania różniczkowe zwyczajne

4.2.1 Metoda Eulera

Jest to bodaj najprostszy sposób na numeryczne rozwiązywanie równań różniczkowych zwyczajnych. Technicznie to metoda pierwszego rzędu. Bazuje na prostej interpretacji definicji pochodnej. Rozpatrujemy równanie postaci

$$\frac{dy}{ds} = y' = f(y, s),$$

z zadanymi warunkami początkowymi $(sx(0), y(0)) = (x_0, y_0)$. Stosując przekształcenie

$$\frac{dy}{ds} = \lim_{\Delta s \rightarrow 0} \frac{\Delta y}{\Delta s}$$

Na chwilę zapomnijmy o tej granicy. Mamy

$$\frac{\Delta y}{\Delta s} = \frac{y(s + \Delta s) - y(s)}{\Delta s}$$

stawiając krok “czasowy” $\Delta x = h$ (jak to zwyczajowo w symulacjach komputerowych) dostajemy

$$\begin{aligned} \frac{y(s + h) - y(s)}{h} &\simeq f(y, s), \\ y(s + h) &\simeq y(s) + f(y, s)h. \end{aligned}$$

Oczywiście pełna równość zachodzi tylko w granicy $h \rightarrow 0$. Na potrzeby numeryczne jednak nie musimy się tym przejmować. Należy mieć tylko świadomość, że zmniejszanie kroku czasowego zbliża nas do wyniku dokładnego (zazwyczaj). Jako, że w iteracjach “czasowych” wykorzystujemy stały krok h , więc w danej, powiedzmy i -tej iteracji czas rzeczywisty zastąpiony zostanie poprzez $s_i = i \cdot h$, możemy równie dobrze wprowadzić indeksowanie nie po czasie s ale po zmiennej iteracyjnej i

$$\begin{aligned}y(s = ih) &= y(s_i) = y_i, \quad s_i = ih, \\y(ih + h) &= y(h(i + 1)) = y(ih) + f(y(ih), ih)h, \\y_{i+1} &= y_i + f(y_i, s_i)h.\end{aligned}$$

Dostając tzw. schemat Eulera. Innymi metodami wyprowadzenia tego prostego schematu będzie użycie rozwinięcia Taylora lub wyciąganie równania różniczkowego od s_0 do $s_0 + h$ za pomocą metody prostokątów (używając tylko jednego prostokąta na całym przedziale całkowania).

Zadanie D1.1 Uzyskaj schemat Eulera z rozwinięcia Taylora.

Zadanie D1.2 Uzyskaj schemat Eulera całkując równanie różniczkowe.

Dokładność metody Eulera mocno zależy od wyboru kroku całkowania h . Musimy też na początku zadać warunki startowe (początkowe) iteracji, zupełnie jak podczas rozwiązywania równań metodami analitycznymi. Dla przykładu obliczmy numerycznie ruch przetłumionego sprężyny o współczynniku sprężystości k .

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

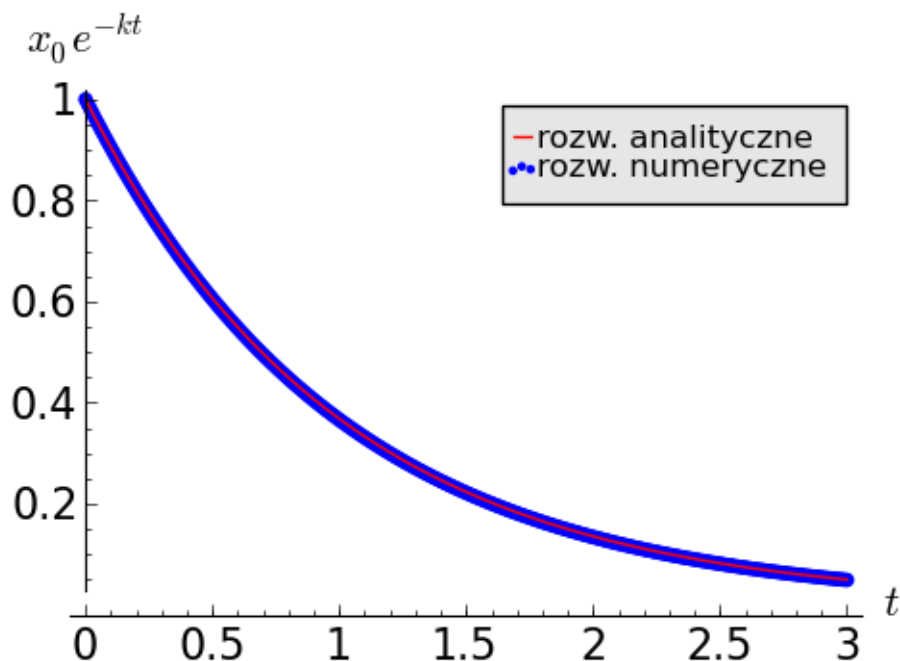
Rozwiązaniem jest zanik eksponencjalny $x(t) = x_0 \exp(-kt)$. Wybierzemy sobie warunek początkowy $x_0 = 1$ i współczynnik sprężystości $k = 0.1$. Policzmy N iteracji i narysujmy wykres. Na początku napiszemy dyskretną wersję równania ruchu

$$x_{i+1} = x_i - kx_ih = x_i(1 - kh).$$

Teraz kod Sage

```
# liczba iteracji rownania
N = 300
# czas początkowy
t = 0
# skok czasowy
h = 0.01
# parametr rownania (sprezystosc)
k = 1
# wartosc początkowa x(t=0)
x_0 = 1
#inicjalizacja listy poprzez wartosc początkowa
lista_x = [x_0]
for i in xrange(N):
    lista_x.append(lista_x[i] * (1 - k*h))
```

wynik (rozwiązanie numeryczne równania) trzymany jest w liście `lista_x`. Wyrzysujmy ją standardowo, razem z wynikiem analitycznym.



Rysunek 4.9: Rozwiązanie analityczne i numeryczne równania $\dot{x} = -kx$.

Pomimo użycia prostej metody (pierwszego rzędu), wykresy wyglądają identycznie. No, ale czy na pewno jest tak pięknie? Poprawność metody możemy łatwo zbadać obliczając błędy względny (E_w) i bezwzględny (E_b).

$$E_b = \bar{y} - y,$$

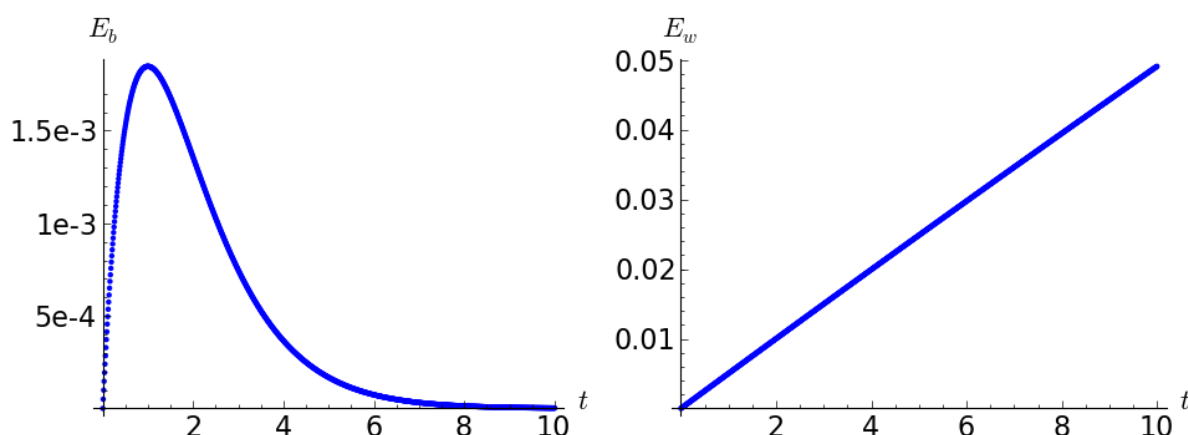
$$E_w = \frac{\bar{y} - y}{\bar{y}} = \frac{E_b}{\bar{y}}.$$

gdzie y to wielkość obliczona algorytmem a \bar{y} to dokładna wielkość analityczna. Dla jasności - nie interesują nas w tym przypadku znaki błędów a jedynie ich wartość bezwzględna (tu proszę zwrócić uwagę na nomenklaturę, żeby nie pomylić wartości bezwzględnej z błędem bezwzględnym). Dlatego najczęściej oblicza się nie E_w a $|E_w|$. Wymaga to drobnej korekty powyższego kodu

```
N, t, h, k, x_0 = 1000, 0, 0.01, 1, 1
g(s) = x_0*exp(-k*s)
lista_x = [x_0]
Eb = [g(0) - x_0]
Ew = [Ew[0]/g(0)]
for i in xrange(1,N):
    lista_x.append(lista_x[i-1] * (1 - k*h))
    Eb.append(abs(g(i*h) - lista_x[i]))
    Ew.append(Eb[i]/g(i*h))
list_plot(zip([i*h for i in xrange(N+1)],Eb)).show()
list_plot(zip([i*h for i in xrange(N+1)],Ew)).show()
```

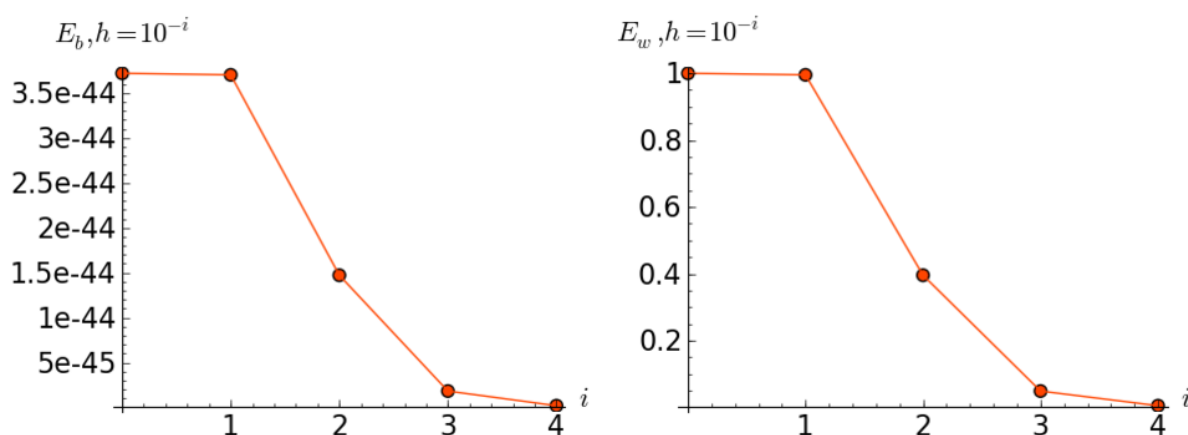
Spójrzmy. Na pierwszym wykresie odłożony mamy błąd bezwzględny. Widzimy, że dla krótkich czasów odbiega on od wartości analitycznej by dla większych czasów zmaleć do zera. Mogą być tego 2 powody: (i) różnica pomiędzy obiema wartościami maleje zo zera lub (ii)

obie wartości maleją do zera, więc ich różnica też. Jako, że funkcja jest eksponencjalna, dużo bardziej prawdopodobny jest ten drugi scenariusz. Aby zobaczyć, czy błąd rośnie z ilością iteracji (w czasie) wykreślimy błąd względny. Mówi on nam o stosunku błędu bezwzględnego do wartości analitycznej (rysunek po prawej). Tu jak widać rośnie on wraz z czasem, z czego możemy wywnioskować, że wraz z ilością iteracji coraz mniej dokładnie obliczamy wartość y .



Rysunek 4.10: Błąd bezwzględny (po lewej) i błąd względny (po prawej).

Najprostsza metoda poprawienia jakości rozwiązań jest zmniejszenie kroku całkowania. Zależności pozostaną podobne, zmniejszy się jednak wartość błędów w danej chwili czasowej.



Rysunek 4.11: Błąd bezwzględny (po lewej) i błąd względny (po prawej) po czasie $t = 10$ dla różnych kroków czasowych $h = 1, 0.1, 0.01, 0.001, 0.0001$.

W tabeli zawarto wartości błędów bezwzględnego i względnego dla różnych wielkości kroku czasowego symulacji, po osiągnięciu czasu końcowego $T_E = 10$. Widać, że pomimo, że za każdym zmniejszeniem kroku zwiększała się ilość kroków czasowych, dokładność obliczeń rosła - malał zarówno błąd bezwzględny jak i względny.

h	N	E_b	E_w
1	100	3.7210^{-44}	1
0.1	1 000	3.7010^{-44}	0.995
0.01	10 000	1.4710^{-44}	0.396
0.001	100 000	1.8210^{-45}	0.0488
0.0001	1 000 000	1.8610^{-46}	0.0499

Przejdźmy teraz do rozwiązywania równania różniczkowego wyższego stopnia. Znów posłużymy się przykładem oscylatora harmonicznego. Tym razem rozwiążemy równanie Newtona dla punktowej cząstki o masie m w potencjale $U(x) = -kx^2/2$. Pomińmy siły tarcia. Również w spokoju zostawimy wymuszenie.

$$m\ddot{x}(t) = -kx(t).$$

Równanie to posiada znane **analityczne rozwiązanie**. Oznaczając $\omega_0^2 = k/m$ dostajemy

$$x(t) = A \sin(\omega_0 t) + B \cos(\omega_0 t).$$

Stałe A , B (amplitudy) zależne są od wyboru warunków początkowych. Spróbujmy numerycznie rozwiązać równanie ruchu tak, aby pokazać zgodność z rozwiązaniem. Aby napisać schemat Eulera dla równania drugiego stopnia najpierw trzeba przepisać równanie do układu równań na x i $v = \dot{x}$.

$$\begin{aligned}\dot{x}(t) &= v(t), \\ \dot{v}(t) &= -\frac{k}{m}x(t) = -\omega_0^2 x(t).\end{aligned}$$

Teraz wystarczy zdyskretyzować te równania, tak samo jak robiliśmy to z równaniem pierwszego rzędu.

$$\begin{aligned}x_{i+1} &= x_i + v_i h, \\ v_{i+1} &= v_i - \omega_0^2 x_i h.\end{aligned}$$

Po ustaleniu warunków początkowych $x(0) = x_0$ oraz $v(0) = v_0$ możemy rozpocząć normalną procedurę symulacji - wybieramy krok czasowy h , ustalamy parametry równania i do dzieła.

```
h = 0.01 # skok
N = 100 # liczba krokow

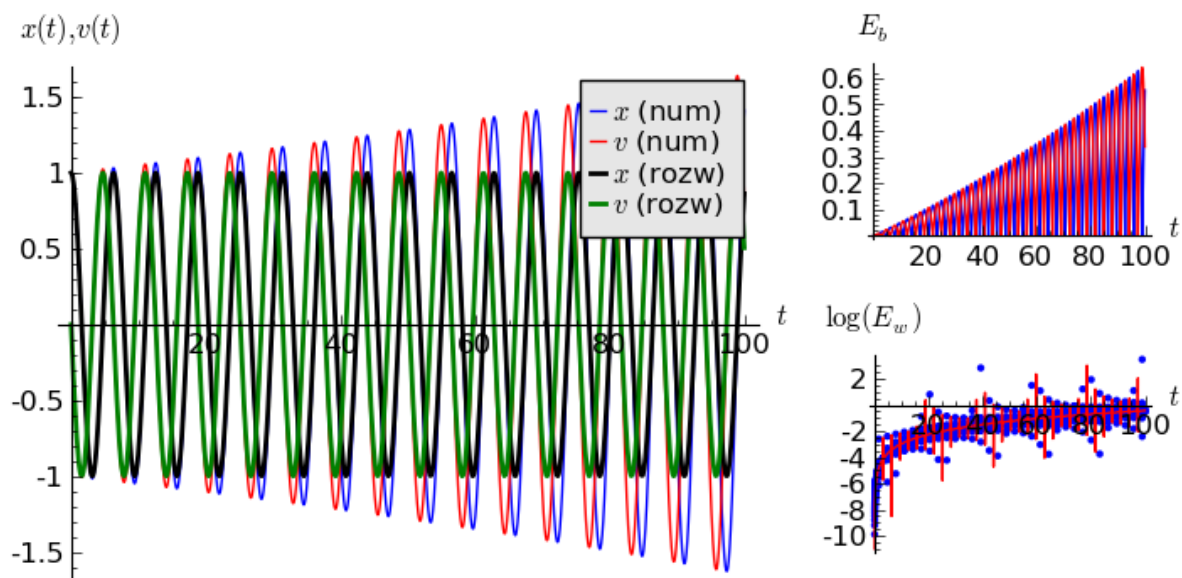
x0 = 1
v0 = 0
omega0 = 1

lista_x = [x0]
lista_v = [v0]
lista_t = [0]
for i in xrange(N):
    lista_x.append(lista_x[i] + lista_v[i] * h)
    lista_v.append(lista_v[i] - omega0^2 * lista_x[i] * h)
    lista_t.append(lista_t[i] + h)
```

Wykorzystamy też Sage do obliczenia rozwiązania analitycznego dla naszego zagadnienia.

```
var('t x omega x_0 v_0')
x = function('x', t)
assume(omega>0)
eq = diff(x, t, 2) + omega^2 * x == 0
solx = desolve(eq, x, ivar=t, ics=[0, x_0, v_0])
solv = diff(solx,t)
```

Teraz możemy zobaczyć jak dokładna jest metoda Eulera w przypadku równań wyższych rzędów. Poniżej znajdziecie wykres rozwiązań dla $h=0.01$ i 10000 kroków.



Rysunek 4.12: Porównanie numerycznego ($x(t)$ linia niebieska, $v(t)$ linia czerwona) i analitycznego ($x(t)$ linia czarna, $v(t)$ linia zielona) rozwiązania zagadnienia oscylatora harmonicznego. Jak widać odchylenia od rozwiązań dokładnych zaczynają być znaczące już dla kilku kroków symulacji. Błąd bezwzględny widnieje na prawym górnym panelu; błąd względny wykreślony jest na prawym dolnym panelu w skali logarytmicznej dla lepszej czytelności. Parametry użyte dla powyższej symulacji: $x_0 = 1, v_0 = 0, \omega_0 = 1, h = 0.01, N = 10000$.

Inną, aczkolwiek trudniejszą metodą będzie użycie algorytmów wyższego rzędu takich jak schemat Rungego-Kutty (2-giego, 4-tego i wyższych rzędów). Dociekliwy student może zajrzeć [tutaj](#).

Zadanie D1.3 Przeprowadź podobną symulację dla innych wartości h . Wykreśl zależność błędów względnego i bezwzględnego w funkcji wartości h . Błędy badaj po rzeczywistym czasie symulacji $T_E = 100$.

Zadanie D1.4 Rozwiązać numerycznie równania

1. $\dot{x}(t) = -kx^3$
2. $\dot{x}(t) = F$
3. $m\ddot{x}(t) = mg$
4. $\ddot{r}(t) = 4\epsilon \left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^6 \right], r > 0$

$$5. \quad m\ddot{x}(t) = -\gamma\dot{x}(t) - kx(t)$$

$$6. \quad m\ddot{x}(t) = -kx(t) - \gamma\dot{x}(t) + a \cos(\omega t)$$

Dla każdego przypadku należy

- narysować x (dla D-F również v) w funkcji t (opisać osie),
- **odpowiedzieć na pytanie: z jakim ruchem mamy do czynienia [dla jakich parametrów równanie jest cykliczny (periodyczny), dla jakich rozwiązanie jest stałe (niezmienne w czasie)...],**
- znaleźć błąd względny i bezwzględny, wykreślić w funkcji czasu.

Pytania:

1. Czym różni się przypadek 5 od 6?
2. Co opisuje potencjał w 4? Jakie ma zastosowanie w fizyce?

4.3 Stochastyczne równania różniczkowe

Stochastyczne równania różniczkowe (SDE, od angielskiej nazwy Stochastic differential equations) są obecnie uważane za standardowe narzędzie wykorzystywane do analizy niektórych wielkości opisujących dynamikę rynków finansowych. Do tych wielkości należą ceny aktywów, stopy procentowe czy ich pochodne. W przeciwieństwie do zwyczajnych równań różniczkowych, które posiadają jednoznaczne rozwiązanie, rozwiązaniami SDE są ciągłe w czasie procesy stochastyczne. Metody komputerowe wykorzystywane do analizy SDE bazują na klasycznych metodach wykorzystywanych do rozwiązywania tradycyjnych, deterministycznych równań różniczkowych, są jednak uogólnione tak, aby radzić sobie z procesami losowymi.

Zestaw zmiennych losowych X_t indeksowanych liczbami rzeczywistymi t nazywamy procesem losowym ciągłym (ze względu na czas). Każda *realizacja* procesu losowego to przypadkowa wartość zmiennej losowej X_t dla każdego t , jest więc funkcją czasu. Co ciekawe, *każda* deterministyczna funkcja $f(t)$ może być uważana za proces stochastyczny, którego wariancja znika.

4.3.1 Proces Wienera

Najbardziej znanym przykładem procesu losowego szeroko występującego w modelach fizyki, chemii ale i rynków finansowych jest *proces Wienera* $W(t) = W_t$, ciągły proces stochastyczny posiadający następujące własności

1. jest to proces rzeczywisty,
2. startuje z zera ($W_0 = 0$),
3. ma stacjonarne i niezależne przyrosty na nieprzekrywających się przedziałach,
4. jest procesem Gaussa o zerowej wartości średniej $\langle W_t - W_s \rangle = 0$ i wariancji przyrostów $\langle [W_t - W_s]^2 \rangle = 2D(t - s)$,

5. proces Wienera może być reprezentowany ciągłymi trajektoriami.

Wynika z tego, że dla każdej różnicy czasów $t - s$ zmienna losowa $W_t - W_s$ jest zmienną losową gaussowską o zerowej wartości średniej i wariancji $2D(t - s)$. Więc jego rozkład prawdopodobieństwa ma postać

$$f_{W_t - W_s}(x) = \frac{1}{\sqrt{2\pi D(t - s)}} \exp \left[-\frac{x^2}{4D(t - s)} \right]. \quad (4.1)$$

Proces taki może być wyprowadzony jako proces graniczny błędzenia przypadkowego. Wystarczy tylko zbadać granicę dla której wielkość skoku i czas pomiędzy skokami będą maleć do zera. Tak zdefiniowanym procesem posługujemy się zwyczajowo, gdy podczas analizy problemu pojawia się jakaś nieregularna siła czy zaburzenie którego nie możemy opisać równaniami deterministycznymi.

Typowe dla rynków finansowych *równanie dyfuzji* może być modelowane przez równanie różniczkowe posiadające część deterministyczną zwaną **dryftem** oraz część losową zwaną **dyfuzją**. Ta ostatnia jest bardzo często reprezentowana właśnie przez proces Wienera. Możemy sobie napisać ogólne równanie

$$dX = a(t, X)dt + b(t, X)dW_t. \quad (4.2)$$

Jest to postać różniczkowa. W zwykłych równaniach różniczkowych zazwyczaj stosujemy pochodne dx/dt . W tym przypadku postać różniczkowa ma większy sens, jako, że wiele interesujących nas procesów losowych (jak ruch Browna) są procesami ciągłymi aczkolwiek nie są różniczkowalne. Powyższe równanie nabiera większego sensu pod znakiem całki

$$X(t) = X(0) + \int_0^t a(s, y)ds + \int_0^t b(s, y)dW_s. \quad (4.3)$$

Ostatni wyraz z prawej zwany jest całką Ito.

Równanie Blacka-Scholesa

Jednym z bardziej znanych, historycznym już równaniem stochastycznym, jest równanie opisujące geometryczny ruch Browna

$$dX = \mu X dt + \sigma X dW_t, \quad (4.4)$$

gdzie μ, σ to wielkości stałe. Równanie to jest jednym z podstawowych elementów modelu wyceny opcji Blacka-Scholesa. Teoria ta została nagrodzona Nagrodą Nobla z ekonomii w roku 1997, a opracowana przez absolwenta fizyki i doktora matematyki Fischera Blacka oraz ekonomistę Myrona Scholesa. Teoria Blacka-Scholesa pozwala na wycenę wartości tzw. finansowych instrumentów pochodnych, czyli opcji, oraz służy do optymalizacji “bezpiecznego” portfela inwestycyjnego.

Pomimo tego, że równanie to wygląda na proste, żeby nie powiedzieć zbyt proste na to by opisywać jakąkolwiek rzeczywistość na rynkach finansowych, ma ono olbrzymie znaczenie, jako, że może być rozwiązane ściśle dając wynikowy wzór z którego możemy wyliczyć zmianę cen prostych opcji. Jak już powiedzieliśmy, rozwiązanie jest geometrycznym ruchem Browna

$$X(t) = X_0 \exp \left[\left(\mu - \frac{\sigma^2}{2} \right) t + \sigma dW_t \right]. \quad (4.5)$$

Dzięki zamkniętej postaci rozwiązania mamy możliwość testowania omówionych poniżej metod numerycznych.

Schemat Eulera-Maruyamy

Najprostszą metodą numerycznego rozwiązywania równań różniczkowych zwyczajnych jest metoda Eulera. Bazuje ona np. na rozwinięciu Taylora w przybliżeniu pierwszego rzędu. Stochastycznym analogiem tej metody jest metoda Eulera-Maruyamy.

Będziemy chcieli podać przybliżone rozwiązanie ogólnej postaci SDE na przedziale czasowym $t \in [t_0, t_E]$. Na początku zdyskretyzujemy sobie ów przedział czasowy, ustalając na siatkę N punktów

$$t_0 < t_1 < t_2 < \dots < t_{N-2} < t_E. \quad (4.6)$$

Dążymy do tego, aby na tej siatce znaleźć przybliżone wartości zmiennej X . Oznaczmy je

$$w_0 < w_1 < w_2 < \dots < w_{N-2} < w_E. \quad (4.7)$$

Są to oczywiście przybliżone rozwiązania zmiennej x dla odpowiednich czasów z powyższej siatki $\{t_i\}$. Zakładając wartość początkową dla SDE $X(t_0) = X_0$ możemy pokusić się o rozwiązanie numeryczne w następującej postaci

$$\begin{aligned} w_0 &= X_0 \\ w_{i+1} &= w_i + a(t_i, w_i)\Delta t_{i+1} + b(t_i, w_i)\Delta W_{i+1} \\ \Delta t_{i+1} &= t_{i+1} - t_i \\ \Delta W_{i+1} &= W(t_{i+1}) - W(t_i). \end{aligned} \quad (4.8)$$

Kluczową sprawą w tym punkcie jest problem: jak modelować ΔW_i ? Mając do dyspozycji generator liczb losowych z rozkładem $N(0, I)$ każdą losową liczbę ΔW_i obliczamy ze wzoru

$$\Delta W_i = \sqrt{\Delta t_i} z_i, \quad (4.9)$$

gdzie z_i jest losowana właśnie z $N(0, I)$. Aby scałkować proces stochastyczny użyjemy formuły na przyrost procesu Wienera

$$\int_{t_i}^{t_{i+1}} \Gamma(t) dt = \int_{t_i}^{t_{i+1}} dW(t) = W(t_{i+1}) - W(t_i) \quad (4.10)$$

Z definicji procesu Wienera wiemy, że jest on procesem Gaussa o zerowej średniej i wariancji liniowej w czasie $\langle [W(t_{i+1}) - W(t_i)]^2 \rangle = 2D\Delta t_i$, co daje nam w sensie średnio-kwadratowym $\Delta W \propto \sqrt{\langle [\Delta W(t)]^2 \rangle} = \sqrt{\Delta t_i}$. Scałkowanie procesu Wienera prowadzi do

$$\int_{t_i}^{t_{i+1}} dW_t = \sqrt{\Delta t_i} z_i. \quad (4.11)$$

Jeżeli założymy sobie, że krok czasowy (odległości na siatce rozwiązań) jest stały i wynosi $\Delta t_i = h$ możemy napisać schemat jawnie

$$w_{i+1} = w_i + ha(t_i, w_i) + \sqrt{h}b(t_i, w_i)z_i. \quad (4.12)$$

Jako, że każdy zestaw wartości $\{w_i\}, i = 0, \dots, E$ wyprodukowany przez powyższą formułę będzie przybliżonym rozwiązaniem procesu losowego, to i każda realizacja (każdy zestaw) będzie również losowa, a co za tym idzie - każda realizacja procesu będzie inna.

Schemat Millsteina

Dodaje on poprawkę do poprzedniego rozwiązania, powodując, że schemat staje się schematem pierwszego rzędu w sensie silnym. Dany jest on wzorem iteracyjnym

$$w_0 = X_0 \quad (4.13)$$

$$w_{i+1} = w_i + a(w_i, t_i)h - \frac{h}{2}b(x_i, t_i)\frac{\partial b'}{\partial x}(w_i, t_i)(z_i^2 - 1) + \sqrt{h}b(w_i, t_i)z_i.$$

Obie metody (Millsteina i Eulera-Maruyamy) redukują się do tego samego schematu gdy część losowa nie jest zależna od zmiennej x . Jeżeli zależność istnieje, schemat Millsteina będzie szybciej zbieżny od schematu EM.

Stochastyczna wersja algorytmu Rungego-Kutty 2 rzędu

Najczęściej spotykane rozwiązania układów różniczkowych równań stochastycznych bazują na powyższych schematach pierwszego rzędu. Aby uzyskać nieco wyższą dokładność i precyzję obliczeń można pokusić się o implementację algorytmów wyższego rzędu. Algorytmy rzędu 2 lub 4 dla zwyczajnych równań różniczkowych nazywane są metodami Rungego-Kutty odpowiednio 2 i 4 rzędu (RK2, RK4). Analogiem ich będą Stochastyczne metody Rungego-Kutty. Zostały one podane przez Rebeccę Honeycutt w 1992. Idea jest rozszerzenie klasycznej metody RK2 tak by poprawnie całkowała również człon stochastyczny.

Wrócimy na chwilę do równania (4.3). Dość łatwo wyobrazić sobie scałkować numerycznie część deterministyczną metodą RK dowolnego w zasadzie rzędu. Jeżeli np: posłużymy się metodą RK4 a część stochastyczną obliczymy metodą Millsteina to otrzymamy metodę która co prawda zgadza się co do rzędu rozwinięcia dla części deterministycznej, ale część stochastyczna będzie raptem rzędu 1, co w wyniku da jedynie metodę rzędu 1. Do tego wszystkiego niepoprawną. Można wyobrazić sobie, że w jednym kroku algorytmu rozpatrywana cząstka Browna na chwilę zapomina, że działa na nią siła losowa, więc ewoluuje jak zwykła, deterministyczna cząstka, po czym znów zaczyna czuć siłę losową jednocześnie zmieniając się w cząstkę Browna.

Nie jest to oczywiście zbyt rozsądne. Należy rozważyć rozszerzenie RK to SRK. Zaczniemy od równania, zakładając stałą funkcję $b(t, x)$.

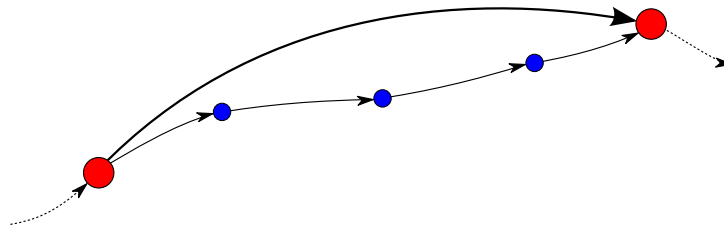
$$x(h) = x_0 + \frac{h}{2}(F_1 + F_2) + \sqrt{2Dh}\phi, \quad (4.14)$$

gdzie

$$F_1 = a(t, x_0), \quad (4.15)$$

$$F_2 = a(t, x_0 + hF_1 + \sqrt{2hD}\phi),$$

dla $\langle \phi \rangle = 0$ oraz $\langle \phi^2 \rangle = 1$.



Rysunek 4.13: Schematyczny krok całkowania stochastycznego równania różniczkowego. Po pojedynczej iteracji o $x_i \rightarrow x_{i+h}$ to przeskoczenie od lewego do prawego czerwonego punktu. Jeżeli zastosujemy przybliżenie RK4 do części deterministycznej, otrzymamy w rezultacie trajektorję zaznaczoną niebieskimi punktami. Gdy odpowiadająca stochastyczna część równania (4.3) obliczona zostanie schematem Millsteina to dostaniemy bezpośredni skok z punktu czerwonego do czerwonego, co oznaczać będzie, że cząstka wykonuje czysto deterministyczny ruch po niebieskiej siatce, zyskując dodatkowy przyrost stochastyczny tylko na końcu przedziału h (w czerwonym prawym punkcie). Ideą SRK jest obliczenie wkładu do dynamiki zarówno od części deterministycznej jak i stochastycznej w *każdym* punkcie w którym próbkujemy system (czyli zarówno w niebieskich jak i czerwonych punktach).

4.4 Przykłady całkowania procesów stochastycznych

4.4.1 Proces dyfuzji

Jest to prawdopodobnie najprostszy proces stochastyczny wykorzystujący biały szum gaussowski jako proces losowy. Przez matematyków nazywany jest po prostu procesem Wienera ponieważ prawa strona równania ruchu zawiera tylko i wyłącznie ów proces. Z drugiej strony jest obok procesu Poissona najważniejszym procesem losowym na bazie którego można zdefiniować całą rodzinę procesów losowych o ciągłych realizacjach. Równanie to można przedstawić używając równania Ito

$$dx(t) = \sqrt{2D}dW(t).$$

Realizacja jest funkcją ciągłą, ale nigdzie nieróżniczkowalną (jako że pochodna procesu Wienera nie istnieje). Za pomocą znanego już schematu Eulera-Maruyamy (EM) możemy sobie wygenerować pojedynczą realizację takiego procesu. Parametr D reguluje natężenie szumu.

$$\begin{aligned} x_0 &= 0 \\ x_{i+1} &= x_i + \sqrt{2hD}N(0, 1). \end{aligned}$$

Wiemy, że dla procesu Wienera $W(0) = 0$, wystartujemy więc z $x(0) = 0$. Weźmy krok $h = 0.01$ i 5000 kroków czasowych. Dla przejrzystości weźmiemy natężenie szumu $D = 1$. Jako, że wiemy jak generować zmienne z rozkładem $N(0, 1)$ użyjemy sobie “symbolicznego” oznaczenia na funkcję zwracającą takie zmienne. Funkcję taką nazwiemy `std_norm`. Konkretna realizacja takiej funkcji może odbywać np: poprzez algorytm Boxa-Mullera. Funkcja ta będzie przy wywołaniu zwracała jedną liczbę losową z $N(0, 1)$.

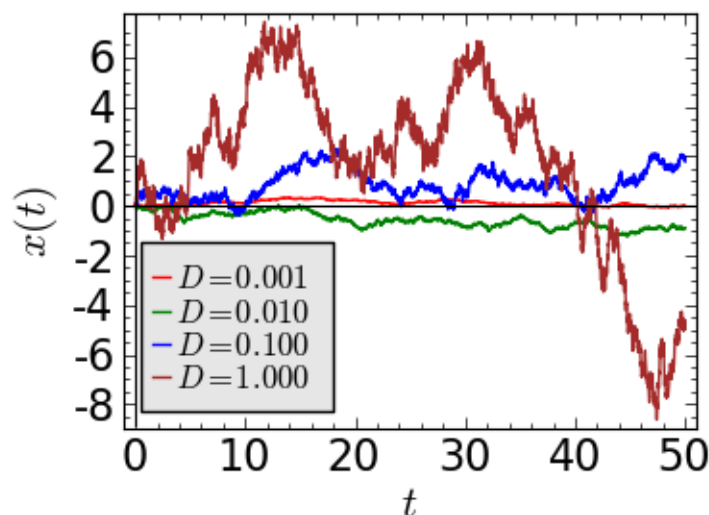
```
h = 0.01
N = 5000
x0 = 0
```



```
D = 1
```

```
x = [x0]
for i in xrange(1,N):
    n01 = std_norm()
    x.append(x[i-1] + sqrt(2*h*D) * n01)
```

Teraz narysujmy sobie takie realizacje dla kilku różnych wartości parametru D .



Rysunek 4.14: Proces dyfuzji dla kilku różnych wartości parametru D .

Na pierwszy rzut oka trajektorie (czy realizacje, przebiegi...) wyglądają kompletnie inaczej. Dla małych wartości D krzywe są bardziej regularne niż dla tych parametryzowanych przez większe wartości D , dla których to wykres jest mocno poszarpany i nieregularny. Jeżeli jednak narysowalibyśmy je osobno, nie oznaczając osi, identyfikacja byłaby niemożliwa - nie widzimy bowiem relacji pomiędzy wartościami (przyrostami).

Jako, że rozwiązanie równania dyfuzji znane jest od dawien dawna, możemy potestować na ile dokładnie rozwiązujemy nasze równanie. Oczywiście jedyne co możemy określić, to rozkład gęstości prawdopodobieństwa. Zakładając warunki początkowe

$$P(x, 0) = \delta(x)$$

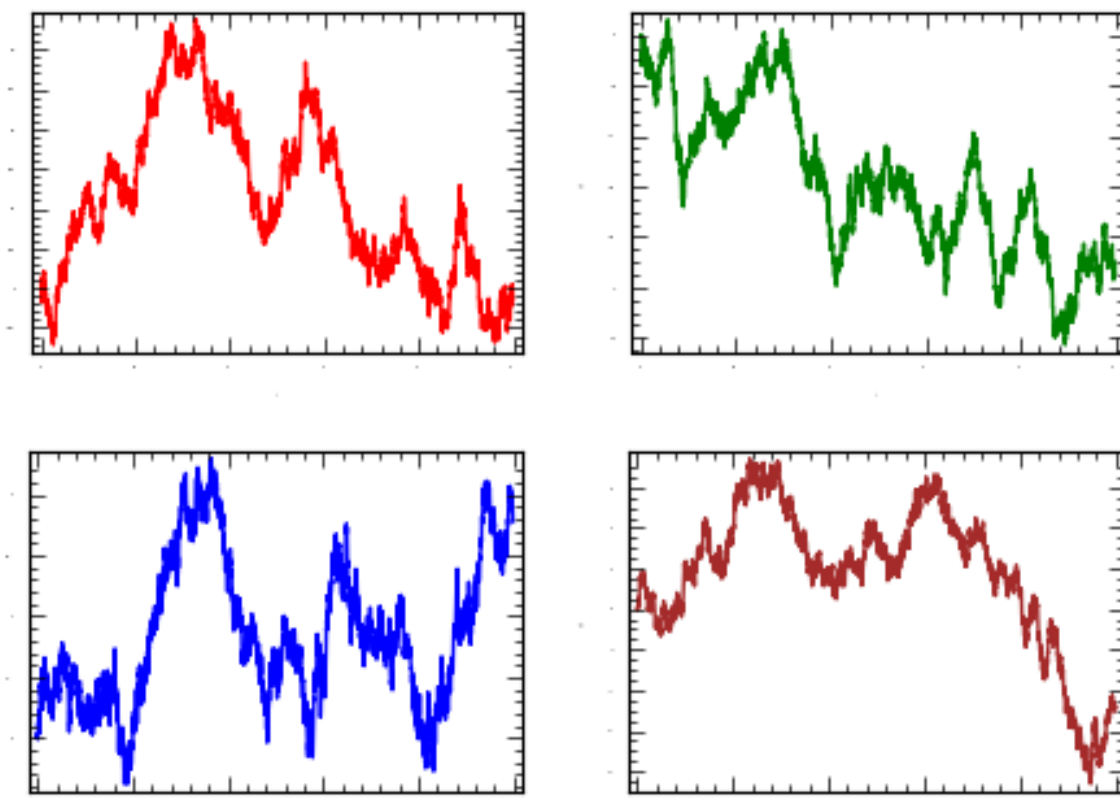
co oznacza, że cały zespół cząstek podlegających dyfuzji wystartuje z $x(0) = 0$, możemy podać odpowiedź, bazując na równaniu dyfuzji

$$\frac{\partial P(x, t)}{\partial t} = D \frac{\partial^2 P(x, t)}{\partial x^2}.$$

Wynikiem jest rozkład Gaussa postaci

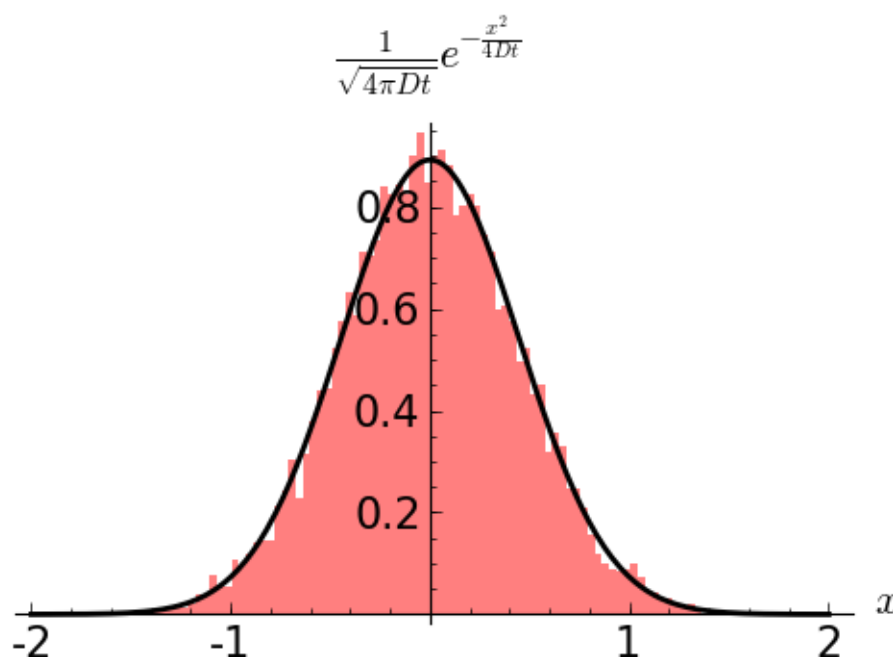
$$P(x, t) = \frac{1}{\sqrt{4\pi Dt}} \exp \left[-\frac{x^2}{4Dt} \right].$$

w którym wariancja rośnie liniowo z czasem a średnia jest równa zero. Zobaczmy, czy jesteśmy w stanie zweryfikować powyższy wzór numerycznie. Postaramy się znaleźć histogram pozycji



Rysunek 4.15: Proces dyfuzji dla kilku różnych wartości parametru D . Są to te same przebiegi co w poprzednim wykresie. Kolory tu użyte odpowiadają kolorom z wykresu poprzedniego.

10000 cząstek po 100 krokach i porównamy go z powyższym wynikiem. Założymy sobie $D = 0.1$ i krok czasowy $h = 0.01$. Oznacza to, że po 100 krokach symulacji osiągniemy rzeczywisty czas równy 1.



Rysunek 4.16: Proces dyfuzji po 100 krokach symulacji dla $D = 0.1$, $h = 0.01$, $N = 100$ a co za tym idzie $t = Nh$.

Zadanie 6.1.1 Wykonaj w Sage powyższy rysunek. Metodę generowania histogramów znajdziesz w pierwszej części skryptu. Listę położeń, którą należy podać do histogramu wygeneruj w podobny sposób jak na początku tego rozdziału.

4.4.2 Dyfuzja ze stałym dryfem

Proces taki otrzymuje się bezpośrednio jako graniczny przypadek niesymetrycznego błędzenia przypadkowego (polecam poczytać [ten podrozdział](#)). Równanie opisujące ten proces ma postać zbliżoną do poprzedniego, bogatsze jest jednak o dodatkowy czynnik, zwany dryfem

$$\frac{\partial P(x, t)}{\partial t} = -V \frac{\partial P(x, t)}{\partial x} + D \frac{\partial^2 P(x, t)}{\partial x^2}.$$

Rozwiązaniem równania dyfuzji z dryfem, z takim samym warunkiem początkowym jak poprzednio (wszystkie cząstki, bądź realizacje zaczynają z tego samego położenia $x(0) = 0$ jest następująca funkcja

$$P(x, t) = \frac{1}{\sqrt{4\pi Dt}} \exp \left[-\frac{(x - Vt)^2}{4Dt} \right].$$

Jest to funkcja Gaussa opisująca zmienne losowe normalne, a dwa pierwsze momenty wynoszą odpowiednio $\xi(t) = Vt$ oraz $\sigma_\xi^2 = 2Dt$. Łatwo zauważyć, że zarówno średnia jak i wariancja zależne są liniowo od czasu. Ponadto wariancja jest identyczna jak w procesie dyfuzji bez

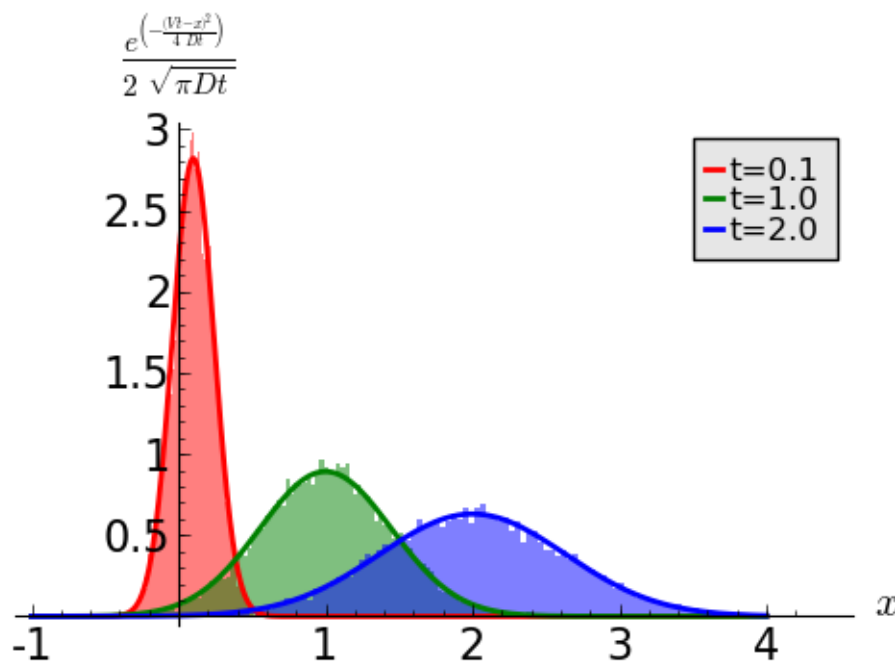
dryftu. Dryft ów stały można z punktu widzenia fizyki rozumieć jako stałą siłę przyłożoną do cząstki (coś na kształt cząstki umieszczonej na równi pochyłej) - położenie cząstki rośnie liniowo z czasem (jak w ruchu jednostajnie przyspieszonym), ale fluktuacje rosną w czasie jak pierwiastek \sqrt{t} .

Podobną analizę numeryczną jak poprzednio możemy przeprowadzić i tutaj. Tym razem, wykreślimy sobie stroboskopowo histogram położenia po kilku krokach: $N = 10, 100, 200$. Po lekkiej modyfikacji numeryczny schemat EM będzie wyglądał tak

```
h = 0.01
N = 5000
x0 = 0
V = 1
D = 1

x = [x0]
for i in xrange(1, N):
    n01 = std_norm()
    x.append(x[i-1] + V*h + sqrt(2*h*D) * n01)
```

Teraz wystarczy zobaczyć, czy histogramy położenia po czasie $t=0.1, 1, 2$ będą odpowiadały obliczonej powyżej funkcji rozkładu.



Rysunek 4.17: Proces dyfuzji ze stałym dryftem po 10, 100 i 200 krokach symulacji dla $D = 0.1$, $h = 0.01$, a co za tym idzie $t = 0.1, 1, 2$.

Możemy policzyć sobie teraz średnie, odchylenie standardowe oraz błędy względny i bezwzględny przybliżenia dokładnych rozwiązań procesu dyfuzji z dryftem.

czas		teoria	symulacje	E_b	E_w [%]
$t=0.1$ std	średnia 0.1414	0.10 0.1413	0.09972 0.00007544	0.0002766 0.05335	0.2766
$t=1$ std	średnia 0.4472	1.0 0.444	1.005 0.003026	0.005345 0.6766	0.5345
$t=2$ std	średnia 0.6324	2.0 0.6362	2.001 0.003708	0.001460 0.5863	0.07302

Jak widzimy błędy bezwzględne dochodzą do około pół punktu procentowego różnicy dla 10000 realizacji. Zwiększenie próby spowoduje jeszcze lepsze dopasowanie, zmniejszenie spowoduje większe odchylenia od wartości rzeczywistych.

Zadanie 6.1.2 Oblicz błędy przybliżenia rozwiązania problemu dyfuzji ze stałym dryftem dla 10, 100, 500 i 1000 różnych realizacji. Zestawienia podaj w tabeli.

4.5 Szum dychotomiczny

Jest to prawdopodobnie najprostszy z możliwych rodzaj skorelowanego szumu, który może wzbudzić transport w układach jakie analizujemy. Proces $\xi(t)$, czyli stacjonarny, dychotomiczny, stochastyczny proces Markowa - szum telegraficzny, przyjmuje losowo dwie wartości:

$$\begin{aligned}\xi(t) &= \{-a, b\}, \quad a, b > 0, \\ P(-a \rightarrow b) &= \mu_a = \frac{1}{\tau_a}, \\ P(b \rightarrow -a) &= \mu_b = \frac{1}{\tau_b},\end{aligned}\tag{4.16}$$

gdzie $P(x_1 \rightarrow x_2)$ jest prawdopodobieństwem skoku z wartości x_1 do x_2 na jednostkę czasu, a τ_a i τ_b wyrażają średni czas przebywania w stanie odpowiednio $-a$ i b . Przy założeniu, że rozkład początkowy szumu $\xi(t)$ jest stacjonarny, czyli:

$$\begin{aligned}P(\xi(0) = b) &= \frac{\mu_a}{\mu_a + \mu_b}, \\ P(\xi(0) = -a) &= \frac{\mu_b}{\mu_a + \mu_b},\end{aligned}\tag{4.17}$$

średnia wartość wynosi

$$\langle \xi(t) \rangle = \frac{b\mu_a - a\mu_b}{\mu_a + \mu_b}.\tag{4.18}$$

Jednak aby średnia siła losowa była zerowa, musimy założyć $b\mu_a = a\mu_b$, co spełnia nam automatycznie zakładany warunek. Funkcja korelacyjna wynosi wówczas

$$\langle \xi(t)\xi(s) \rangle = \frac{Q}{\tau} e^{-|t-s|/\tau},\tag{4.19}$$

gdzie $Q = \tau ab$ jest natężeniem procesu oraz $\tau = 1/(\mu_a + \mu_b)$ jest czasem korelacji. Z powyższego wynika, że szum dychotomiczny jest charakteryzowany trzema parametrami: Q , τ i $\theta = b - a$. Ostatni mówi o symetrii procesu - gdy θ znika proces jest symetryczny, w przeciwnym wypadku jest asymetryczny.

4.5.1 Realizacja szumu dychotomicznego

Proces $\xi(t)$ oznacza, że cząstka losowo przeskakuje ze stanu $-a$ do b i z powrotem, średnio spędzając odpowiednio $1/\mu_a$ czasu w stanie $-a$ i $1/\mu_b$ w stanie b . Czas pobytu w obu stanach jest zmienną losową o rozkładzie eksponencjalnym i średniej $1/\mu_k$, dla $k = a$ lub b , zatem algorytm generujący trajektorie $\xi(t)$ polega na losowaniu liczb z rozkładem eksponencjalnym. Postępować należy więc w sposób następujący:

- generujemy realizację procesu na jednostkę czasu
- całkujemy otrzymane równanie różniczkowe.

Gdy startujemy z wartości przykładowo $-a$, bierzemy skok czasowy $\Delta t = \frac{1}{\mu_a} \hat{z}$ (czas określający jak długo cząstka przebywać będzie w stanie $-a$), gdzie \hat{z} to generator liczb losowych z rozkładem

$$P(\Delta t) = \exp(-\Delta t), \Delta t > 0. \quad (4.20)$$

Aby zapamiętać całą realizację procesu $\xi(t)$ musimy tablicować liczby t_i - skoki pomiędzy wartościami $\{a, b\}$, oraz wartość startową.

```
from scipy import stats
from numpy import cumsum

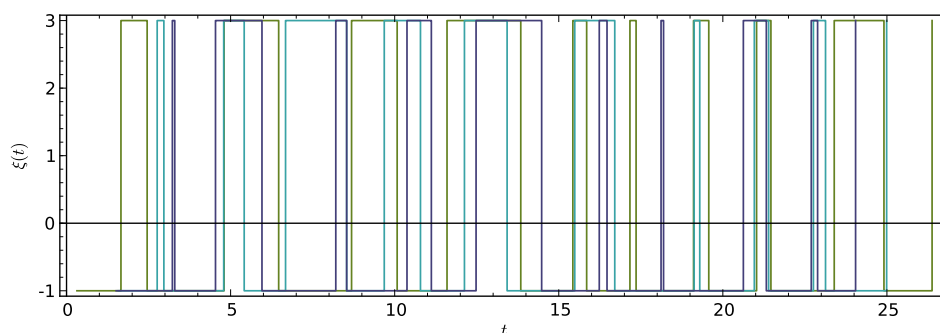
# definicja szumu dychotomicznego
a = -1
b = 3
stan = [a,b]
mu_a = 1
mu_b = mu_a * abs(b) / abs(a)
mu = [mu_a, mu_b]

# realizacja
N = 20
czasy = [-log(random())/mu[i%2]] for i in range(N)
punkty = cumsum(czasy)
stany = [stan[i%2] for i in range(N)]

# wizualizacja
p = plot_step_function(zip(punkty, stany))
```

4.6 Szum kangura

Jako, że kangury skaczą zwykle na skończone odległości, to realizacja procesu ograniczonego wydaje się bardziej sensowna. Niemniej jednak dla komputera sytuacja jest bez różnicy. Metoda symulacji procesu kangura polega na generowaniu stanów do których proces skacze (miejsca w które skacze kangur) oraz przedziałów czasowych w których to proces będzie przebywał (czasy odpoczynku kangura pomiędzy skokami). Jeżeli chodzi o proces kangura, to owe przedziały czasowe zależne są od stanu w jakim znajduje się proces (coś jakby powiedzieć, że im dalej kangur skoczy tym dłużej musi odpocząć...). Algorytm polega na powtarzaniu



Rysunek 4.18: Realizacje szumu dychotomicznego.

1. Obliczamy wartość zmiennej losowej korzystając z jednego z rozkładów
 1. gdy proces $\xi(t)$ jest nieograniczony, określony na przedziale $(-\infty, \infty)$ i stacjonarna gęstość prawdopodobieństwa jest gaussowska:

$$p(z) = Q(z) = \frac{1}{\sqrt{2\pi}\sigma} \exp(-z^2/2\sigma^2), \quad \xi(t) \in (-\infty, \infty) \quad (4.21)$$

2. gdy proces $\xi(t)$ jest ograniczony na przedziale $[-l, l]$ i ma jednostajną gęstość prawdopodobieństwa

$$p(z) = Q(z) = \frac{1}{2l} \theta(z+l) \theta(l-z), \quad \xi(t) \in [-l, l], \quad (4.22)$$

gdzie $\theta(x)$ jest funkcją schodkową Heaviside'a.

2. Teraz obliczamy częstość $\nu(\xi)$.
3. Kiedy już znamy częstość, to obliczamy przedział czasowy w którym będzie przebywał proces kangura w stanie ξ zgodnie z

$$P(T) = \nu(\xi) e^{-T\nu(\xi)} \quad (4.23)$$

Oczywiście rozpatrując szum Kubo-Andersona, gdzie $\nu(\xi) = \nu_0$ jest stałe, wyrażenia mocno się upraszczają.

```
#szum kangura
#szum Kubo - Andersona
#stała częstość \nu
#proces ograniczony

from scipy import stats

l = 2
N = 20
p=[]
```

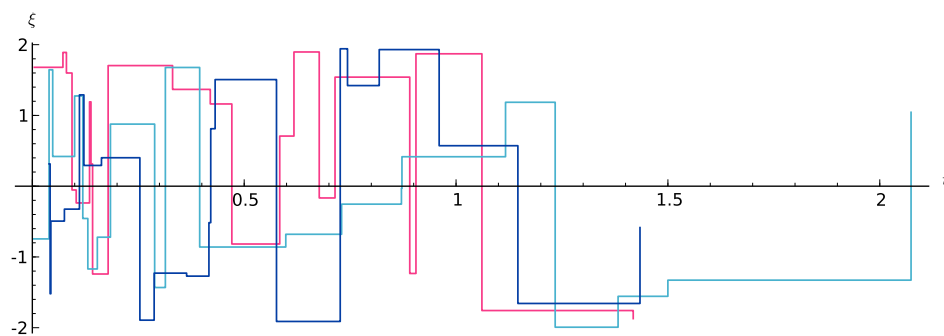
```

for iii in range(3):
    ksi = [2*1*random() - 1 for i in range(N)]
    #list_plot(ksi)

    nu_0 = 2.2
    czasy = stats.expon.rvs(scale=1/nu_0, size=N)

    c=random(), random(), random()
    p.append(plot_step_function(zip(czasy, ksi), color=c))

```



Rysunek 4.19: Realizacje ograniczonego szumu kangura.