

OPRACOWANIE ZAGADNIEŃ NA EGZAMIN DYPLOMOWY

SPIS TREŚCI

1	Zagadnienia obejmujące podstawowe treści programowe kierunku studiów Fizyka Techniczna do egzaminu dyplomowego na studiach II stopnia	2
1.1	Ruch w mechanice newtonowskiej i relatywistycznej	2
1.2	Zasady zachowania i symetrie w fizyce	3
1.3	Klasyczny i kwantowy oscylator harmoniczny	4
1.4	Fizyczna treść równań Maxwella i równania falowego.	5
1.5	Właściwości fal elektromagnetycznych.	7
1.6	Interferencja i dyfrakcja fal.	8
1.7	Zasady termodynamiki.	11
1.8	Entropia w ujęciu termodynamicznym i statystycznym.	12
1.9	Klasyczne i kwantowe rozkłady statystyczne.	13
1.10	Budowa materii.	16
1.11	Dualizm korpuskularno-falowy i jego eksperymentalne potwierdzenie.	18
1.12	Hamiltonian w mechanice klasycznej i kwantowej.	18
1.13	Równanie Schrödingera zależne i niezależne od czasu.	20
1.14	Podstawy formalizmu kwantowego – wielkości fizyczne, stany, operatory.	21
1.15	Kwantowy opis atomu wodoru. Liczby kwantowe.	22
2	Specjalność: Eksploracja Danych i Modelowanie Interdyscyplinarne	23
2.1	Automaty komórkowe.	23
2.2	Podstawowe twierdzenie algorytmów genetycznych.	25
2.3	Sieci bezskalowe.	26
2.4	Agentowe modele dynamiki opinii.	27
2.5	Sieć neuronowa typu Hopfielda.	30
2.6	Naiwny klasyfikator Bayesa.	31
2.7	Metody badania sygnałów w dziedzinie czasu i częstości.	32
2.8	Synchronizacja w nieliniowych układach dynamicznych.	34
2.9	Równanie Fokkera-Plancka.	36
2.10	Fraktalne i multifraktalne metody analizy sygnałów.	37
2.11	Teoria pola średniego dla modelu Isinga.	40
2.12	Zjawisko złamania symetrii w układach rzeczywistych: przykłady, modele.	41
2.13	Informacja wzajemna w przetwarzaniu i transmisji danych.	43
2.14	Metody redukcji wymiaru w eksploracji danych.	44
2.15	Metody wnioskowania statystycznego.	46

1 ZAGADNIENIA OBEJMUJĄCE PODSTAWOWE TREŚCI PROGRAMOWE KIERUNKU STUDIÓW FIZYKA TECHNICZNA DO EGZAMINU DYPLOMOWEGO NA STUDIACH II STOPNIA

1.1 *Ruch w mechanice newtonowskiej i relatywistycznej*

1 Zasada dynamiki Newtona:

I zasada dynamiki Newtona zakłada istnienie inercjalnego układu odniesienia. Układ inercjalny to taki, w którym cząstka nie podlegająca oddziaływaniu z otoczeniem, spoczywa lub porusza się po prostej ze stałą prędkością (układy inercjalne poruszają się ruchem jednostajnym lub spoczywają względem siebie).

2 Zasada dynamiki Newtona:

W inercjalnym układzie odniesienia jeśli siły działające na ciało nie równoważą się ($\vec{F}_w \neq 0$) to ciało porusza się z przyspieszeniem wprost proporcjonalnym do siły wypadkowej, a odwrotnie proporcjonalnym do masy ciała:

$$\vec{a} = \frac{1}{m} * \vec{F}_w$$

$$\vec{F}_w = \frac{d\vec{p}}{dt} = \frac{d}{dt}(m\vec{v}) = m \frac{d\vec{v}}{dt} = m\vec{a}$$

Pierwsza zasada dynamiki Newtona jest szczególnym przypadkiem drugiej zasady dynamiki Newtona (gdy $\vec{F}_w = 0$).

3 Zasada dynamiki Newtona:

Oddziaływania ciał są zawsze wzajemne. Jeżeli ciało A działa na ciało B siłą \vec{F} (akcja), to ciało B działa na ciało A siłą o takiej samej wartości i kierunku, lecz przeciwnym zwrocie (reakcja).

Szczególna teoria względności

1 postulat:

We wszystkich układach inercjalnych prawa fizyki są jednakowe (zasada względności).

2 postulat:

Dla wszystkich obserwatorów inercjalnych prędkość światła w próżni (c) jest taka sama i nie zależy od prędkości źródła światła.

Powyższe postulaty Einsteina prowadzą do transformacji Lorentza:

Rozważmy układ K oraz układ K' poruszający się względem K z prędkością v_x wzdłuż osi OX (dla $t = t' = 0$ początki układów współrzędnych 0_K i $0_{K'}$ pokrywają się), wtedy:

$$t' = \gamma(t - \frac{v_x x}{c^2}), \gamma = \frac{1}{\sqrt{1 - \frac{v_x^2}{c^2}}}$$

$$x' = \gamma(x - v_x t), y' = y, z' = z$$

Konsekwencje szczególnej teorii względności:

- Względność jednoczesności - dwa zdarzenia określone przez jednego obserwatora jako jednoczesne, mogą nie być jednoczesne dla innego obserwatora.
- Dylatacja czasu - czas, jaki mija pomiędzy dwoma zdarzeniami, nie jest jednoznacznie określony, lecz zależy od ruchu obserwatora (paradoks bliźniąt).

- Relatywistyczne składanie prędkości.
- Masa jest równoważna energii $E = mc^2$.
- Ciała bezmasowe poruszają się z prędkością c , dla ciał z niezerową masą niemożliwe jest osiągnięcie prędkości c .
- Skrócenie Lorentza.

1.2 Zasady zachowania i symetrie w fizyce

Jeśli układ posiada pewną symetrię, oznacza to, że równania opisujące ten układ nie zmieniają swojej postaci po dokonaniu przekształceń symetrii.

Dyskretne przekształcenia symetrii to takie, których nie można sparametryzować np.

- Teoria grup i symetrii translacyjnej dla sieci periodycznej w kryształach.
- Symetria permutacyjna funkcji falowej dla układu wielu ciał - związana z nierozróżnialnością cząstek elementarnych (zamiana miejscami cząstek układu nie zmieniłaby równań opisujących układ).
- Symetria zwierciadlana **P** związana z przekształceniem odbicia przestrzennego (zmiana znaków składowych przestrzennych wektorów na przeciwne).
- Odwracalność w czasie **T** (zmiana znaku czasu w równaniach).
- Parzystość ładunkowa **C** (zmiana znaku ładunku).

Elektromagnetyzm, grawitacja i oddziaływania silne są niezmiennicze względem każdej z ostatnich trzech wymienionych symetrii (CPT) osobno, jednakże w przypadku oddziaływań słabych niezmienniczość jest zachowana tylko w przypadku łącznego ich działania CPT (rozpad β łamie symetrię **P** i **C**, ale zachowuje połączoną symetrię **CP**, która dla odmiany jest łamana w przypadku rozpadu mezonów **K**).

Symetrie związane z ciągłymi przekształceniami są bezpośrednio związane z istnieniem zasad zachowania - związek ten opisuje twierdzenie Noether. Zgodnie z tym twierdzeniem, z daną symetrią układu jest związanych tyle praw zachowania, ile ciągłych rzeczywistych parametrów potrzebnych jest do sparametryzowania odpowiadających tej symetrii przekształceń np.

- Zasada zachowania energii wynika z symetrii związanej z przesunięciem w czasie - niezmienniczości działania **S** opisującego ruch danego układu od czasu (t - parametr). Jeżeli układ absorbuje lub emituje energię, wówczas to działanie jest funkcją czasu (t) - odpowiada to w konsekwencji zmianie energii układu.
- Zasada zachowania pędu wynika z symetrii związanej z przesunięciem układu w przestrzeni.
- Zasada zachowania momentu pędu wynika z symetrii związanej z obrotem układu.
- Zasada zachowania ładunku wynika z niezmienniczości funkcji falowej elektronu względem transformacji cechowania.

1.3 Klasyczny i kwantowy oscylator harmoniczny

Klasyczny oscylator harmoniczny – to ciało o masie m na które działa siła proporcjonalna do wychylenia x ciała od stanu równowagi i mająca przeciwny zwrot:

$$\vec{F} = -k\vec{x},$$

gdzie k jest stałą wielkością (tzw. stałą sprężystości). Przykładem oscylatora harmonicznego jest ciało na sprężynie, wykonujące niewielkie drgania od położenia równowagi, co zapewnia słuszność założenia o proporcjonalności siły do wychylenia (dla dużych wychyleń założenie to nie byłoby słuszne). Układ drgający ma energię potencjalną:

$$U(\vec{x}) = \frac{1}{2}kx^2 = \frac{1}{2}m\omega^2x^2,$$

która jest tym większa, im większe jest rozciągnięcie sprężyny ($\omega = \sqrt{\frac{k}{m}}$ jest częstotliwością kołową ruchu drgającego). Energia całkowita układu jest sumą energii kinetycznej i potencjalnej:

$$E = \frac{p^2}{2m} + U(x),$$

gdzie $p = mv$ oznacza pęd ciała drgającego w położeniu x . Całkowita energia układu drgającego harmonicznie nie ulega zmianie w czasie, mimo że energia potencjalna zamienia się cyklicznie w energię kinetyczną i odwrotnie, kinetyczna przechodzi w potencjalną.

W mechanice kwantowej do opisu ruchu układów fizycznych stosuje się zamiast równania Newtona równanie Schrödingera. Konkretna jego postać zależy od opisywanej sytuacji fizycznej. Jedną z metod znalezienia postaci równania Schrödingera w konkretnych przypadkach jest tzw. metoda kwantowania, polegająca na zamianie w równaniach ruchu mechaniki klasycznej pędu ciała p na operator pędu \hat{p} . Współrzędne położenia ciała, np. x pozostawia się przy tym bez zmian (nadając mu teraz nazwę operatora położenia). Słuszność tej metody uzasadnia fakt, że otrzymane za jej pomocą równania dają przewidywania zgodne z wynikami eksperymentów. W przypadku ruchu jednowymiarowego operator pędu ma postać:

$$\hat{p} = -i\hbar \frac{d}{dx}$$

Ponieważ poszukiwany jest opis stanu układu w zależności od współrzędnych x , dlatego trzeba znaleźć jawną postać równania Schrödingera w reprezentacji położeniowej, przy czym dla uproszczenia założymy, że energia układu jest niezmienna. (Podobnie zakłada się, rozwiązując zagadnienie poziomów energetycznych atomu wodoru). Jest to uzasadnione, jeżeli układ drgający pozostaje dłuższy czas w izolacji od otoczenia. Dlatego stosuje się równanie Schrödingera niezależne od czasu:

$$\hat{H}\Psi(x, t) = E\Psi(x, t),$$

gdzie E oznacza energię układu. Pozostaje znalezienie jawnej postaci operatora Hamiltona \hat{H} . W tym celu do wyrażenia na energię całkowitą E oscylatora klasycznego (patrz wyżej) w miejsce klasycznego pędu p podstawia się operator pędu \hat{p} :

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2x^2.$$

Podstawiając jawną postać operatora pędu, otrzymuje się ostatecznie:

$$\hat{H} = \frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \frac{1}{2}m\omega^2x^2.$$

Równanie Schrödingera niezależne od czasu przyjmuje więc postać:

$$\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \psi(x) + \frac{1}{2}m\omega^2x^2 \psi(x) = E\psi(x).$$

Rozwiązując powyższe równanie otrzymujemy zbiór możliwych stanów stacjonarnych (niezależnych od czasu) układu $\psi_n(x)$, dla $n = 0, 1, 2, \dots$. Stanom własnym odpowiadają wartości własne (energie oscylatora) o następujących wartościach:

$$E_n = \hbar\omega(n + \frac{1}{2}), \text{ dla } n = 0, 1, 2, \dots$$

Układ kwantowy drgający harmonicznie przyjmuje tylko wyróżnione wartości energii, czym różni się od układu klasycznego (makroskopowego) – ten ostatni może drgać, mając dowolną wartość energii. Ponieważ drgające układy mikroskopowe faktycznie przyjmują dyskretne poziomy energii, widoczne się staje, że teoria Schrödingera dostarcza właściwego ich opisu. Najmniejsza energia drgań nie jest zerowa, gdyż $E_0 = \frac{\hbar\omega}{2}$. Jest to tzw. energia drgań zerowych, która nie jest znana fizyce klasycznej. Istnienie tej energii oznacza, że układ kwantowy nigdy nie może być w absolutnym spoczynku.

1.4 Fizyczna treść równań Maxwella i równania falowego.

$1) \oint_S \vec{E} \cdot d\vec{S} = \frac{Q}{\varepsilon};$ $2) \oint_l \vec{E} \cdot d\vec{l} = -\frac{\partial \Phi_B}{\partial t};$ $3) \oint_S \vec{B} \cdot d\vec{S} = 0;$ $4) \oint_l \vec{B} \cdot d\vec{l} = \mu I + \mu\varepsilon \frac{\partial \Phi_E}{\partial t}$	$1) \nabla \cdot \vec{E} = \frac{\rho}{\varepsilon};$ $2) \nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t};$ $3) \nabla \cdot \vec{B} = 0;$ $4) \nabla \times \vec{B} = \mu \vec{j} + \mu\varepsilon \frac{\partial \vec{E}}{\partial t}.$
(a) Postać całkowa.	(b) Postać różniczkowa.

Rysunek 1: Równania Maxwella.

Rysunek 8 przedstawia równania Maxwella, gdzie:

ρ - gęstość ładunku,

ε - przenikalność dielektryczna,

μ - przenikalność magnetyczna,

\vec{j} - gęstość prądu,

Φ_B - strumień indukcji magnetycznej,

Φ_E - strumień natężenia pola elektrycznego,

$\nabla * \vec{F} = \frac{\partial F_x(x,y,z)}{\partial x} + \frac{\partial F_y(x,y,z)}{\partial y} + \frac{\partial F_z(x,y,z)}{\partial z}$ - dywergencja pola wektorowego

$\vec{F} = [F_x, F_y, F_z]$,

$\nabla \times \vec{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ F_x & F_y & F_z \end{vmatrix} = (\frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z})\mathbf{i} + (\frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial x})\mathbf{j} + (\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y})\mathbf{k}$ - ro-

tacja pola wektorowego \vec{F} .

Sens fizyczny praw Maxwella:

- 1) Prawo Gaussa dla elektryczności - źródłem pola elektrycznego są ładunki, a strumień tego pola przez dowolną powierzchnię zamkniętą zależy tylko od ładunku zamkniętego przez tę powierzchnię.
- 2) Prawo Faradaya - zmiana strumienia indukcji magnetycznej przez powierzchnię zamkniętej pętli powoduje powstanie w tej pętli siły elektromotorycznej indukcji (SEM), a kierunek płynącego prądu jest taki, żeby przeciwdziałać zmianom powodującym indukcję (reguła Lenza).

- 3) Prawo Gaussa dla magnetyzmu - nie istnieją ładunki magnetyczne, a strumień pola magnetycznego przez dowolną powierzchnię zamkniętą jest równy 0.
- 4) Prawo Ampere'a - zmienne pole elektryczne i płynący prąd powodują powstanie pola magnetycznego.

Dla fali elektromagnetycznej w próżni wektory \vec{E} i \vec{B} drgają w płaszczyznach wzajemnie prostopadłych i dla fali rozchodzącej się w kierunku osi x możemy przyjąć taki układ odniesienia aby wektor \vec{E} drgał w kierunku osi y a wektor \vec{B} w kierunku osi z. Zatem wektory \vec{E} i \vec{B} mają tylko po jednej składowej:

$$\vec{E} = [0, E, 0],$$

$$\vec{B} = [0, 0, B].$$

Liczymy rotacje wektorów \vec{E} i \vec{B} (wykorzystując fakt, że nasza fala jest falą płaską i pola \vec{E} i \vec{B} zmieniają się tylko względem współrzędnej x, czyli że $\frac{\partial E}{\partial z} = 0, \frac{\partial B}{\partial y} = 0$):

$$\nabla \times \vec{E} = \mathbf{k} \frac{\partial E}{\partial x},$$

$$\nabla \times \vec{B} = -\mathbf{j} \frac{\partial B}{\partial x}.$$

Korzystając z równań Faradaya oraz Ampera w postaci różniczkowej otrzymujemy (poszukujemy równania dla fal elektromagnetycznych rozchodzących się w próżni gdzie nie będą występowały prądy przewodzenia, czyli $\vec{j} = 0$):

$$\mathbf{k} \frac{\partial E}{\partial x} = -\frac{\partial \vec{B}}{\partial t} = -\mathbf{k} \frac{\partial B}{\partial t},$$

$$\mathbf{j} \frac{\partial B}{\partial x} = -\mu_0 \epsilon_0 \frac{\partial \vec{E}}{\partial t} = -\mathbf{j} \mu_0 \epsilon_0 \frac{\partial E}{\partial t}.$$

W efekcie dostajemy układ dwóch równań:

$$\frac{\partial E}{\partial x} = -\frac{\partial B}{\partial t},$$

$$\frac{\partial B}{\partial x} = -\mu_0 \epsilon_0 \frac{\partial E}{\partial t}.$$

Równania falowe dla \vec{E} i \vec{B} będą miały identyczną postać. Jeżeli zdecydujemy się szukać równania dla \vec{E} , to eliminujemy z naszego układu równań \vec{B} przez utworzenie pochodnych mieszanych \vec{E} względem x i t. Różniczkujemy zatem pierwsze równanie po x, a drugie po t (jeśli chcemy szukać równania dla \vec{B} eliminujemy w ten sam sposób z naszych równań \vec{E}):

$$\frac{\partial^2 E}{\partial x^2} = -\frac{\partial^2 B}{\partial t \partial x},$$

$$\frac{\partial^2 B}{\partial t \partial x} = -\mu_0 \epsilon_0 \frac{\partial^2 E}{\partial t^2}.$$

Z powyższego układu równań otrzymujemy poszukiwane równanie falowe dla pola \vec{E} :

$$\frac{\partial^2 E}{\partial x^2} - \mu_0 \epsilon_0 \frac{\partial^2 E}{\partial t^2} = 0$$

Znając ogólną postać równania falowego dla fali rozchodzącej się z prędkością v w kierunku osi x:

$$\frac{\partial^2 \xi}{\partial x^2} - \frac{1}{v^2} \frac{\partial^2 \xi}{\partial t^2} = 0$$

otrzymujemy związek pomiędzy prędkością światła w próżni (c) a wartościami przenikalności elektrycznej i magnetycznej próżni:

$$c = \frac{1}{\sqrt{\mu_0 \epsilon_0}}.$$

Rozwiązanie równania falowego dla pola \vec{E} ma postać:

$$\vec{E} = E_0 \sin(kx - \omega t),$$

gdzie: $k = \frac{2\pi}{\lambda}, \omega = ck.$

1.5 Właściwości fal elektromagnetycznych.

Falą elektromagnetyczną nazywamy rozchodzące się w przestrzeni zaburzenie pola elektromagnetycznego. Składowa elektryczna i magnetyczna fali indukują się wzajemnie – zmieniające się pole elektryczne wytwarza zmieniające się pole magnetyczne, a z kolei zmieniające się pole magnetyczne wytwarza zmienne pole elektryczne.

Promieniowanie elektromagnetyczne przejawia właściwości falowe ulegając interferencji, dyfrakcji, spełnia prawo odbicia i załamania. W wyniku superpozycji fal elektromagnetycznych może powstać fala stojąca.

Strumień energii przenoszony przez falę elektromagnetyczną w każdym punkcie przestrzeni określa wektor Poyntinga zdefiniowany jako:

$$\vec{S} = \frac{1}{\mu_0} \vec{E} \times \vec{B},$$

gdzie:

μ_0 - przenikalność magnetyczna próżni,

\vec{E} - natężenie pola elektrycznego,

\vec{B} - indukcja pola magnetycznego.

Choć w elektrodynamice klasycznej energię promieniowania elektromagnetycznego uważa się za wielkość ciągłą, zależną jedynie od natężenia pola elektrycznego i indukcji pola magnetycznego, to zjawiska zachodzące na poziomie atomowym dowodzą, że jest ona skwantowana. Energia pojedynczego kwantu jest zależna tylko od częstotliwości fali ν i wynosi:

$$E = h\nu, \text{ gdzie } h - \text{stała Plancka.}$$

Właściwości fal elektromagnetycznych zależą od długości fali. Promieniowaniem elektromagnetycznym o różnej długości fali są:

- Fale radiowe (długość fali powyżej 1 m) - znajdują bardzo szerokie zastosowanie w telekomunikacji, radiofonii, telewizji, radioastronomii i wielu innych dziedzinach nauki i techniki. Naturalne źródła fal radiowych to między innymi wyładowania atmosferyczne, zorze polarne, radiogalaktyki.
- Mikrofale (od 1 mm do 1 m) - podstawowe zastosowania mikrofali to łączność (na przykład telefonia komórkowa, radiolinie, bezprzewodowe sieci komputerowe) oraz technika radarowa. Fale zakresu mikrofalowego są również wykorzystywane w radioastronomii, a odkrycie mikrofalowego promieniowania tła miało ważne znaczenie dla rozwoju i weryfikacji modeli kosmologicznych. Wiele dielektryków mocno absorbuje mikrofale, co powoduje ich rozgrzewanie i jest wykorzystywane w kuchenkach mikrofalowych, przemysłowych urządzeniach grzejnych i w medycynie.
- Podczerwień (od 700 nm do 1 mm) - promieniowanie podczerwone jest nazywane również ciepłym, szczególnie gdy jego źródłem są nagrzane ciała. Każde ciało o temperaturze większej od zera bezwzględnego emituje takie promieniowanie, a ciała o temperaturze pokojowej najczęściej promieniowania emitują w zakresie długości fali rzędu 10 μm . Przedmioty o wyższej temperaturze emitują promieniowanie o większym natężeniu i mniejszej długości, co pozwala na zdalny pomiar ich temperatury i obserwację za pomocą urządzeń rejestrujących wysyłane promieniowanie (termowizja).
- Światło widzialne (od 380 nm do 700 nm) - światło (promieniowanie widzialne) to ta część widma promieniowania elektromagnetycznego, na którą reaguje zmysł wzroku człowieka. Różne zwierzęta mogą widzieć w nieco różnych zakresach. Światło ma bardzo duże znaczenie w nauce i wiele

zastosowań w technice. Dziedziny nauki i techniki zajmujące się światłem noszą nazwę optyki.

- Ultrafiolet (od 10 nm do 380 nm) - promieniowanie ultrafioletowe jest zaliczane do promieniowania jonizującego, czyli ma zdolność odrywania elektronów od atomów i cząsteczek. W technice ultrafiolet stosowany jest powszechnie. Powoduje świecenie (fluorescencję) wielu substancji chemicznych. W świetłówkach ultrafiolet wytworzony na skutek wyładowania jarzeniowego pobudza luminofor do świecenia w zakresie widzialnym. Niektóre owady, na przykład pszczoły, widzą w bliskiej światłu widzialnemu części widma promieniowania ultrafioletowego, również rośliny posiadają receptory ultrafioletu.
- Promieniowanie rentgenowskie (od 5 pm do 10 nm) - promieniowanie rentgenowskie jest promieniowaniem jonizującym. Technicznie promieniowanie rentgenowskie uzyskuje się przeważnie poprzez wyhamowywanie rozprędzonych cząstek naładowanych. W lampach rentgenowskich są to rozprędzone za pomocą wysokiego napięcia elektrony hamowane na metalowych anodach. Źródłem wysokoenergetycznego promieniowania rentgenowskiego są również przyspieszane w akceleratorach cząstki naładowane. Promieniowanie rentgenowskie jest wykorzystywane do wykonywania zdjęć rentgenowskich do celów defektoskopii i diagnostyki medycznej. W zakresie promieniowania rentgenowskiego są również prowadzone obserwacje astronomiczne.
- Promieniowanie gamma (0,03 pm do 300 pm) - promieniowania gamma jest promieniowaniem jonizującym. Promieniowanie gamma towarzyszy reakcjom jądrowym, powstaje w wyniku anihilacji – zderzenie cząstki i antycząstki, oraz rozpadów cząstek elementarnych. Otrzymywane w cyklotronach promieniowanie hamowania i synchrotronowe również leży w zakresie długości fali promieniowania gamma, choć niekiedy bywa nazywane wysokoenergetycznym promieniowaniem rentgenowskim. Promienie gamma mogą służyć do sterylizacji żywności i sprzętu medycznego. W medycynie używa się ich w radioterapii oraz w diagnostyce. Zastosowanie w przemyśle obejmują badania defektoskopowe. Astronomia promieniowania gamma zajmuje się obserwacjami w tym zakresie długości fal.

1.6 Interferencja i dyfrakcja fal.

Interferencja jest to zjawisko powstawania nowego, przestrzennego rozkładu amplitudy fali (wzmocnienia i wygaszenia) w wyniku nakładania się (superpozycji fal) dwóch lub więcej fal. Warunkiem trwałej interferencji fal jest ich spójność, czyli korelacja faz i równość częstotliwości.

Dla najprostszego przypadku dwóch fal harmoniczných o jednakowych amplitudach A , jednakowej długości fali λ i zgodnych fazach początkowych, rozchodzących się z dwóch różnych źródeł, które leżą w odległościach odpowiednio d_1 i d_2 od punktu P , zaburzenie w punkcie P opisuje wzór:

$y(P) = A \sin(\omega t + \phi_1) + A \sin(\omega t + \phi_2)$, gdzie:

$$\phi_1 = 2\pi \frac{d_1}{\lambda},$$

$$\phi_2 = 2\pi \frac{d_2}{\lambda}.$$

Gdy spełniony jest warunek:

$$\phi_1 - \phi_2 = 2\pi \frac{d_1 - d_2}{\lambda} = 2k\pi,$$

gdzie k - dowolna liczba naturalna, to fale w punkcie P ulegają wzmocnie-

niu (są w tej samej fazie) i:

$$y(P) = 2A \sin(\omega t).$$

Gdy natomiast w pewnym punkcie P_1 :

$$\phi_1 - \phi_2 = 2\pi \frac{d_1 - d_2}{\lambda} = (2k + 1)\pi$$

wtedy fale wygaszają się (są w przeciwfazie) i:

$$y(P_1) = 0.$$

Interferencja pozwala na bardzo precyzyjny pomiar zmian długości drogi od źródła do detektora fali.

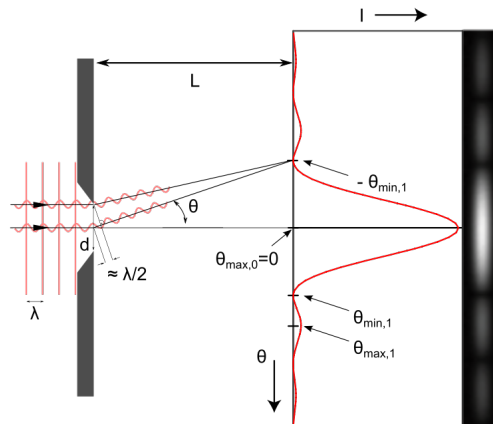
Dyfrakcja (ugięcie fali) – zespół zjawisk związanych ze zmianą kierunku rozchodzenia się fali będący odstępstwem od praw optyki geometrycznej (dział optyki zajmujący się wytłumaczeniem zjawisk optycznych przy użyciu pojęcia promienia). Dyfrakcję w węższym znaczeniu określa się jako ugięcie światła wokół krawędzi przeszkody lub otworu w obszarze cienia przeszkody.

Zjawisko dyfrakcji rozpatruje się jako interferencję fal cząstkowych powstających zgodnie z zasadą Huygensa. Jest to zasada stosowana do określenia rozchodzenia się fali w ośrodku, mówiąca, iż każdy punkt ośrodka, do którego dotarło czoło fali, można uważać za źródło nowej fali kulistej. Wyпадkową powierzchnię falową tworzy powierzchnia styczna do wszystkich powierzchni fali cząstkowych i ją właśnie obserwuje się w ośrodku.

Rysunek 2 przedstawia dyfrakcję na pojedynczej szczelinie. Gdy różnica dróg fali ze skrajnego i środkowego elementu równa jest połowie długości fali, to fale z obu połówek szczeliny wygaszają się. Jeżeli wiązki są niemal równoległe, to różnica odległości pojawia się tylko przy wyjściu promieni ze szczeliny, wówczas:

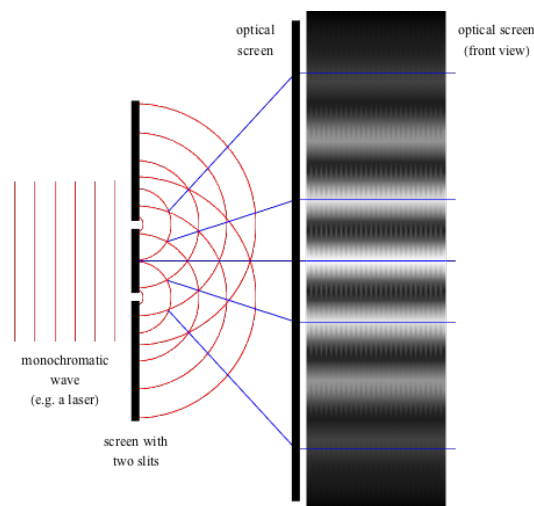
$$d \sin(\theta_{min}) = \lambda.$$

Przepuszczenie fali przez szczelinę dyfrakcyjną pozwala na określenie kierunku rozchodzenia się fali. Im mniejsza jest szerokość szczeliny, tym dokładniej można to zrobić. Jednocześnie zmniejszanie szczeliny powoduje, że trudniej jest określić energię fali, ponieważ rozprasza się ona na większy obszar. W efekcie iloczyn błędu określenia energii oraz błędu pomiaru kierunku musi być większy od pewnej stałej. Oznacza to, że istnieje granica dokładności pomiaru parametrów rozchodzącej się fali. Próba dokładniejszego określenia jednego z parametrów fali powoduje zwiększenie niepewności pomiaru drugiego sprzężonego z nim. Zjawisko to ma fundamentalne znaczenie, jeżeli weźmie się pod uwagę, że każda materialna cząstka jest falą. Zjawisko to w mechanice kwantowej odpowiada zasadzie nieoznaczoności



Rysunek 2: Dyfrakcja na pojedynczej szczelinie.

Pierwszą najprostszą formę eksperymentu przejścia światła przez podwójną szczelinę (rysunek 3), zwaną obecnie doświadczenie Younga wykonał Thomas Young w 1801 roku. Eksperyment ten był zaczątkiem do uznania w XIX w falowej teorii światła.



Rysunek 3: Przejście światła przez podwójną szczelinę.

Dla obrazów dyfrakcyjnych powstałych po przejściu światła przez otwór kołowy definiuje się warunek Rayleigha określający zdolność rozdzielczą elementów i układów optycznych:

$\phi \approx \sin(\phi) = 1.22 \frac{\lambda}{d}$ (przybliżenie dla małych kątów), gdzie:

ϕ - minimalny kąt między promieniami, których obrazy mają być rozróżnialne, czyli inaczej – ich odległość kątowa,

λ - długość fali światła,

d - średnica otworu.

Aby wzmocnić falę przechodzącą przez szczelinę stosuje się w optyce układy wielu takich szczelin, nazywane siatką dyfrakcyjną (rysunek 4). Efekty optyczne od każdej szczeliny dodają się, przez co zachowanie fali zależy tylko od stałej siatki (odległości dzielącej najbliższe sobie rysy). Zjawisko dyfrakcji zachodzi również, kiedy fale przechodzą przez wiele blisko

siebie położonych warstw. Jeżeli odległość między warstwami jest stała, kolejne maksima fali można opisać zależnością:

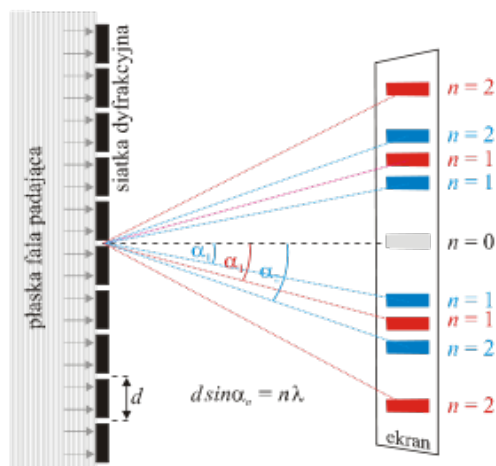
$$\sin(\theta) = \frac{n\lambda}{d}, \text{ gdzie:}$$

d - stała siatki,

θ - kąt od osi wiązki światła,

λ - długość fali,

d - przyjmuje wartości całkowite dodatnie od 1,2,3,...



Rysunek 4: Siatka dyfrakcyjna.

Zjawisko dyfrakcji pozwoliło na rozwój krystalografii rentgenowskiej, dzięki której badano strukturę kryształów, odkryto w ten sposób także strukturę spirali DNA.

1.7 Zasady termodynamiki.

Zerowa zasada termodynamiki - jeśli układy A i B mogące ze sobą wymieniać ciepło są ze sobą w równowadze termicznej, i to samo jest prawdą dla układów B i C, to układy A i C również są ze sobą w równowadze termicznej. Z zasady tej wynika istnienie temperatury empirycznej. Temperatura to taka wielkość fizyczna, która dla układów A, B i C jest równa, gdy ustaje przepływ ciepła. Układy będą ze sobą w równowadze termodynamicznej. Zerowa zasada termodynamiki stwierdza także, że ciało w równowadze termodynamicznej ma wszędzie tę samą temperaturę.

Pierwsza zasada termodynamiki - zmiana energii wewnętrznej (będącej funkcją stanu) układu zamkniętego (nie wymienia masy z otoczeniem) jest równa energii, która przepływa przez jego granice na sposób ciepła i pracy: $\Delta U = W + Q$, gdzie:

ΔU - zmiana energii wewnętrznej układu,

W - praca wykonana na układzie,

Q - ciepło przekazane do układu.

Alternatywne sformułowanie - nie istnieje perpetuum mobile pierwszego rodzaju (maszyna, która wytwarza więcej energii, niż sama zużywa).

Druga zasada termodynamiki - w układzie termodynamicznie izolowanym istnieje funkcja stanu zwana entropią, która nie maleje z czasem. Matematyczny zapis tego faktu to następujące sformułowanie: zmiana entropii

ΔS w dowolnym procesie odwracalnym jest równa całce z przekazu ciepła dQ podzielonego przez temperaturę T . W procesie nieodwracalnym natomiast zmiana entropii jest większa od tej całki: $\Delta S \geq \int \frac{dQ}{T}$. Różnica ta jest miarą nieodwracalności procesu i jest związana z rozpraszaniem energii. Entropia (S) jest funkcją stanu będącą miarą liczby sposobów (W), na jakie może być zrealizowany określony stan termodynamiczny danego układu w określonej temperaturze (T). Układ dąży do stanu, który może być w danych warunkach zrealizowany na jak najwięcej sposobów; dąży więc on do maksymalizacji entropii. Nie można bez wkładu pracy przesyłać energii termicznej między ciałami mającymi tę samą temperaturę. Oznacza to, że perpetuum mobile drugiego rodzaju (maszyna, która zamienia energię cieplną na pracę mechaniczną bez wzrostu całkowitej entropii) nie istnieje.

Trzecia zasada termodynamiki - nie można za pomocą skończonej liczby kroków uzyskać temperatury zera bezwzględnego (zero kelwinów), jeżeli za punkt wyjścia oberzemy niezerową temperaturę bezwzględną. Podstawą takiego zdefiniowania III zasady termodynamiki jest analiza sprawności lodówki. Jak wiemy, lodówka działa na zasadzie odwrotnego cyklu Carnota, a jej sprawność dana jest wzorem:

$$\eta = \frac{Q_{odebrane}}{W} = \frac{T_2}{T_1 - T_2}$$

Jeżeli ciało o określonej temperaturze T_1 chcielibyśmy schłodzić do $T_2 \rightarrow$, odbierając przy tym skończone ciepło Q , to analizując wzór widzimy, że w takim wypadku $\frac{Q}{W} \rightarrow 0$, czyli $W \rightarrow$ nieskończoności. Gdybyśmy podstawili $T_2 = 0$, równanie nie miałoby sensu matematycznego, co oznacza, że nie da się osiągnąć temperatury zera bezwzględnego w skończonej liczbie kroków. Mówiąc inaczej, gdyby udało się schłodzić jakąś substancję do 0 K i gdyby utworzyła ona kryształ doskonały nieposiadający zamrożonych defektów krystalicznych, to jej entropia musiałaby przyjąć wartość 0 . Jest to jednak technicznie, a także formalnie, niewykonalne: $\lim_{T \rightarrow 0} S = 0$

1.8 Entropia w ujęciu termodynamicznym i statystycznym.

Entropia jest to termodynamiczna funkcja stanu, określająca kierunek przebiegu procesów spontanicznych (samorzutnych) w odosobnionym układzie termodynamicznym. Entropia jest miarą stopnia nieuporządkowania układu i rozproszenia energii. Jest wielkością ekstensywną. Zgodnie z drugą zasadą termodynamiki, jeżeli układ termodynamiczny przechodzi od jednego stanu równowagi do drugiego, bez udziału czynników zewnętrznych (a więc spontanicznie), to jego entropia zawsze rośnie. W ramach II zasady termodynamiki zmiana entropii (w procesach kwazistatycznych) jest zdefiniowana przez swoją różniczkę zupełną jako:

$$dS = \frac{1}{T} \delta Q, \text{ gdzie:}$$

T - temperatura bezwzględna,

δQ - ciepło elementarne, czyli niewielka ilość ciepła dostarczonego do układu.

Entropię pewnego stanu termodynamicznego P można wyznaczyć ze wzoru:

$$S(P) = \int_0^{T_p} \frac{C(T)dT}{T}, \text{ gdzie:}$$

$C(T)$ - pojemność cieplna (ilość ciepła, jaka jest niezbędna do zmiany temperatury ciała o jednostkę temperatury),

T_p - temperatura w stanie P . Podstawowe równanie termodynamiki fenomenologicznej, w którym występuje entropia, ma postać:

$$dU = TdS - pdV + \sum_{i=1}^k \mu_i dN_i, \text{ gdzie:}$$

U - energia wewnętrzna,
 k - liczba różnych składników,
 T - temperatura,
 p - ciśnienie,
 μ_i - potencjał chemiczny i-tego składnika.

W termodynamice statystycznej całkowita entropia układu makroskopowego jest równa:

$$S = k \ln(W)$$

lub

$$S = -k \sum_i p_i \ln(p_i), \text{ gdzie:}$$

k - stała Boltzmanna,

W - liczba sposobów, na jakie makroskopowy stan termodynamiczny układu (makrostan) może być zrealizowany poprzez stany mikroskopowe (mikrostany),

p_i - prawdopodobieństwo i-tego mikrostanu.

Zatem $\log_2(W)$ jest liczbą bitów potrzebnych do określenia, którą realizację przyjął dany układ.

1.9 Klasyczne i kwantowe rozkłady statystyczne.

Gaz doskonały - zwany też gazem idealnym, jest to abstrakcyjny, matematyczny model fizyczny gazu, spełniający następujące warunki:

- brak oddziaływań międzycząsteczkowych z wyjątkiem odpychania w momencie zderzeń cząsteczek,
- objętość cząsteczek jest znikoma w stosunku do objętości gazu.
- zderzenia cząsteczek są doskonale sprężyste,
- cząsteczki znajdują się w ciągłym, chaotycznym ruchu.

Rozkład Maxwella – wzór określający rozkład prędkości cząstek gazu doskonałego. Założenie: funkcja rozkładu f zależy tylko od wartości prędkości v :

$$f(v_x, v_y, v_z) = f(v_x^2 + v_y^2 + v_z^2)$$

Żaden kierunek nie jest uprzywilejowany, ruch w każdym z kierunków x, y i z odbywa się niezależnie:

$$f(v_x, v_y, v_z) = h(v_x)h(v_y)h(v_z)$$

Funkcja rozkładu h jest taka sama dla każdego kierunku co wynika z symetrii.

$$\ln[h(v_x)] + \ln[h(v_y)] + \ln[h(v_z)] = \ln[f(v_x^2 + v_y^2 + v_z^2)], \text{ stąd:}$$

$$h(v_x) = C_x e^{(-Bv_x^2)}$$

$$f(v) = C_x C_y C_z e^{-B(v_x^2 + v_y^2 + v_z^2)} = C e^{-Bv^2}$$

Stałą C znajdziemy z warunku normalizacji natomiast stałą B znajdujemy na podstawie równania łączącego średni kwadrat prędkości cząsteczek z temperaturą oraz wyrazić wartość średnią kwadratu prędkości poprzez funkcję gęstości prawdopodobieństwa:

$$\frac{m \langle v^2 \rangle}{2} = \frac{3}{2} kT = \frac{m}{2} \int_0^\infty v^2 f(v) * 4\pi v^2 dv, \text{ gdzie: } m - \text{masa cząsteczki,}$$

k - stała Boltzmanna,

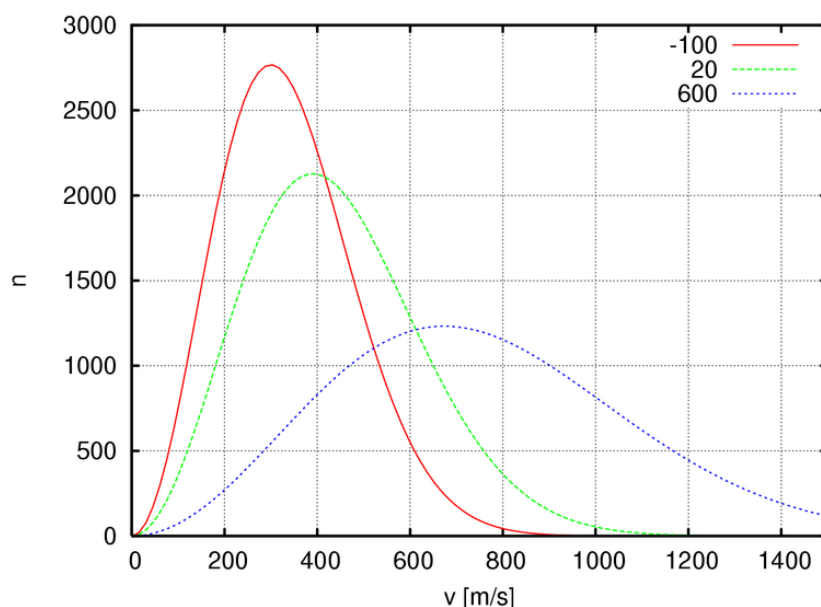
$4\pi v^2 dv$ - element objętości przestrzeni prędkości ($v, v+dv$) we współrzędnych sferycznych.

Otrzymujemy ostatecznie:

$$f(v) = \left(\frac{m}{2\pi kT}\right)^{\frac{3}{2}} e^{-\frac{mv^2}{2kT}}$$

$\frac{dP}{dv} = 4\pi \left(\frac{m}{2\pi kT}\right)^{\frac{3}{2}} v^2 e^{\frac{mv^2}{2kT}}$ - gęstość prawdopodobieństwa wystąpienia cząsteczki o prędkości v .

Rozkład Maxwella pokazuje, że prędkości cząsteczek zależą od temperatury oraz masy molowej. Wraz ze wzrostem temperatury rozkład się poszerza (rysunek 5), a jego prędkość najbardziej prawdopodobna (tam gdzie dP/dv jest największe), jak i średnia prędkość ($\langle v \rangle$) oraz średnia prędkość kwadratowa ($\sqrt{\langle v^2 \rangle}$) zwiększają się.



Rysunek 5: Rozkład Maxwella dla tlenu dla trzech temperatur ($-100\text{ }^{\circ}\text{C}$, temperatura pokojowa i $600\text{ }^{\circ}\text{C}$). Wartość funkcji odpowiada liczbie cząsteczek spośród 1 miliona cząsteczek, jaka będzie poruszać się z prędkością $v \pm 0,5\text{ m/s}$.

Wzór na koncentrację cząsteczek, czyli ich liczbę w jednostce objętości:

$$n = n_0 \exp\left(-\frac{E_p}{kT}\right), \text{ gdzie:}$$

E_p - energia potencjalna cząstek w danym stanie (np. na danej wysokości),
 n_0 - koncentracja cząsteczek dla $E_p = 0$ (np. koncentracja na powierzchni ziemi).

Wzór ten wyraża zależność koncentracji cząsteczek od ich wysokości lub energii potencjalnej. Wynikający z niego rozkład koncentracji nosi nazwę rozkładu Boltzmanna i odnosi się nie tylko do pola sił przyciągania ziemskiego, ale do dowolnego pola potencjalnego, jeśli tylko cząsteczki poruszają się chaotycznym ruchem cieplnym. Liczba cząsteczek w objętości $dV = dz \cdot dy \cdot dx$, której położenie określają współrzędne (x, y, z) wynosi:

$$dn(x, y, z) = n_0 \exp\left(-\frac{E_p(x, y, z)}{kT}\right) dx \cdot dy \cdot dz$$

Nietrudno tu zauważyć podobieństwo rozkładów Maxwella i Boltzmanna, z tą różnicą, że pierwszy wyraża zależność od kwadratu prędkości lub energii kinetycznej, drugi - od energii potencjalnej. Rozkład zawierający obie te zależności nosi nazwę rozkładu Maxwella-Boltzmanna i ma postać:

$$dn(x, y, z, v_x, v_y, v_z) = A \exp\left(-\frac{E_p + E_k}{kt}\right) dx \cdot dy \cdot dz \cdot v_x \cdot v_y \cdot v_z$$

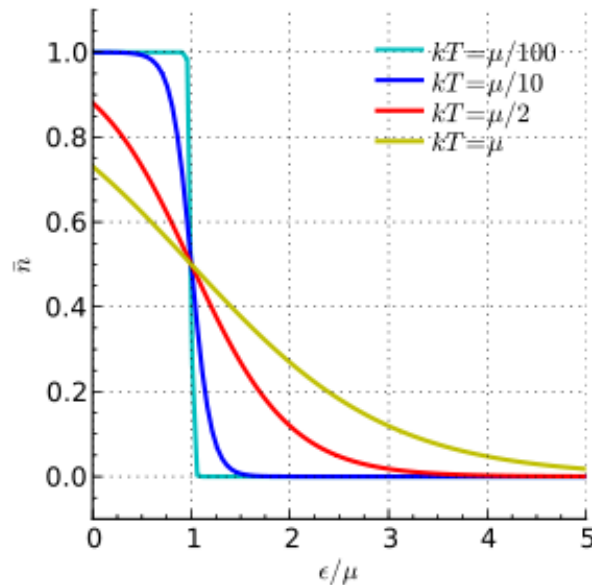
Zakaz Pauliego głosi, że prawdopodobieństwo znalezienia w układzie fermionów (cząstek o spinie połowkowym) pary cząstek o jednakowych liczbach

bach kwantowych jest równe zero (czyli np. możemy znaleźć na tej samej orbicie atomowej dwa elektrony, ale muszą mieć przeciwne spiny - jest to przykład degeneracji).

Zgodnie z rozkładem Fermiego-Diraca (rysunek 6) średnia liczba cząstek w niezdegenerowanym stanie energetycznym E dana jest przez:

$$\langle n \rangle = \frac{1}{e^{\beta(E-\mu)} + 1}, \text{ gdzie:}$$

E - energia tego stanu,
 μ - potencjał chemiczny,
 $\beta = 1/k_B T$



Rysunek 6: Rozkład Fermiego-Diraca. Oś pozioma: E/μ . Oś pionowa: $\langle n \rangle$.

W temperaturze zera bezwzględnego wprowadza się oznaczenie $\mu = \mu(0) = E_F$ - jest to energia najwyższej obsadzonego stanu (k_F - poziom Fermiego) w temperaturze zera bezwzględnego. W tej temperaturze obsadzone są wszystkie stany o energii mniejszej lub równej energii Fermiego (E_F) a wyższe stany nie są obsadzone. Dla każdej temperatury T zachodzi $P(E_k) = 0.5$ gdy $E_k = \mu$. Dla takich energii, że $E_k - \mu \gg k_B T$ rozkład przechodzi w klasyczny rozkład Boltzmanna.

Statystyka Bosego-Einsteina dotyczy bozonów (cząstek o spinie całkowitym, których nie obowiązuje zakaz Pauliego) traktowanych jako gaz bozonowy. Zgodnie z rozkładem Bosego-Einsteina średnia liczba cząstek w danym stanie kwantowym jest równa:

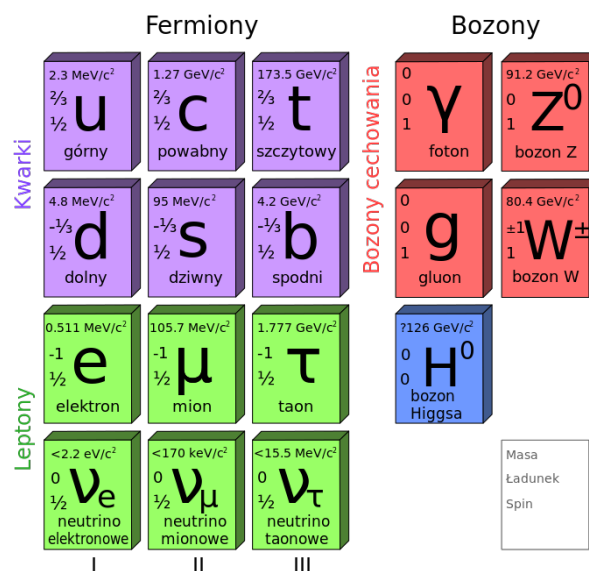
$$\langle n_i \rangle = \frac{n}{Z} \frac{g_i}{e^{\beta(E_i - \mu)} - 1}, \text{ gdzie:}$$

E_i - energia i-tego stanu,
 g_i - degeneracja i-tego stanu,
 n - całkowita liczba cząstek,
 μ - potencjał chemiczny,
 $Z = \sum_i \frac{g_i}{e^{\beta(E_i - \mu)} - 1}$ - suma statystyczna.

Potencjał chemiczny w tym rozkładzie jest zawsze ujemny lub równy zero. Gdy temperatura jest wysoka, można zaniedbać składnik -1 i rozkład przechodzi w rozkład fizyki klasycznej - klasyczny rozkład Boltzmanna. Rozkładowi Bosego-Einsteina podlegają fotony (o spinie 1) – nosi on wtedy nazwę rozkładu Plancka, który tłumaczy promieniowanie ciała doskonale czarnego. Jego wprowadzenie przez Plancka zapoczątkowało mechanikę kwantową. Zakaz Pauliego nie dotyczy bozonów, umożliwia to ich kondensację. Kondensacja Bosego-Einsteina – efekt kwantowy zachodzący w układach podległych rozkładowi Bosego-Einsteina. W temperaturach niższych od temperatury krytycznej część cząstek (bozonów) przechodzi w zerowy stan pędowy – cząstki te mają identyczny pęd. Oznacza to, że w zerowej objętości przestrzeni pędów może znajdować się niezerowa liczba cząstek. Mówi się wtedy o makroskopowym obsadzeniu stanu podstawowego. Efektem kondensacji jest kolektywne zachowanie wszystkich cząstek biorących w niej udział (w przybliżeniu wszystkie zachowują się jak jedna cząstka). Nie chodzi tu o kondensację w zwykłym sensie w przestrzeni położeniowej – cząstki nie znajdują się w jednym miejscu, lecz o "kondensację" cząstek w przestrzeni pędów – znaczna liczba cząstek ma taki sam pęd. Rozkład przestrzenny cząstek "skondensowanych" pozostaje równomierny (jeśli nie ma pól zewnętrznych). W kondensacie Bosego-Einsteina zachodzi zjawisko nadciekłości.

1.10 Budowa materii.

Model standardowy – teoria fizyki cząstek podstawowych, zwanych też cząstkami elementarnymi, które są podstawowymi składnikami każdej materii (rys. 7). Opisuje trzy z czterech (z wyjątkiem grawitacji) oddziaływań podstawowych: elektromagnetyczne, słabe i silne. Sformułowana jest w języku matematyki, opisując relacjami matematycznymi zależności między elementami tej teorii. Opiera się na koncepcji pola Yanga-Millsa (pole rządzące oddziaływaniem wszystkich znanych cząstek we wszechświecie).



Rysunek 7: Cząstki elementarne według Modelu Standardowego.

Fermiony są podstawowymi elementami budującymi materię. Materię trwałą, która nas otacza, tworzą następujące cząstki: elektron, kwark górny (u) oraz kwark dolny (d) np. dwa kwarki górne i jeden dolny (uud) tworzą proton. Kwarki mają kolor i zapach (są to nazwy nadane umownie pewnym liczbom kwantowym). Do tej grupy cząstek należy też neutrino elektronowe, a grupa ta tworzy pierwszą generację. Kolejne dwie generacje zawierają po cztery cząstki odpowiadające cząstkom z pierwszej generacji (lecz o różnej masie).

Bozony cechowania przenoszą oddziaływania:

- elektromagnetyczne - przenoszone jest przez foton. Oddziaływanie to odbywa się poprzez wytworzenie lub pochłonięcie fotonu.
- słabe - powodujące między innymi rozpady beta, przenoszone jest przez bozony W^+ i W^- oraz Z^0 .
- silne - łączące kwarki w hadrony, przenoszone jest przez osiem rodzajów gluonów, sposób oddziaływania (rodzaj gluonu) oznaczany jest właściwością nazywaną kolorem gluonu. Hadrony złożone są z kwarków występujących przeważnie trójkami (trzy kwarki lub trzy antykwarki tworzą bariony - spin połówkowy, liczba barionowa = 1) lub parami (pary kwark i antykwark tworzą mezony - spin całkowity, liczba barionowa = 0). Właściwością hadronów jest również ich całkowity ładunek elektryczny.

Bozon Higgsa (istnienie którego udało się potwierdzić doświadczalnie w 2012 roku) oddziałując z innymi cząstkami nadaje im masę (głównie dotyczy nadawania masy elektronowi, nie dotyczy nadawania masy protonowi i neutronowi, których masa wynika z innego mechanizmu).

Model standardowy jest potwierdzony doświadczalnie, lecz nie jest w pełni satysfakcjonujący z teoretycznego punktu widzenia:

- Ma 19 swobodnych parametrów (np. masy cząstek), które należy wyznaczyć doświadczalnie, gdyż teoria nie wyjaśnia ich wartości.
- Obliczenia masy Wszechświata nie zgadzają się z obserwowaną ilością materii we Wszechświecie, brakującą materię nazywa się ciemną materią.
- W podstawowej wersji nie uwzględnia mas neutrin.

Protony i neutrony związane oddziaływaniem silnym (za pośrednictwem mezonów μ) tworzą jądra atomowe. Protony odpychają się elektrostatycznie, jednak jądro utrzymywane jest w całości przez oddziaływanie silne. Działa ono tylko na niewielką odległość, dlatego jądra zbyt duże i masywne stają się nietrwałe, co prowadzi do samorzutnego ich rozpadu. Liczba protonów w jądrze atomowym to liczba atomowa, a suma liczb protonów i neutronów to liczba masowa.

Obiekt fizyczny złożony z jądra atomowego i znajdujących się w otoczeniu jądra, związanych z nim oddziaływaniem elektromagnetycznym (siłą elektrostatyczną), elektronów to atom.

Atomy mogą łączyć się w cząsteczki, których względną trwałość zapewniają wiązania chemiczne.

Atomy mogą łączyć się w cząsteczki, których względną trwałość zapewniają wiązania chemiczne. Wiązania chemiczne powstają dzięki wymianie elektronów między atomami, która może odbywać się na dwa sposoby:

- kowalencyjny – polegający na uwspólnianiu par elektronów przez dwa lub więcej atomów,
- jonowy – polegający na trwałym przeniesieniu elektronów z jednego atomu na drugi, w którego wyniku na jednym z atomów tworzy się całkowity ładunek ujemny, a na drugim dodatni; w efekcie powstaje para jonowa, która jest związana z sobą zwykłymi oddziaływaniami elektrostatycznymi.

1.11 Dualizm korpuskularno-falowy i jego eksperymentalne potwierdzenie.

Dualizm korpuskularno-falowy – cecha obiektów kwantowych (np. fotonów czy elektronów) polegająca na przejawianiu, w zależności od sytuacji, właściwości falowych (dyfrakcja, interferencja) lub korpuskularnych (dobrze określona lokalizacja, pęd).

Zgodnie z mechaniką kwantową cała materia charakteryzuje się takim dualizmem, chociaż uwidacznia się on bezpośrednio tylko w bardzo subtelnych eksperymentach wykonywanych na atomach, fotonach, czy innych obiektach kwantowych.

Dualizm korpuskularno-falowy jest ściśle związany z falami de Broglie'a – koncepcją, która przyczyniła się do powstania mechaniki kwantowej, a w szczególności do wyprowadzenia równania Schrödingera: $\lambda = \frac{h}{p}$, gdzie h jest stałą Plancka, łączącą wielkości falowe (długość fali λ) z korpuskularnymi (pęd p).

Dokonując pomiaru położenia cząstki zawsze znajdujemy ją w przybliżeniu w konkretnym miejscu w przestrzeni (rejestruje ją konkretny detektor). W przypadku eksperymentów z podwójną szczeliną uzyskuje się interferencję bądź nie w zależności od tego czy obiekt przejawia właściwości falowe czy cząsteczkowe. Właściwości cząsteczkowe są obserwowane, gdy w szczelinach będzie umieszczony detektor, wykrywający przez którą szczelinę się poruszał obiekt. Przyczyną tego jest istnienie splątania kwantowego i dostępność informacji o obserwacjach. Po detekcji cząstki nieoznaczoność jej pędu stopniowo wzrasta, przez co maleje widoczność prążków interferencyjnych.

Young w 1801 r. pierwszy eksperyment przeprowadził ze światłem. Następnie przeprowadzano eksperymenty z elektronami, atomami oraz cząsteczkami. Największe układy, dla których zaobserwowano dualizm korpuskularno-falowy miały 58 (ftalocyjanina), 114 atomów (pochodna ftalocjaniny) i nawet po kilkaset atomów.

1.12 Hamiltonian w mechanice klasycznej i kwantowej.

Hamiltonian (funkcja Hamiltona) – funkcja współrzędnych uogólnionych i pędów uogólnionych, opisująca układ fizyczny:

$H = H(q_1, \dots, q_N, p_1, \dots, p_N, t)$, gdzie:

q_j - współrzędne uogólnione,

p_j - pędy uogólnione,

N - liczba stopni swobody,

t - czas.

Funkcję Hamiltona otrzymuje się z wyrażenia na energię całkowitą układu (przy czym prędkości wyraża się za pomocą pędów) np:

- Hamiltonian punktu materialnego poruszającego się z prędkością nie-relatywistyczną w potencjale V :

$$H(t, \mathbf{q}, \mathbf{p}) = \frac{\mathbf{p}^2}{2m} + V(\mathbf{q})$$

- Hamiltonian oscylatora harmonicznego poruszającego się w kierunku x :

$$H(x, p) = \frac{p^2}{2m} + \frac{m\omega_0^2 x^2}{2}$$

Funkcję Hamiltona można też otrzymać z funkcji Lagrange'a (za pomocą tzw. transformacji Legendre'a). Funkcja Lagrange'a:

$$L(q_1(t), \dots, q_n(t), \dot{q}_1(t), \dots, \dot{q}_n(t), t) = T(q_1(t), \dots, q_n(t), \dot{q}_1(t), \dots, \dot{q}_n(t), t) - U(q_1(t), \dots, q_n(t), \dot{q}_1(t), \dots, \dot{q}_n(t), t),$$

gdzie:

T - energia kinetyczna,

U - uogólniona energia potencjalna.

q_j - współrzędna uogólniona,

\dot{q}_j - prędkość uogólniona.

Dla każdej prędkości uogólnionej \dot{q}_j wyznacza się odpowiadający jej pęd uogólniony p_j , zdefiniowany jako pochodna funkcji Lagrange'a po prędkości uogólnionej \dot{q}_j :

$$p_j = \frac{\partial L}{\partial \dot{q}_j}$$

Hamiltonian można znaleźć teraz z funkcji Lagrange'a za pomocą tzw. transformacji Legendre'a:

$$H = H(q_1, \dots, q_N, p_1, \dots, p_N, t) = \sum_i \dot{q}_i p_i - L(q_1(t), \dots, q_n(t), \dot{q}_1(t), \dots, \dot{q}_n(t), t)$$

przy czym konieczne jest wyrażenie prędkości uogólnionych występujących w funkcji Lagrange'a przez pędy uogólnione, gdyż funkcja Hamiltona musi być zapisana jako funkcja pędów uogólnionych. Nie dla wszystkich układów taka transformacja jest możliwa. W przypadku współrzędnych kartezjańskich pędy uogólnione są zwykłymi pędami. W współrzędnych walcowych jako jedną ze współrzędnych uogólnionych cząstki przyjmuje się kąt; wtedy prędkość uogólniona jest prędkością kątową, a pęd uogólniony – obliczany jako pochodna funkcji Lagrange'a po prędkości kątowej – okazuje się być momentem pędu cząstki. W ogólnym przypadku pędy uogólnione mogą nie posiadać prostej interpretacji fizycznej, co wynika z dowolności wyboru współrzędnych uogólnionych.

Operator Hamiltona \hat{H} - operator definiowany w mechanice kwantowej, będący odpowiednikiem funkcji Hamiltona H (hamiltonianu) mechaniki klasycznej. Operator Hamiltona działa na wektory stanu układu kwantowego:

$$\hat{H}|\Psi(t)\rangle = i\hbar \frac{\partial}{\partial t} |\Psi(t)\rangle$$

tworząc zależne od czasu równanie Schrodingera. Operator Hamiltona jest jedną z obserwabli, jakie wprowadza mechanika kwantowa, czyli operatorem takim, że jego wartości własne są wielkościami, które można otrzymać w eksperymencie. Wartości własne operatora Hamiltona przedstawiają wartości energii, jakie układ kwantowy może posiadać. Ponieważ energie wyraża się za pomocą liczb rzeczywistych, to implikuje, że operator Hamiltona musi być operatorem hermitowskim (operatorem samosprzężonym) $\hat{H} = \hat{H}^\dagger$. Jest tak dlatego, że tylko operator hermitowski ma wartości własne będące zawsze liczbami rzeczywistymi.

Aby uzyskać postać operatora Hamiltona dla danego układu, w klasycznej funkcji Hamiltona zamienia się współrzędne uogólnione i pędy na odpowiadające im operatory. Najprościej dokonuje się tego, gdy współrzędnymi

uogólnionymi są współrzędne kartezjańskie x_i , które pozostawia się bez zmian, a np. w przypadku cząstki swobodnej pędом przypisuje się operatory różniczkowania po sprzężonych z nimi współrzędnych:

$$p_i \rightarrow \hat{p}_i = -i\hbar \frac{\partial}{\partial x_i}$$

Operator hamiltona dla cząstki swobodnej w polu $V(x, t)$ będzie miał więc postać:

$$H(x, p_x, t) = \frac{\hat{p}_x^2}{2m} + V(x, t) = \frac{-\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x, t)$$

Jeżeli klasyczna funkcja Hamiltona opisująca dany układ jest niezależna od czasu, to także operator Hamiltona \hat{H} nie zależy od czasu. Wtedy zamiast ogólnego równania Schrödingera wystarczy rozwiązać równania Schrödingera niezależne od czasu, które jest równaniem własnym hamiltonianu:

$$\hat{H}|\psi_E\rangle = E|\psi_E\rangle, \text{ gdzie:}$$

E - wartości własne operatora Hamiltona - wartości te są wartościami energii, jakie układ może posiadać,

$|\psi_E\rangle$ - stany własne operatora Hamiltona o energiach E .

1.13 Równanie Schrödingera zależne i niezależne od czasu.

Równanie Schrödingera – jedno z podstawowych równań nierelatywistycznej mechaniki kwantowej opisujące funkcję falową albo stan układu kwantowego.

Najbardziej ogólna forma zależnego od czasu równania Schrödingera:

$$i\hbar \frac{d}{dt}|\Psi(t)\rangle = \hat{H}|\Psi(t)\rangle, \text{ gdzie:}$$

i - jednostka urojona,

$\hbar = h/2\pi$, h - stała Plancka,

\hat{H} - operator Hamiltona,

$|\Psi(t)\rangle$ - wektor stanu układu.

Aby rozwiązać równanie Schrödingera dla danego układu kwantowego, należy znaleźć właściwą postać operatora Hamiltona oraz wyrazić wektor stanu w odpowiedniej reprezentacji, np. dla pojedynczej cząstki w jednym wymiarze x , w potencjale $V(x)$ hamiltonian wygląda następująco:

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(x) = \frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x)$$

Gdy potencjał nie zależy od czasu, funkcję falową możemy zapisać jako iloczyn części przestrzennej i czasowej:

$\Psi(x, t) = \psi(x)\psi(t)$ otrzymując następującą formę równania Schrödingera:

$$i\hbar \frac{d(\psi(x)\psi(t))}{dt} = \frac{\hbar^2}{2m} \frac{d^2(\psi(x)\psi(t))}{dx^2} + V(x)\psi(x)\psi(t), \text{ następnie odseparo-}$$

wywołujemy zmienne i wiedząc, że lewa strona jest tylko zależna od czasu, a prawa tylko od położenia są one równe stałej (na podstawie prawej strony równania stwierdzamy, że stała ta ma wymiar energii):

$$i\hbar \frac{1}{\psi(t)} \frac{d\psi(t)}{dt} = \frac{1}{\psi(x)} \left[\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x) \right] \psi(x) = E$$

Z powyższego równania otrzymujemy dwie zależności:

$\Psi(x, t) = \psi(x)e^{-iEt/\hbar}$ - całkowita funkcja falowa różni się od części przestrzennej tylko czynnikiem fazowym,

$\frac{\hbar^2}{2m} \frac{d^2\psi(x)}{dx^2} + V(x)\psi(x) = E\psi(x)$ - niezależne od czasu równanie Schrödingera dla przypadku jednowymiarowego cząstki w potencjale $V(x)$.

Funkcje $\psi(x)$ spełniające powyższe równanie przy zadanym potencjale $V(x)$ nazywamy funkcjami własnymi, a wartości energii, dla których istnieją te rozwiązania, nazywamy wartościami własnymi.

Funkcja własna jak i jej pochodna musi być skończona, jednoznaczna i ciągła. Gęstość prawdopodobieństwa znalezienia cząstki opisuje wzór:

$$P(x, t) = \Psi^*(x, t)\Psi(x, t) = \psi^*(x)\psi(x) = P(x),$$

Warunek normalizacji dla gęstości prawdopodobieństwa:

$$\int_{-\infty}^{\infty} \Psi^*(x, t)\Psi(x, t) = 1$$

Równanie Schrödingera jest równaniem liniowym, co oznacza, że jeśli $\Psi_1(x, t)$, $\Psi_2(x, t)$ są rozwiązaniami tego równania, to również liniowa kombinacja tych funkcji jest rozwiązaniem tego równania. Uogólniając:

$\Psi(x, t) = \sum_N a_N \Psi_N(x, t)$ jest rozwiązaniem równania Schrödingera przy dowolnych a_N gdy Ψ_1, \dots, Ψ_N są rozwiązaniami tego równania.

1.14 Podstawy formalizmu kwantowego – wielkości fizyczne, stany, operatory.

Stan układu kwantowego reprezentowany jest przez wektor stanu w przestrzeni Hilberta (przestrzeń wektorowa nad ciałem liczb zespolonych). Do oznaczania wektorów stanu stosuje się notację Diraca. Wektor stanu oznacza się przez ket $|\Psi\rangle$, a wektor dualny (sprzężony) do niego przez $\langle\Psi|$.

Każda wielkość fizyczna (obserwable) reprezentowana jest przez hermitowski (lub samosprzężony) operator liniowy działający w przestrzeni stanów (przestrzeni Hilberta). Zbiór wartości własnych tego operatora, nazywany widmem punktowym operatora, interpretuje się jako zbiór możliwych wartości obserwowalnych (pomiarowych). Dla hermitowskich operatorów wartości w widmie są liczbami rzeczywistymi. Stany własne (wektory w łasne) tego operatora do tych wartości własnych interpretuje się jako możliwe stany, w których znajdzie się układ po dokonaniu pomiaru.

Wektor stanu można zapisać w postaci poniższej sumy:

$$|\Psi\rangle = \sum_i |i\rangle\langle i|\Psi\rangle$$

Prawdopodobieństwa otrzymania poszczególnych wartości pomiarów:

$$P_i = |\langle i|\Psi\rangle|^2$$

Operator pędu w reprezentacji położeniowej można otrzymać, używając rozwiązania fali płaskiej do równania Schrödingera pojedynczej cząstki:

$$\Psi(x, t) = e^{\frac{i}{\hbar}(px - Et)}, \text{ gdzie:}$$

p - pęd cząstki w kierunku x ,

E - energia cząstki.

Pochodna cząstkowa 1-go rzędu wzdłuż x wynosi:

$$\frac{\partial \Psi(x, t)}{\partial x} = \frac{ip}{\hbar} e^{\frac{i}{\hbar}(px - Et)}, \text{ stąd:}$$

$$\hat{p} = -i\hbar \frac{\partial}{\partial x}$$

Ważną cechą kwantowego operatora położenia jest to, że nie komutuje on z operatorem pędu (są to operatory kanonicznie sprzężone). Operatory te spełniają relację komutacyjną:

$$[x_i, p_j] = i\hbar\delta_{ij}, \text{ gdzie:}$$

$$[\mathbf{A}, \mathbf{B}] = \mathbf{AB} - \mathbf{BA}$$

Powyższa zależność jest matematycznym zapisem zasady nieoznaczoności, która można zapisać również jako:

$$\Delta x \Delta p_x \geq \frac{\hbar}{2}, \text{ gdzie:}$$

$\Delta x, \Delta p_x$ - niepewności wynikające z istoty samego pomiaru (wpływu pomiaru na obiekt mierzony)

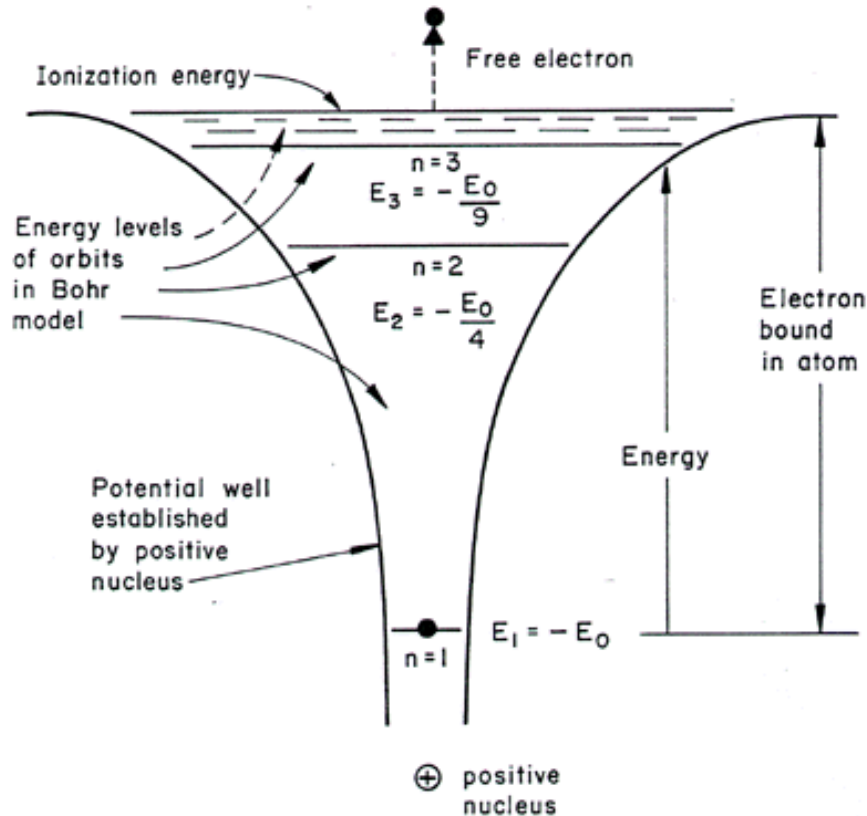
1.15 Kwantowy opis atomu wodoru. Liczby kwantowe.

Energię potencjalną będącą wynikiem oddziaływania kulombowskiego pomiędzy pojedynczym elektronem i protonem w atomie wodoru opisuje poniższy wzór:

$$U(r) = \frac{e^2}{4\pi\epsilon_0 r}, \text{ gdzie:}$$

r - odległość elektronu od protonu.

Energia potencjalna jest funkcją odległości r od jądra. Jest to pole centralne.



Rysunek 8: Energia potencjalna oddziaływania kulombowskiego pomiędzy elektronem, a jądrem w atomie wodoru.

Równanie Schrodingera dla takiego potencjału ma postać:

$$-\frac{\hbar}{2m_e} \nabla^2 \Psi(\vec{r}) = (E - U(\vec{r})) \Psi(\vec{r})$$

Równanie to można rozwiązać przechodząc do układu współrzędnych sferycznych, a następnie sprowadzając je do układu równań jednowymiarowych stosując separację zmiennych:

$$\Psi(\vec{r}) = R(r)\Theta(\theta)\phi(\varphi)$$

Rozwiązanie dla funkcji $\phi(\varphi)$:

$$\phi(\varphi) = e^{im_l\varphi} \text{ implikuje:}$$

$$e^{im_l0} = e^{im_l2\pi}, \text{ warunek ten jest spełniony gdy:}$$

$$|m_l| = 0, 1, 2, 3, \dots$$

Liczba kwantowa m_l , nazywana magnetyczną liczbą kwantową, jest związana z orientacją w przestrzeni wektora momentu pędu. Jeżeli atom znajduje się w zewnętrznym polu magnetycznym, to jego energia zależy od tej liczby kwantowej. Liczba kwantowa m_l może być tylko liczbą całkowitą.

Rozwiązanie dla funkcji $\Theta(\theta)$ implikuje istnienie liczby kwantowej $l = |m_l|, |m_l| + 1, |m_l| + 2 \dots$ nazywanej orbitalną liczbą kwantową, będącą miarą wielkości momentu pędu związanego ze stanem kwantowym:

$L = \hbar \sqrt{l(l+1)}$ - orbitalny moment pędu.

$L_z = \hbar m_l$ - rzut wektora momentu pędu na oś OZ.

Obie powyższe wielkości są skwantowane.

Dostępne stany energii opisuje równanie:

$$E = -\frac{1}{2} m_e c^2 (Z\alpha)^2 \frac{1}{n_r + l + 1}, \text{ gdzie:}$$

Z - liczba atomowa (dla wodoru $Z = 1$),

$\alpha \approx 1/137$ - stała struktury subtelnej,

$n_l = 0, 1, 2 \dots$ - radialna liczba kwantowa,

$n = n_r + l + 1$ - główna liczba kwantowa (jest zawsze całkowitą liczbą dodatnią).

Jest to dokładnie ten sam wzór, który wynika z modelu Bohra dla energii dozwolonych stanów związanych.

Np. dla energii $E_0 = E(n = 1)$ mamy tylko jeden możliwy stan ($n = 1, l = 0, m_l = 0$). Energii $E_1 = E(n = 2)$ odpowiadają cztery zdegenerowane stany:

$(n = 2, l = 0, m_l = 0),$

$(n = 2, l = 1, m_l = +1),$

$(n = 2, l = 1, m_l = 0),$

$(n = 2, l = 1, m_l = -1).$

2 SPECJALNOŚĆ: EKSPLOACJA DANYCH I MODELOWANIE INTERDYSCYPLINARNE

2.1 Automaty komórkowe.

Automat komórkowy – system składający się z pojedynczych komórek, znajdujących się obok siebie. Ich układ przypomina szachownicę lub planszę do gry.

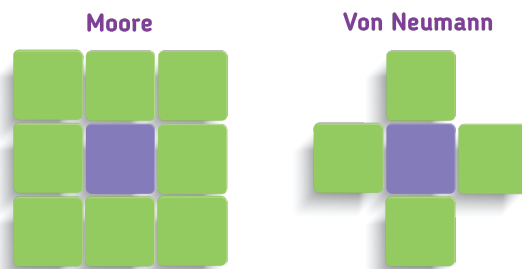
Każda z komórek może przyjąć jeden ze stanów, przy czym liczba stanów jest skończona, ale dowolnie duża.

Przestrzeń stanów określamy poprzez zdefiniowanie wartości wybranej ze skończonego zbioru Q , który może być podzbiorem zbiorów podstawowych (np. liczb, liter) lub złożonych (struktury, obiekty).

Poprzez przestrzeń stanów opisane zostają wszystkie możliwe stany komórki. Od dobrego określenia przestrzeni stanów zależy szybkość automatu komórkowego, jak i jego podstawowe cechy charakterystyczne. Stan komórki zależy od aktualnych stanów komórek z otoczenia, jak i komórka swoim stanem wpływa bezpośrednio na stany swoich sąsiadów.

Istnieją trzy konstrukcyjne czynniki automatu komórkowego, które w zasadniczy sposób wpływają na strukturę siatki:

- wymiar przestrzeni, zależny od wielkości badanego problemu (siatka 1D, 2D, 3D, nD),
- warunek regularności, mówiący o całkowitym wypełnieniu siatki przez jednakowe komórki (komórki trójkątne, kwadratowe, sześciokątne dla 2D, sześciennie lub w kształcie dwunastościanów rombów dla 3D etc.),
- liczba sąsiadów (zależna od obu powyższych, rysunek 9 przedstawia podstawowe typy sąsiedztwa).



Rysunek 9: Podstawowe typy sąsiedztwa.

Budowa wszystkich komórek musi być identyczna (muszą mieć tyle samo sąsiadów, takie same zbiory stanów itp.)

Ważnym elementem konstrukcyjnym automatów komórkowych są warunki początkowe, czyli stany poszczególnych komórek w zerowej iteracji, czyli na samym początku. To od ustawienia początkowego komórek zależy dalsza ewolucja automatu, jego zachowanie, stan końcowy, tym samym powodzenie całej symulacji. Niektóre automaty komórkowe z założenia muszą mieć w odpowiedni sposób ustalone warunki początkowe

Zachowanie automatu jest definiowane m.in. poprzez warunki brzegowe, które mogą być np. okresowe, pochłaniające, czy odbijające.

Stan komórki x w chwili t oznaczmy jako x_t , a stan sąsiedztwa jako $u(x_t)$. Dla takich kryteriów stan komórki x w kolejnym kroku iteracji można opisać wzorem:

$x_{t+1} = f(u(x_t), x_t)$, gdzie:

f - funkcja przejścia, która może być opisana różnego rodzaju zależnościami, np. jako tabela przejść, w postaci algorytmicznej lub jako zbiór reguł.

Po ustaleniu i zdefiniowaniu wszystkich elementów składowych automatu komórkowego można przejść do nakładania reguł na siatkę, czyli aktualizowania jej.

Podział automatów komórkowych według Wolframa:

- Klasa I: Automaty niezmiennie – ewoluują do czasu, kiedy wszystkie komórki osiągną identyczny stan niezależnie od stanu początkowego (zbieżne).
- Klasa II: Automaty ewoluujące do stanu stabilnego lub okresowych wzorców (okresowe).
- Klasa III: Automaty wykazujące nieporządek zarówno lokalnie, jak i globalnie, nie wykazujące żadnego wzorca (chaotyczne).
- Klasa IV: Automaty wykazujące bardziej złożone, długotrwałe zachowanie („żywe”).

Technika automatów komórkowych jest używana do symulacji komputerowych w wielu problemach nauki i techniki. Automaty te dostarczają również wielu zagadnień teorii dynamiki nieliniowej.

Jednym z pierwszych i najbardziej znanych przykładów automatu komórkowego jest *Gra w życie* (*Life, The game of life*). Reguły tego automatu według Conwaya:

- Martwa komórka, która ma dokładnie 3 żywych sąsiadów, staje się żywa w następnej jednostce czasu (rodzi się),

- Żywa komórka z 2 albo 3 żywymi sąsiadami pozostaje nadal żywa; przy innej liczbie sąsiadów umiera (z „samotności” albo „zatłoczenia”).

2.2 Podstawowe twierdzenie algorytmów genetycznych.

Algorytmy genetyczne stanowią podgrupę algorytmów ewolucyjnych. Są to algorytmy poszukiwania oparte na mechanizmach doboru naturalnego oraz dziedziczności. Oznacza to, iż wykorzystujemy ewolucję aby poprzez operatory (mutacje i krzyżowanie) otrzymać z niemal losowych danych to co będziemy uważać za poprawne. AG pracują na zakodowanej formie rozwiązań oraz na całej populacji rozwiązań. Bazują tylko na funkcji celu (nie musi być ona ciągła ani różniczkowalna, ani także być funkcją analityczną).

Schematem H nazywamy wzorec opisujący podzbiór ciągów podobnych ze względu na ustalone pozycje. Alfabet składający się z k symboli i ciągów l elementowych posiada k^l ciągów i $(k+1)^l$ schematów. Np. dla alfabetu dwuelementowego $\{0, 1\}$ alfabet schematu $V^+ = \{0, 1, *\}$, gdzie $*$ - oznacza pozycję nieustaloną.

Rząd schematu to liczba ustalonych pozycji we wzorcu $o(H)$ np:

$$o(\{01**01**\}) = 4$$

Rozpiętość schematu to odległość między dwiema skrajnymi ustalonymi pozycjami $\delta(H)$ np:

$$\delta(\{01**01**\}) = 5$$

$m(H, t)$ to liczba reprezentantów schematu H , czyli jak dużo osobników populacji pasuje do schematu H w czasie t .

Wartość oczekiwana liczby reprezentantów schematu H w następnym pokoleniu:

$$E(m(H, t+1)) = m(H, t) \cdot n \cdot \frac{f(H)}{\sum_i f_i} = m(H, t) \cdot \frac{f(H)}{\bar{f}}, \text{ gdzie:}$$

$f(H)$ - współczynnik dostosowania schematu H ,

$\bar{f} = \frac{\sum_i f_i}{n}$ - średni współczynnik dostosowania całej populacji.

Prawdopodobieństwo przeżycia schematu:

$$p_s(H) \geq \frac{\delta(H)}{l-1}$$

Oczekiwana liczba reprezentantów schematu H w następnym pokoleniu, otrzymanym w wyniku reprodukcji, krzyżowania i mutacji, spełnia następującą nierówność:

$$m(H, t+1) \geq m(H, t) \cdot \frac{f(H)}{\bar{f}} [1 - p_c \frac{\delta(H)}{l-1} - p_m o(H)], \text{ gdzie:}$$

p_c - prawdopodobieństwo krzyżowania,

p_m - prawdopodobieństwo mutacji.

Dla schematów niskiego rzędu i o małej rozpiętości destruktywne efekty krzyżowania i mutacji można pominąć, a wtedy:

$m(H, t+1) = m(H, t) \cdot (1 + \epsilon)$, przy założeniu, że $f(H, t) = (1 + \epsilon)\bar{f}$ czyli schemat H ma dopasowanie o ϵ większe niż średnia z populacji. Zakładając że ϵ nie zmienia się w czasie, wówczas długoterminowy efekt przetwarzania schematu H będzie następujący:

$m(H, t+s) = m(H, t)(1 + \epsilon)^s$ - oceniane "powyżej średniej" schematy o "niskiego rzędu" i o "małej rozpiętości" uzyskują wykładniczo rosnącą liczbę reprezentantów w kolejnych pokoleniach. Ta konkluzja jest nazywana podstawowym twierdzeniem algorytmów genetycznych.

2.3 Sieci bezskalowe.

Sieć bezskalowa to sieć o potęgowym rozkładzie liczby połączeń między węzłami:

$$P(k) = \frac{C}{k^\gamma}, \text{ gdzie:}$$

k - liczba węzłów,

γ - parametr właściwy dla danej sieci i zwykle przyjmującym wartości z zakresu $(2, 3)$.

Potęgowy rozkład stopni wierzchołków jest cechą charakterystyczną wielu sieci rzeczywistych (np. sieć energetyczna, sieć www). Rozkład potęgowy nie ma naturalnej skali, jest bezskalowy. Oznacza to, że mówienie o średnich wartościach stopni wierzchołków w takich sieciach jest co najmniej niewskazane, a wielu wypadkach posługiwanie się ideą średniego stopnia prowadzi do poważnych błędów. W sieci o potęgowym rozkładzie stopni wierzchołków wiele węzłów ma tylko jedną krawędź (tłusty ogon), ale można też znaleźć węzły z ogromną liczbą krawędzi, tzw. huby.

Nie wszystkie tłustoogoniaste rozkłady są potęgowe. Zlogarytmizujemy obie strony równania opisujuącego rozkład potęgowy:

$$\ln(P(k)) = -\gamma \ln(k) + \ln(C)$$

Rozkład potęgowy charakteryzuje się tym, że przedstawiony w nowych współrzędnych tj. $X = \ln(k)$, $Y = \ln(P(k))$ staje się równoważny liniowej zależności $Y = aX + b$.

Sieci bezskalowe charakteryzują się łatwością w niszczeniu topologii sieci, poprzez usunięcie wierzchołków o dużej liczbie połączeń tzw. "hub-ów", które są krytyczne dla łączności sieci np. awaria istotnego węzła sieci energetycznej, albo atak hakerski na ruchliwy serwer.

Sieci bezskalowe mogą być generowane zgodnie z modelem Barabasi-Alberta. Proces konstrukcji takiej sieci wygląda następująco:

- Rozpocznij proces od pewnego niewielkiego zbioru wierzchołków (zazwyczaj jednego).
- W każdym kroku dodaj jeden wierzchołek i dołącz do już istniejących dając preferencje tym którzy już mają dobrą łączność (prawdopodobieństwo dodania krawędzi do wierzchołka v jest proporcjonalne do aktualnego stopnia d_v - wiązanie preferencyjne). Ilość dodawanych w każdym kroku krawędzi jest parametrem modelu m .

Powstałe w ten sposób sieci charakteryzują się rozkładem krawędzi:

$$P(k) \approx \frac{2m^2}{k^3}.$$

Prawo potęgowe wraz z wykładnikiem są niezmiennicze dla losowego podgrafu (z dużym prawdopodobieństwem). Tutaj pojawia się wyjaśnienie terminu "bezskalowy".

Sieci małych światów:

W eksperymencie przeprowadzonym w 1967 roku amerykański socjolog Stanley Milgram wysłał do przypadkowo wybranych ludzi 160 listów zawierających wyjaśnienie eksperymentu, zdjęcie, nazwisko i adres pewnej osoby oraz instrukcję postępowania. Jeżeli adresat znał osobiście człowieka wymienionego w liście, miał przesłać list bezpośrednio do niego. W przeciwnym wypadku list powinien zostać przesłany do innego znajomego, o którym adresat mógł sądzić, że może znać poszukiwaną osobę lub przynajmniej znać kogoś, kto tę osobę zna osobiście. Celem eksperymentu było ustalenie, jak długi jest łańcuch znajomych gwarantujący dostarczenie przesyłki do adresata. Choć większość listów zaginęła, to jednak te 42, które dotarły do poszukiwanej osoby, dostarczyły zaskakujących wyników. Okazało się, że paru listom

wystarczyło zaledwie dwóch pośredników by dotrzeć do celu. W kilku innych przypadkach pośredników było kilkunastu. Jednak, po uśrednieniu wyników, okazało się, że statystyczny list przeszedł przez ręce jedynie sześciu osób. Eksperyment Milgrama dowiódł prawdziwości obiegowego porządku (znanego również Barneyowi), że świat jest mały. Mimo że sieć społeczna liczy kilka miliardów ludzi, to średnia droga między dowolną parą węzłów w takiej sieci wynosi około sześciu. Ponad dwadzieścia lat po eksperymencie Milgrama ukute zostało nawet sformułowanie „sześć uścisków dłoni” albo „sześć stopni separacji”. Źródło: (<http://www.if.pw.edu.pl/~agatka/moodle/charakterystyki.html#x1-2002r1>)

2.4 Agentowe modele dynamiki opinii.

Model wpływu społecznego

Wprowadźmy następujące oznaczenia:

- $\sigma_i(t)$ - opinia i -tego agenta, $\sigma_i(t) \in \{-1, 1\}$
- $I_i(t)$ - "pole społeczne" oddziałujące na i -tego agenta (wynika ono z obecności innych agentów)
- S_i - możliwość popierania (podtrzymania opinii) i -tego agenta, $S_i > 0$
- \prod_i - umiejętność perwazji (przekonania do swojej opinii) i -tego agenta, $\prod_i > 0$
- d_{ij} - odległość społeczna między i -tym, a j -tym agentem
- $g(d_{ij})$ - osłabienie oddziaływań pomiędzy agentami w zależności od ich odległości społecznej, zazwyczaj $g(d_{ij}) = A \cdot d_{ij}^\alpha$, $A, \alpha > 0$, $g(d_{ij}) = \frac{1}{\beta}$
- β - parametr przekonania do własnej opinii

Poniższy wzór opisuje dynamikę modelu:

$\sigma_i(t + \Delta t) = -\text{sign}[\sigma_i(t)I_i(t)]$, gdzie:

$$I_i(t) = \sum_j \frac{-[\sigma_i(t)\sigma_j(t) + 1]S_j}{2g(d_{ij})} + \sum_j \frac{[\sigma_i(t)\sigma_j(t) - 1]\prod_j}{2g(d_{ij})} + \eta_i(t), \text{ gdzie:}$$

$\eta_i(t)$ – czynnik temperaturowy (szum losowy) o właściwościach $\langle \eta_i(t) \rangle = 0$, $\langle (\eta_i(t))^2 \rangle = T$. Kiedy stworzymy dwie sieci (grupy) agentów, o różnych opiniach, w których $T_1 > T_2$ i porównamy je ze sobą to grupa o wyższej temperaturze jest bardziej podatna na zmianę opinii.

Wybrany agent zmieni swoją opinię na przeciwną, kiedy pole społeczne przyjmie wartość dodatnią. Innymi słowy, jego opinia ulegnie zmianie, wtedy gdy presja otoczenia będzie wystarczająco duża.

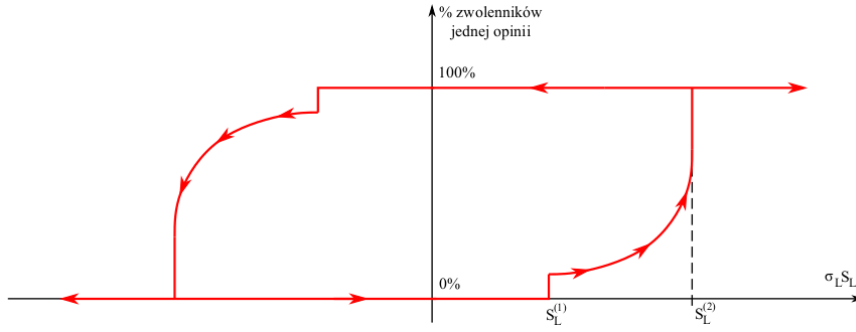
Model silnego lidera

Zakładamy poniższe wartości parametrów:

$$\forall i \neq L, S_i = \prod_i = 1, S_L = \prod_L \gg 1$$

Otrzymujemy poniższe równanie na "pole społeczne":

$$I_i(t) = \sum_{i \neq L} \frac{-\sigma_i(t)\sigma_j(t)}{g(d_{ij})} - \frac{\sigma_i(t)\sigma_L(t)}{g(d_{iL})} S_L$$



Rysunek 10: Proces zwolenników jednej opinii w funkcji iloczynu siły perswazji lidera i jego opinii. Wykres ten przypomina pętle histerezy, obserwowaną m.in. w zjawiskach magnetycznych.

Dla modelu silnego lidera istnieje wartość krytyczna S_L po której wszyscy agenci w grupie przyjmują wspólną opinię (rys. 10).

Model votera (wyborcy) - agenci umiejscowieni są w sieci, w której możemy wyróżnić najbliższe sąsiedztwo każdego z nich. Każdy agent posiada pewną opinię (stan) $s_i(t) \in -1, 1$. Dynamika przebiega w sposób następujący sposób:

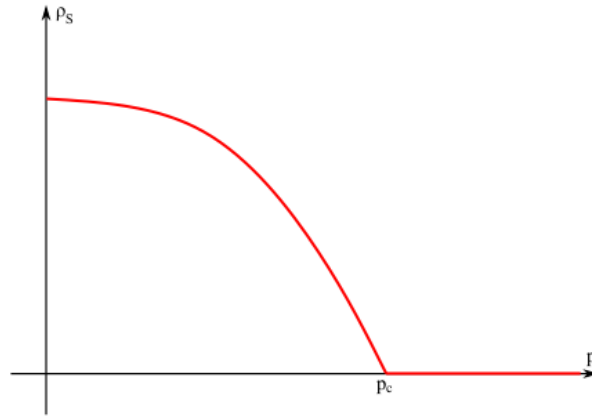
- Wylosuj agenta i ,
- Wylosuj agenta j będącego sąsiadem agenta i ,
- Agent i przyjmuje opinię agenta j : $s_i(t + \Delta t) = s_j(t)$.

W przypadku kiedy sieć jest prosta (tzn. każdy agent ma jednakową liczbę sąsiadów) wtedy układ osiąga stan równowagi $\langle s_i \rangle = const$. Innymi słowy liczba aktywnych interfejsów $\rho(t)$ (połączeń agentów o przeciwnych opiniach) zanika w czasie.

Model votera można wzbogacić o koewolucję (ewoluują zarówno stany węzłów jak i topologia samej sieci). Dynamika koewolucji:

- Z prawdopodobieństwem $1 - p$ agent i przyjmuje stan agenta j , tak jak w zwykłym modelu votera,
- z prawdopodobieństwem p agent i zrywa połączenie z agentem j i nawiązuje połączenie z losowym agentem, który nie jest jego sąsiadem i posiada taką samą opinię.

Założmy, że po bardzo długim czasie ($t \rightarrow \infty$) średnia liczba linków aktywnych wynosi $\rho_s = const$. Istnieje wtedy pewne prawdopodobieństwo krytyczne $p = p_c$, dla którego $\rho_s = 0$ (rys. 11).



Rysunek 11: Wykres zależności średniej gęstości aktywnych połączeń po bardzo długim czasie w funkcji prawdopodobieństwa p .

Model Axelroda - każdemu agentowi przypisujemy wektor cech (np. wyznawana religia, poglądy polityczne) o długości F , z których każda może posiadać q wariantów. Przebieg dynamiki:

- wybieramy losowo agenta i ,
- wybieramy losowo agenta j z najbliższego sąsiedztwa agenta i
- obliczamy liczbę wspólnych cech agentów i, j :

$$l(i, j) = \sum_{f=1}^F \delta_{\sigma_{i,f}, \sigma_{j,f}},$$
gdzie $\sigma_{i,f}$ - wariant cechy f , agenta i
- Spośród wszystkich cech, takich, że $\sigma_{i,f} \neq \sigma_{j,f}$ wybieramy cechę k i zmieniamy opinię agenta i : $\sigma_{i,k}(t + \Delta t) = \sigma_{j,k}(t)$ z prawdopodobieństwem $p = \frac{l(i, j)}{F}$

Model większościowy - zawiera N agentów, z których każdy ma opinię ± 1 . W dowolnej chwili czasu wybieramy m połączonych ze sobą agentów (zwykle losuje się jednego agenta oraz jego $m - 1$ różnych sąsiadów). Każdy agent z tak wylosowanej grupy przyjmuje opinię większości. Prawdopodobieństwo wylosowania agenta z opinią $+1$:

$$a(t) = \frac{N_a(t)}{N}, \text{ gdzie } N_a(t) - \text{liczba zwolenników opinii } +1 \text{ w chwili } t.$$

W najprostszym przypadku zakładamy, że wszyscy agenci są połączeni oraz $m = 3$. Wtedy:

$$N_a(t + \Delta t) = N_a(t) + a(t)^2 \cdot [1 - a(t)] \cdot (+1) + a(t) \cdot [1 - a(t)]^2 \cdot (-1),$$

$$a(t + \Delta t) - a(t) = \frac{1}{N} a(t) [1 - a(t)] [2a(t) - 1].$$

Jeśli za Δt przyjmiemy N wtedy otrzymamy poniższe równanie różniczkowe:

$$\dot{a} = a(t + \Delta t) - a(t) = a(1 - a)(2a - 1) = f(a)$$

Punkty stałe (miejsca zerowe funkcji $f(a)$) tego dynamicznego układu:

$a_- = 0, a_0 = 0.5, a_+ = 1$ Stabilność otrzymanych trzech punktów stałych badamy obliczając wykładnik Liapunowa $\lambda(a)$ równy pochodnej $f'(a) = -6a^2 + 6a - 1$ w tych punktach:

$$\lambda(a_-) = -1 < 0 - \text{punkt stabilny},$$

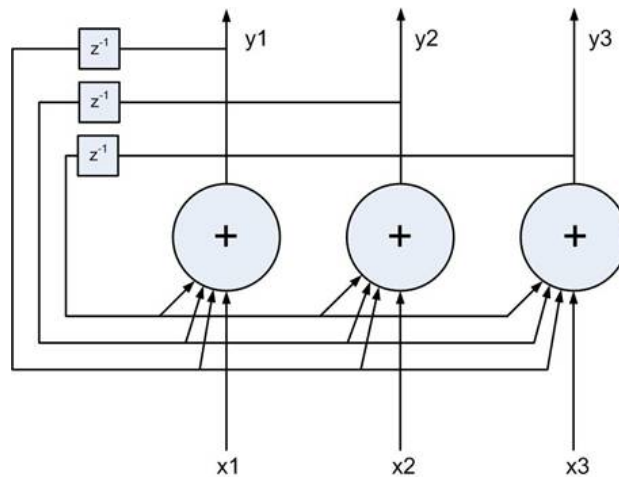
$$\lambda(a_+) = -1 < 0 - \text{punkt stabilny},$$

$$\lambda(a_0) = 1/2 > 0 - \text{punkt niestabilny}.$$

2.5 Sieć neuronowa typu Hopfielda.

Sieć Hopfielda jest przykładem sieci ze sprzężeniem zwrotnym (tzw. sieć rekurencyjna), gdzie wyjścia poszczególnych neuronów są podawane z odpowiednimi wagami na wejścia każdego z neuronów. Oczywiście do każdego neuronu podpięte jest wejście, którym doprowadzana jest odpowiednia składowa wektora testowego. Założenia:

- całkowite połączenie neuronów,
- neurony 2-stanowe,
- symetryczne połączenia synaptyczne
- synchroniczna dynamika sieci



Rysunek 12: Równania Maxwella.

W sieciach Hopfielda waga połączenia wyjścia neuronu z własnym wejściem jest zerowana. Wymagane jest także, aby macierz wag W była symetryczna.

Funkcja aktywacji pojedynczego neuronu wygląda następująco:

$$y_i(t+1) = 1, \text{ gdy } \sum_{j=1}^N w_{ij}y_j(t) + x_i(t) + \xi_i(t) > 0,$$

$$y_i(t+1) = y_i(t), \text{ gdy } \sum_{j=1}^N w_{ij}y_j(t) + x_i(t) + \xi_i(t) = 0,$$

$$y_i(t+1) = -1, \text{ gdy } \sum_{j=1}^N w_{ij}y_j(t) + x_i(t) + \xi_i(t) < 0, \text{ gdzie:}$$

i oraz j oznaczają numery neuronów N -neuronowej sieci Hopfielda, ξ oznacza szum (np. gaussowski o $\mu = 0$), natomiast t przedstawia chwilę czasową. Należy również zwrócić uwagę na to, że składowe wektora x w chwili czasowej $t = 0$ są przepisywane na wyjścia neuronów, natomiast dla $t > 0$ wejścia są „odpinane” od neuronów ($x = 0$).

W trybie uczenia, na podstawie zbioru wektorów uczących obliczane są wagi. Najprostszą metodą (wykorzystywaną w naszym przykładzie) uczenia sieci jest reguła Hebb’a. Wagi obliczane są według wzoru:

$$w_{ij} = \frac{1}{N} \sum_{k=1}^K x_i^{(k)} x_j^{(k)}, \text{ gdzie } k \text{ oznacza numer wektora uczącego, a } K \text{ liczbę wszystkich wektorów uczących.}$$

Okazuje się, że gdy na wejście sieci podamy wektor identyczny z wzorcem, wówczas sieć nie zmieni swojego stanu, rozpoznaje ona również obrazy „niewiele” różniące się od wzorców. W niektórych przypadkach pojawia się kolejna cecha sieci Hopfielda - pamiętanie zależności pomiędzy

sąsiednimi pikselami, a nie ich wartości. Wtedy otrzymujemy obraz wzorca z odwróconymi kolorami.

Miara podobieństwa wzorca k -tego (pokrycie ang. *overlap*):

$$m^k(t) = \frac{1}{N} \sum_{i=1}^N y_i(t) x_i^{(k)}.$$

Gdy:

$m^k = 1$ - całkowita zgodność wzorca,

$m^k = 0$ - całkowita niezgodność wzorca,

$m^k = -1$ - stan sieci jest negatywnym wzorca.

Sieć Hopfielda jest ograniczona ilością wzorców, które może zapamiętać i skutecznie rozróżniać. Jeśli p - ilość zapamiętanych wzorców, a $N \rightarrow \infty$ - liczba neuronów, to $p_c = \frac{p}{N} \approx 0,138$, przy założeniu, że średnie pokrycie wzorców nie spada poniżej 95%.

Liczbę identycznych bitów w dwóch stanach sieci nazywamy odległością Hamminga. Proces ewaluacji wzorca wejściowego zatrzymujemy, gdy wzorec wyjściowy zbliży się na pewną, ustaloną odległość Hamminga do któregoś z zapamiętanych wzorców.

2.6 Naiwny klasyfikator Bayesa.

Naiwny klasyfikator Bayesa - klasyfikator oparty na założeniu o wzajemnej niezależności zmiennych objaśniających (predyktorów). Często nie mają one żadnego związku z rzeczywistością i właśnie z tego powodu nazywa się je naiwnymi. Bardziej opisowe jest określenie - „model cech niezależnych”. Model prawdopodobieństwa można wyprowadzić korzystając z twierdzenia Bayesa wiążącego prawdopodobieństwa warunkowe dwóch zdarzeń warunkujących się nawzajem:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}, \text{ gdzie:}$$

A i B - zdarzenia,

$P(A|B)$ - prawdopodobieństwo warunkowe, tj. prawdopodobieństwo zajścia zdarzenia A , o ile zajdzie zdarzenie B .

Przykład zastosowania twierdzenia Bayesa:

B - zdarzenie "u pacjenta występuje wysoka gorączka" (w badanej populacji $P(B) = 0.2$),

A - zdarzenie "pacjent ma grypę" (w badanej populacji $P(A) = 0.1$),

$P(B|A) = 0.7$ - na podstawie statystyki populacji wiemy, że jeśli pacjent ma grypę to na 70% ma gorączkę.

Na podstawie twierdzenia Bayesa możemy obliczyć prawdopodobieństwo, że pacjent mający gorączkę, jest chory na grypę:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} = \frac{0,7 \cdot 0,1}{0,2} = 0,35$$

Korzystając z twierdzenia Bayesa:

$$p(C|F_1, \dots, F_n) = \frac{P(C)P(F_1, \dots, F_n|C)}{P(F_1, \dots, F_n)}, \text{ gdzie:}$$

C - klasa zależna od n opisujących zmiennych niezależnych F_i .

W praktyce interesujący jest tylko licznik ułamka, bo mianownik nie zależy od C . Licznik ułamka jest równoważny do łącznego modelu prawdopodobieństwa:

$$\begin{aligned} &P(C, F_1, \dots, F_n) \\ &= P(C)P(F_1, \dots, F_n|C) \\ &= P(C)P(F_1|C)P(F_2, \dots, F_n|C, F_1) \\ &= P(C)P(F_1|C)P(F_2|C, F_1)P(F_3, \dots, F_n|C, F_1, F_2) \end{aligned}$$

$$= P(C)P(F_1|C)P(F_2|C, F_1)P(F_3|C, F_1, F_2)P(F_4, \dots, F_n|C, F_1, F_2, F_3) \\ \text{itd...}$$

Włącza się teraz "naiwną" warunkową zależność, zakładając, że każda cecha F_i jest warunkowo niezależna od każdej innej cechy. Oznacza to, że:

$p(F_i|C, F_j) = P(F_i|C)$ więc model można wyrazić jako:

$$P(C, F_1, \dots, F_n) = p(C)p(F_1|C)p(F_2|C)\dots p(F_n|C) = p(C) \prod_{i=1}^n p(F_i|C)$$

Oznacza to, że pod powyższymi niezależnymi założeniami, warunkowe rozmieszczenie nad klasą zmiennych C można zapisać jako:

$$p(C|F_1, \dots, F_n) = \frac{1}{Z} p(C) \prod_{i=1}^n p(F_i|C), \text{ gdzie:}$$

Z - współczynnik skalowania zależny wyłącznie od zmiennych F_1, \dots, F_n .

Modele tej formy są łatwiejsze do zrealizowania, gdy rozłoży się je na czynniki zwane klasą „prior” $p(C)$ i niezależny rozkład prawdopodobieństwa $p(F_i|C)$. Jeśli są k klasy i jeśli model dla $p(F_i)$ może być wyrażony przez r parametrów, wtedy odpowiadający naiwny model Bayesa ma $(k-1) + nrk$ parametrów. W praktyce często $k = 2$ (klasyfikacja binarna) i $r = 1$ (zmienna Bernoulliego jako cecha), wtedy całkowita liczba parametrów naiwnego modelu Bayesa to $2n + 1$, gdzie n jest liczbą cech.

W przypadku uczenia z nadzorem, parametry modelu można obliczyć stosując wnioskowanie Bayesa lub inną parametryczną procedurę estymacji.

Dotychczasowe omówienie problemu wyprowadziło model niezależnych cech, które są naiwnym probabilistycznym modelem Bayesa. Naiwny klasyfikator bayesowski łączy ten model z regułą decyzyjną:

$$\text{class}(f_1, \dots, f_n) = \underset{c}{\operatorname{argmax}} p(C = c) \prod_{i=1}^n p(F_i = f_i|C = c)$$

Przykład

Chcemy sprawdzić czy dany dokument D jest spamem - S czy nie - $\neg S$. Prawdopodobieństwo, że i -te słowo zdarza się w dokumencie klasy C zapisujemy, jako:

$$p(w_i|C)$$

Wtedy prawdopodobieństwo, że dokument D należy do klasy C wynosi (na podstawie twierdzenia Bayesa):

$$p(C|D) = \frac{p(C)}{P(D)} \prod_i p(w_i|C), \text{ gdzie } P(D) = \prod_i p(w_i)$$

$$p(S|D) = \frac{p(S)}{P(D)} \prod_i p(w_i|S)$$

$$p(\neg S|D) = \frac{p(\neg S)}{P(D)} \prod_i p(w_i|\neg S)$$

$\text{class}(D) = S$, jeżeli $p(S|D) > p(\neg S|D)$ oraz $\text{class}(D) = \neg S$ w przeciwnym wypadku.

2.7 Metody badania sygnałów w dziedzinie czasu i częstotliwości.

Transformata Fouriera - badanie sygnału stacjonarnego w dziedzinie częstotliwości.

$$FT(f) = \int_{-\infty}^{\infty} x(t) \cdot e^{-2i\pi ft} dt$$

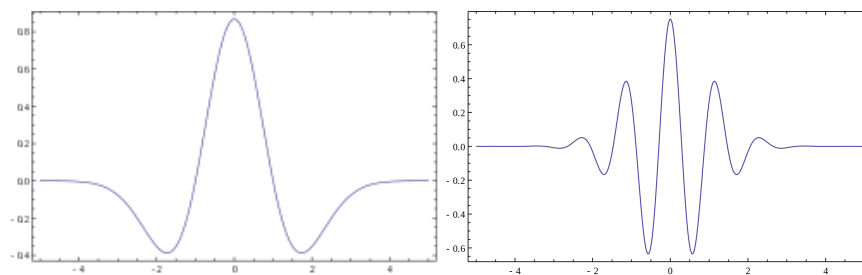
Jeśli wynik tego całkowania ma dużą wartość w punkcie f oznacz to, że sygnał ma wysoką wartość składowej widmowej o częstotliwości f . W przeciwnym wypadku widmo sygnału nie posiada istotnej składowej o tej częstotliwości.

Transformata Fouriera w oknie czasowym STFT - stała się alternatywą dla transformacji Fouriera z tego względu, że nadaje się do analizowania sygnałów niestacjonarnych. Niestacjonarny sygnał dzieli się na okna czasowe.

we, w których powstałe sygnały są w przybliżeniu stacjonarne. Okno czasowe powstają poprzez iloczyn sygnału oraz okna reprezentowanego przez prostokąt ($f_O = 1$ w oknie, $f_O = 0$ poza oknem). Stosując szerokie okno otrzymujemy dokładne informacje o niskich częstotliwościach, natomiast stosując wąskie okno – dokładne informacje o wysokich częstotliwościach. Wymiary analizowanego okna zmieniają się w zależności od tego, którą z informacji (o czasie, czy o częstotliwości), chcemy uzyskać z większą wiarygodnością.

$$STFT(f)_{f_O} = \int_{-\infty}^{\infty} f_O(t_1, t_2) x(t) \cdot e^{-2i\pi f t} dt$$

Ciągła transformata falkowa CWT - nie zatracając charakteru badanego szeregu, polega na dzieleniu sygnału na mniejsze części, za pomocą okien wagowych o zmieniających się wymiarach, a następnie na analizowaniu każdej z nich poprzez porównywanie z przesuniętą i przeskalowaną funkcją podstawową. W wyniku zastosowania CWT otrzymujemy macierz współczynników mówiących o tym, w jakim stopniu dana część sygnału pokrywa się z porównywaną falką podstawową (rys. 13).



Rysunek 13: Przykładowe falki, po lewej - *sbrero*, po prawej - *morlet*.

$$C(\text{skala}, \text{pozycja}) = \int f(t) \Psi(\text{skala}, \text{pozycja}, t) dt, \text{ gdzie:}$$

C - macierz współczynników,

$f(t)$ - oryginalny sygnał,

Ψ - falka.

Dobór skali odbywa się na zasadzie kostki Heisenberga - rozdzielczości czasowa oraz częstotliwościowa zmieniają się, natomiast ich stosunek jest stały.

Ciągła transformacja falkowa zdefiniowana jest jako suma całego omawianego sygnału pomnożonego przez przeskalowane i przesunięte wersje falki podstawowej:

$$CWT_x^{\Psi^*}(s, \tau) = \int_{-\infty}^{\infty} x(t) \Psi_{s, \tau}^* dt, \text{ gdzie:}$$

s - skala,

τ - przesunięcie (pozycja),

$x(t)$ - oryginalny sygnał,

Ψ^* - falka przeskalowana i przesunięta.

$$\Psi^*(s, \tau) = \frac{1}{\sqrt{|s|}} \Psi\left(\frac{t - \tau}{s}\right)$$

Empiryczna analiza modów - EMD

Funkcja modów własnych spełnia dwa warunki:

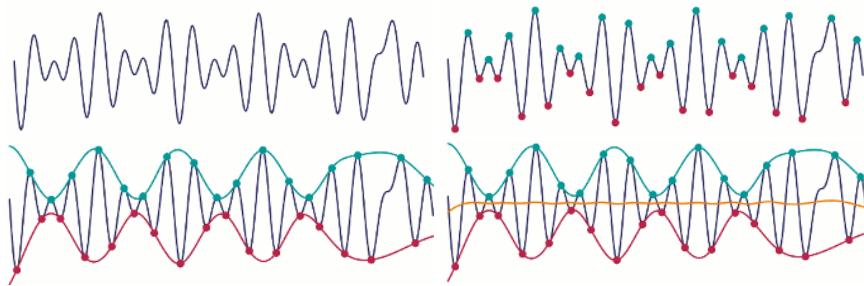
- Na całej szerokości sygnału liczba przejść przez zero oraz liczba ekstremów musi być, albo jednakowa, albo różnić się maksymalnie o jeden.
- W każdym punkcie sygnału średnia wartość obwiedni dolnej oraz górnej wynosi zero.

Dekompozycja na mody empiryczne oparta jest na założeniach:

- Badany sygnał posiada przynajmniej dwa ekstrema - jedno minimum oraz jedno maksimum.
- Charakterystyczna dla sygnału skala czasowa wynika z odstępów między ekstremami.
- Jeśli sygnał nie ma żadnych ekstremów, a jedynie punkty przecięcia, można go zróżniczkować jednokrotnie lub wielokrotnie.

Procedura "przesiewania" (otrzymywania funkcji IMF - *intrinsic mode functions*) dla metody EMD (rys 14):

- Wyznaczyć wszystkie ekstrema sygnału.
- Połączyć wszystkie maksyma za pomocą *spline lines* - otrzymujemy obwiednię max_i
- To samo dla minimów - otrzymujemy obwiednię min_i
- obliczamy średnią $m_i = \frac{min_i + max_i}{2}$
- $h_i = h_{i-1} - m_i$, gdzie $h_0 = X(t)$. Jeśli w kroku i prawdziwym będzie kryterium stop np. $std(h_i, h_{i-1}) < \epsilon$ wtedy nasze h_i jest modem IMF.



Rysunek 14: Kolejne kroki pojedynczej iteracji procedury "przesiewania" dla metody EMD.

Kolejne mody otrzymujemy poprzez metode przesiewania dla różnicy sygnału oraz znalezionych modów $r_k = X(t) - \sum_{i=0}^k c_i$, gdzie k - liczba znalezionych modów. Jeśli r_k jest funkcją monotoniczną, więcej modów nie może być znalezione - kończymy procedurę.

EMD może być wykorzystywana np. do badania wad osi układów mechanicznych.

2.8 Synchronizacja w nieliniowych układach dynamicznych.

Synchronizacja tożsamościowa - oba sygnały opisujące dynamikę układów są identyczne.

Synchronizacja uogólniona - odpowiadające sobie zmienne są proporcjonalne (np. amplituda)

Aby doszło do powyższych synchronizacji układy muszą być sprzężone.

Synchronizacja fazowa - wprowadza się pojęcie sygnału analitycznego:

$\psi(t) = s(t) + j\tilde{s}(t) = Ae^{j\omega\phi(t)}$, gdzie:

$\tilde{s}(t)$ - transformata Hilberta sygnału $s(t)$,

$\phi(t)$ - faza sygnału $s(t)$.

Układy są zsynchronizowane fazowo jeśli różnica faz $|n\psi_1(t) + m\psi_2(t)| = \text{const.}$

Istnieje wiele układów pomiędzy którymi może wystąpić jedynie synchronizacja fazowa. Przykładem są układy impulsowe jak pewien gatunek świetlików żyjący w Indonezji.

Model Kuramoto - model opisujący synchronizację w grupie sprzężonych oscylatorów.

Opisuje układ N rotatorów. Tempo zmiany fazy θ_i , i -tego rotatora:

$$\frac{d\theta_i}{dt} = \omega_0^i + \frac{k}{N} \sum_{j=1}^N \sin(\theta_i - \theta_j), \text{ gdzie:}$$

ω_0^i - naturalna częstość rotatora i ,

k - amplituda sprzężenia.

Gdy $k > 0$ fazy rotatorów powinny się zsynchronizować:

- Jeśli rotator i spóźnia się względem średniego rotatora $\tilde{\theta}(t) - \theta_i(t) > 0$, wtedy rotator i przyspiesza.
- Gdy rotator i śpieszy się względem średniego rotatora $\tilde{\theta}(t) - \theta_i(t) < 0$, wtedy rotator i zwalnia.

Istnieje pewna wartość krytyczna amplitudy sprzężenia k_C , powyżej której scyntylatory się zsynchronizują. Dla $N = 2$ wartość ta jest zależy od różnicy naturalnych częstości rotatorów $|\omega_1^0 - \omega_2^0|$. Dla $N \gg 1$, gdy naturalne częstości pochodzą z rozkładu $\mathcal{N}(0, \sigma^2)$ wtedy wartość krytyczna zależy od σ . **Wniosek:** Im większe zróżnicowanie naturalnych częstości rotatorów tym trudniej zsynchronizować układ.

Synchronizacja grupy świetlików

Układ składa się z N oscylatorów.

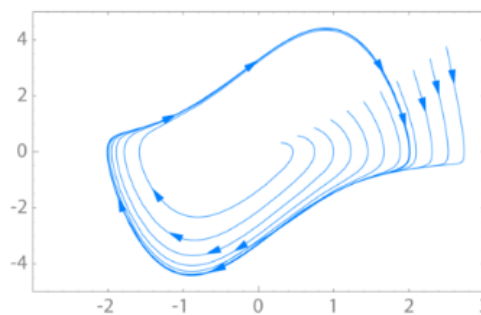
Naładowanie oscylatora definiuje się następująco:

$$U(\phi(t)) = C(1 - e^{-\gamma\phi(t)}), \text{ gdzie:}$$

$\phi(t)$ - faza oscylatora.

Gdy $U(\phi) = 1$, następuje rozładowanie oscylatora ($U(\phi) = 0$) i każdy ze sprzężonych oscylatorów ładuje się o ϵ co jednocześnie "przesuwa" fazy tych oscylatorów.

Kształt funkcji naładowania powoduje, że po pewnym czasie dochodzi do sprzężenia faz oscylatorów.



Rysunek 15: "Wpadanie" w cykl graniczny oscylatora Pola dla różnych warunków początkowych.

W oscylatorach nieliniowych z cyklem granicznym faza, amplituda oraz częstość są własnościami stanu oscylatora. Po pewnym stanie nieustalonym

układ osiąga cykl graniczny (rys 15), niezależnie od warunków początkowych. Przykładem takiego oscylatora jest oscylator van der Pola opisany równaniem:

$$\frac{d^2x}{dt^2} + \mu(x^2 - 1)\frac{dx}{dt} + x = 0$$

Sprzężone oscylatory nieliniowe zazwyczaj starają się "śpiewać wspólnie" tzn. z jedną częstotliwością przy czym zazwyczaj częstotliwość "silniejszego" (np. tego o większej amplitudzie) bierze górę, lub znajdują wspólną częstotliwość - zazwyczaj ustalony stosunek ich częstotliwości własnych. Zjawisko sprzężenia częstotliwości (drżania synfazowe) jest rodzajem synchronizacji. Oscylatory liniowe tego nie potrafią: nie są w stanie zmieniać swej częstotliwości. Nawet sprzężone, nieliniowe oscylatory nie zawsze sobie radzą ze znalezieniem wspólnej częstotliwości, stąd wynika kwaziperiodyczność (ruch z kilkoma niewspółmiernymi częstotliwościami) lub ruch chaotyczny (trajektoria układu w przestrzeni fazowej jest otwarta, ale ograniczona przestrzennie).

Dzięki swoim własnościom oscylatory nieliniowe są dobrymi kandydatami do:

- Modelowania zjawisk biologicznych - wiele rytmów jest cechą własną gatunku, niezależną od warunków początkowych oraz względnie małych zmian otoczenia, a jednocześnie muszą być w zgodzie z rytmem natury (np. dzień/noc)
- do zastosowań technicznych w telekomunikacji, czy w technikach pomiarowych.

Synchronizacja fazowa znalazła zastosowanie m.in. w medycynie, gdzie pozwala analizować oddziaływanie oddechu z rytmem serca.

2.9 Równanie Fokkera-Plancka.

Równanie Fokkera-Plancka – równanie różniczkowe cząstkowe drugiego rzędu. Opisuje ewolucję czasową funkcji gęstości prawdopodobieństwa $p(\mathbf{x}, \mathbf{v}, t)$ położenia i prędkości. Po raz pierwszy równanie to zostało użyte do opisu zjawiska ruchów Browna cząstki zanurzonej w cieczy. Ogólna forma równania dla N zmiennych:

$$\frac{\partial}{\partial t} p(\mathbf{x}, \mathbf{v}, t) = \left[- \sum_{i=1}^N \frac{\partial}{\partial x_i} D_i^1(x_1, \dots, x_N) + \sum_{i=1}^N \sum_{j=1}^N \frac{\partial^2}{\partial x_i \partial x_j} D_{i,j}^2(x_1, \dots, x_N) \right] p(\mathbf{x}, \mathbf{v}, t),$$

gdzie D^1 to wektor dryftu, a D^2 oznacza tensor dyfuzji.

Jeśli nie interesują nas prędkości oraz poruszamy się tylko w jednym wymiarze, wtedy równanie F-P upraszcza się do postaci:

$$\frac{\partial}{\partial t} p(x, t) = \left[- \frac{\partial}{\partial x} D^1(x) + \frac{\partial^2}{\partial x^2} D^2(x) \right] p(x, t)$$

Następnie możemy zdefiniować potencjał

$$\Phi(x) = \ln D^2(x) + \int_0^x \frac{D^1(x')}{D^2(x')} dx'$$

oraz prąd prawdopodobieństwa:

$$\begin{aligned} \frac{\partial}{\partial t} p(x, t) &= \frac{\partial}{\partial x} J(x, t) \\ J(x, t) &= D^1(x, t) p(x, t) - \frac{\partial}{\partial x} [D^2(x, t) p(x, t)] \end{aligned}$$

Jeśli J wyrazimy poprzez Φ oraz skorzystamy z warunku brzegowego $J(x, t) = 0$, dla $x \in (a, b)$ to otrzymamy rozwiązanie stacjonarne gęstości prawdopodobieństwa:

$$p_s(x) = \frac{N_0}{D^2(x)} \exp \left[\int_0^x \frac{D^1(x')}{D^2(x')} dx' \right], \text{ gdzie:}$$

$$N_0 = \left(\int_a^b \exp[-\Phi(x)] dx \right)^{-1}$$

Przykład. Dyfuzja w polu grawitacyjnym: $D^1 = -g = \text{const.}$, $D^2 = \frac{D}{2} = \text{const.}$

Rozwiązanie stacjonarne gęstości prawdopodobieństwa:

$$p_s(x) = \frac{2N_0}{D} \exp \left(-\frac{2g}{D} x \right)$$

2.10 Fraktalne i multifraktalne metody analizy sygnałów.

Fraktale są to obiekty geometryczne, których wzorce powtarzają się w różnych skalach. Mimo nietrywialnej struktury w każdej skali mają względnie prostą definicję rekurencyjną.

Gdy sygnały o nieregularnym charakterze (ruchy Browna, kształty w geologii, szeregi czasowe w medycynie) i ich wykresy są zbiorami fraktalnymi w R^2 - określa się te sygnały jako funkcje fraktalne.

Funkcje takie są samoafiniczne: są samopodobne gdy się je przekształca za pomocą anizotropowego skalowania:

$f(x)$ jest funkcją samoafiniczną gdy $\forall x_0 \in \mathbb{R}, \exists H \in \mathbb{R}, \lambda > 0$ takie, że:
 $f(x_0 + \lambda x) - f(x_0) \approx \lambda^H [f(x_0 + x) - f(x_0)]$, gdzie:

H - wykładnik Hursta będący miarą chropowatości funkcji. Wykres funkcji jest samopodobny jeśli $H = 1$.

Funkcje fraktalne są zazwyczaj multifraktalne - ich własności zmieniają się od punktu do punktu:

$|f(x+l) - f(x)| \propto l^{h(x)}$, gdzie:

$h(x)$ - wykładnik Holdera funkcji f punkcie x .

Multifraktalność polega na tym, że dla każdej wartości argumentu x funkcji $f(x)$ możemy mieć inny wykładnik Holdera.

Wymiar fraktalny d jest definiowany jako:

$$N = M^d,$$

$$d = \frac{\ln N}{\ln M}, \text{ gdzie:}$$

N - ilość powstałych części podczas kroku konstrukcyjnego,

M - ułamek liniowej długości nowej części w procesie konstrukcji do liniowej długości części pierwotnej.

Np. dla zbioru Cantora $N = 2$, $M = 3$, $d = \frac{\ln 2}{\ln 3} \approx 0.63$ (rys. 16).



Rysunek 16: Kolejne kroki konstrukcyjne zbioru Cantora.

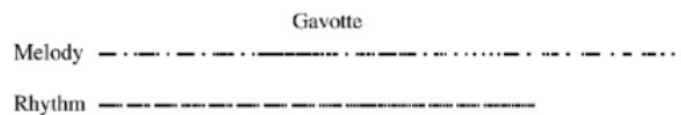
Multifraktalna analiza szeregów muzycznych

Melodia i rytm są dwoma podstawowymi cechami utworu muzycznego. Melodia jest sekwencją zmian wysokości dźwięku. Rytm jest sekwencją długości trwania dźwięków melodii. Melodię odwzorowuje się na rozkład punktów, dzieląc oktawę równomiernie na 12 wysokości tonu.

Konwersja zapisu muzycznego na szeregi punktowe (przykład na rysunku 17):

- pierwszy dźwięk melodii przyjmuje się jako podstawowy i stawia się punkt na pierwszej pozycji, na osi poziomej
- jeśli różnica wysokości dźwięku drugiego wynosi m stawia się punkt na pozycji $1 + m$

Podobnie postępujemy z rytmem. Wybieramy najniższą miarę czasu (np. szesnastka). Jeśli m to różnica odległości kolejnego dźwięku od poprzedniego w szeregu miar czasowych, to stawiamy punkt w miejscu oddalonym o m od poprzedniego. Różnica pomiędzy szesnastką a ósemką to 1, szesnastką, a ćwierćnotą - 2 itd.



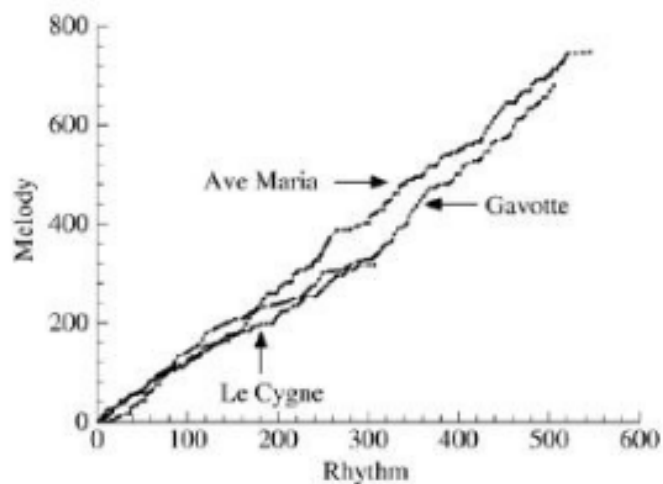
Rysunek 17: Konwersja utworu *Gavotte* na szeregi punktowe.

Obliczanie wykładnika Holdera:

- wybieramy punkt i na osi poziomej (zaczerniony bądź pusty)
- zliczamy liczbę punktów p_i w pokryciu o promieniu r i wyliczamy rozkład $p_i(r)$, zmieniając wartość r ,
- wykładnik Holdera dany jest jako nachylenie funkcji $p_i(r)$ w skali $\log - \log$,
- przysuwając się po szeregu punktowym, otrzymujemy zależność wykładnika od położenia i .

Wnioski:

- Utwory muzyczne posiadają własności multifraktalne.
- Metoda pozwala segregować utwory muzyczne ze względu na własności melodi oraz ze względu na rytmikę utworu.
- Oba szeregi można połączyć, tworząc unikatową charakterystykę każdego utworu - otrzymuje się fraktal podobny do schodków diabelskich (rys. 18).



Rysunek 18: Przykłady połączonych konwersji dla rytmu i melodii.

Multifraktalna analiza beztrendowa (uogólniona multifraktalna procedura DFA)

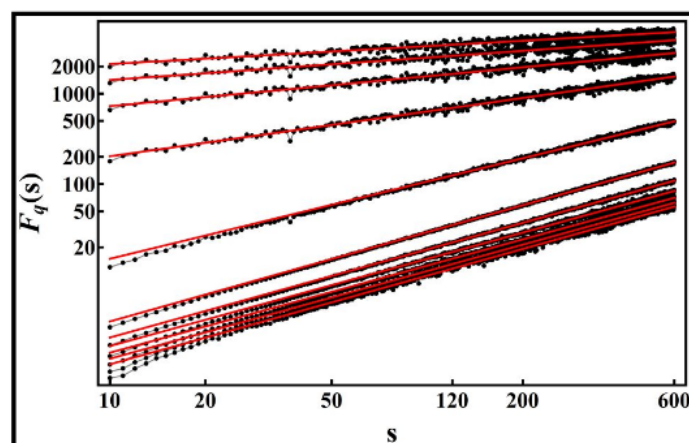
- Dzielimy sygnał na k segmentów o długości s ,
- Dla każdego segmentu wykonujemy regresję (liniową lub nieliniową) i odejmujemy się jej wynik od sygnału w segmencie
- dla różnych długości segmentu s oraz różnych wartości parametru q oblicza się miarę fluktuacji zdefiniowaną następująco:

$$F_q(s) = \frac{1}{2N_s} \left[\sum_{v=1}^{2N_s} F^2(v, s)^{\frac{q}{2}} \right]^{\frac{1}{q}}, \text{ gdzie:}$$

v - indeks okna,

q - parametr, jeśli $q < 0$ - małe fluktuacje, jeśli $q > 0$ - duże.

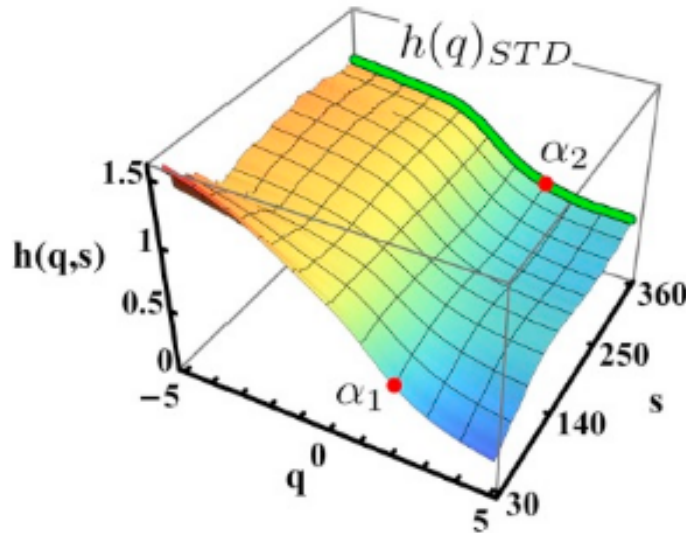
- plotuje się wyniki w skali log-log (rys. 19)
- jeśli analizowany sygnał jest daleko-zasięgowo skorelowany według prawa potęgowego wtedy $F_q(s) \propto s^{h(q)}$, gdzie h - uogólniony wykładnik Hursta.



Rysunek 19: Miara fluktuacji w funkcji szerokości segmentu dla różnych wartości parametru q .

Wielkoskalowa analiza multifraktalna

- Stosujemy przesuwające się i rozszerzające okno
- Regresja wykonywana jest tylko wewnątrz aktualnego położenia okna
- Ponieważ $F_q(s)$ jest w skali log – log okno rozszerza się logarytmicznie
- Powstaje tzw. powierzchnia Hursta (rys 20)



Rysunek 20: Miara fluktuacji w funkcji szerokości segmentu oraz wartości parametru q .

2.11 Teoria pola średniego dla modelu Isinga.

Metoda pola średniego dla modelu Isinga w sieciach regularnych

W sieciach regularnych każdy wierzchołek ma taki sam stopień.

Energia modelu Isinga:

$$E = -J \sum_i s_i \sum_{\langle j \rangle_i} s_j - \sum_i h_i s_i, \text{ gdzie:}$$

J - macierz oddziaływań pomiędzy węzłami,

$\langle j \rangle_i$ - najbliższe sąsiedztwo węzła i ,

h_i - energia spinu i w zewnętrznym polu. Zakładamy pole jednorodne $\forall i h_i = h$.

Z dobrym przybliżeniem można przyjąć, że wszystkie węzły sieci i związane z nimi spiny średnio są takie same. Drugą sumę w pierwszym członie można zastąpić jej wartością uśrednioną:

$$E \approx -J \sum_i s_i \cdot z \cdot \langle s \rangle - h \sum_i s_i, \text{ gdzie:}$$

z - liczba najbliższych sąsiadów,

$$\langle s \rangle = \frac{1}{N} \sum_i s_i,$$

N - liczba węzłów

$$E_i \approx -J \cdot s_i \cdot z \cdot \langle s \rangle - h s_i$$

Zagadnienie przybliżenia średniego pola pozwala uprościć zagadnienie modelu Isinga:

$$p(s_i) = \frac{e^{\beta(Jz\langle s \rangle s_i + h s_i)}}{Z_i}, \text{ gdzie:}$$

$$Z_i = e^{\beta(Jz\langle s \rangle + h)} + e^{-\beta(Jz\langle s \rangle + h)} = 2 \cosh(\beta(Jz\langle s \rangle + h))$$

$$\langle s \rangle = \sum_{s_i=\pm 1} s_i p(s_i) = \tanh(\beta J z \langle s \rangle + h) \rightarrow \text{uwikłane równanie na } \langle s \rangle.$$

Różna od zera wartości $\langle s \rangle$ są możliwe tylko gdy temperatura jest większa od pewnej temperatury krytycznej $T_c = \frac{1}{k_B \beta_c}$. Temperaturę tą obliczamy stosując liniowe przybliżenie \tanh dla małych wartości argumentu (założmy także, że nie ma pola zewnętrznego tzn $h = 0$):

$$\langle s \rangle = \beta_c J z \langle s \rangle,$$

$$\langle s \rangle (1 - \beta_c J z) = 0,$$

$$\beta_c = \frac{1}{J z},$$

$$T_c = \frac{1}{k_B \beta_c}.$$

Dla temperatury niższej od T_c układ może mieć wartość średniego spinu różną od zera.

Dla lepszego przybliżenia tangensa:

$$\tanh x \approx x - \frac{1}{3}x^3$$

otrzymujemy:

$$\langle s \rangle = \beta z J \langle s \rangle - \frac{1}{3} \beta^3 z^3 J^3 \langle s \rangle^3,$$

$$\langle s \rangle = \sqrt{\frac{3\beta z J - 3}{\beta^3 z^3 J^3}}, \text{ podstawiając:}$$

$$\beta z J = \frac{T_0}{T}, \text{ otrzymujemy:}$$

$$\langle s \rangle = \sqrt{3}^{3/2} \sqrt{\frac{T}{T_0}} \sqrt{\frac{T_0 - T}{T}} \rightarrow \langle s \rangle \propto \sqrt{T_0 - T} \text{ wskazuje na ciągłą przemianę fazową.}$$

2.12 Zjawisko złamania symetrii w układach rzeczywistych: przykłady, modele.

Ferromagnetyzm

Oddziaływania magnetyczne dipoli w ferromagnetykach dążą do ustawienia ich równolegle. Zależą więc od ich wzajemnego ustawienia, ale nie wyróżniają żadnego bezwzględnego kierunku w przestrzeni. Są symetryczne względem obrotów w zwykłej przestrzeni konfiguracyjnej. Powyżej pewnej temperatury, zwanej temperaturą Curie (T_C), energia kinetyczna dipoli jest na tyle duża, że nie pozwala na ich wzajemną korelację. Stan ferromagnetyka wykazuje więc też symetrię obrotową. Jeżeli temperatura spadnie poniżej T_C , to oddziaływania dipoli są już na tyle silne, że porządkują ustawienie dipoli i ferromagnetyk ma pewne namagnesowanie wyróżniające jakiś kierunek w przestrzeni. Powyżej T_C wszystkie kierunki były takie same, żaden nie był wyróżniony, poniżej T_C ferromagnetyk "spontanicznie" wybiera jakiś kierunek magnetyzacji. Zwróćmy uwagę na to, że oddziaływanie dipoli nie ulega zmianie, dalej ma symetrię obrotową. W tym sensie teoria nadal jest symetryczna. To stan ferromagnetyka nie wykazuje symetrii. Zjawisko to można symulować, korzystając z modelu Isinga (patrz zagadnienie wyżej).

Model żerowania mrówek

Mrówki mają do dyspozycji dwie ścieżki między mrowiskiem, a pokarmem. Prawdopodobieństwo wyboru ścieżki i :

$\frac{(x_i + k)^\alpha}{(x_i + k)^\alpha + (x_j + k)^\alpha}$, gdzie:
 x_i - ilość feromonu na ścieżce i ,
 α - stopień nieliniowości wyboru,
 k - stopień atrakcyjności ścieżki bez feromonu (im większy tym tym więcej feromonu potrzeba do preferencji danej ścieżki).

Dodatkowe założenia:

- Stała liczba mrówek φ .
- Mrówki wracają tą samą drogą, co przyszły.
- Ilości feromonu pozostawionego na ścieżce zależy od jej długości lub od jej jakości q_i .
- Feromon ulatnia się ze stałą prędkością v .

Tempo zmian koncentracji feromonu:

$$\frac{dx_1}{dt} = \varphi q_1 \frac{(x_1 + k)^\alpha}{(x_1 + k)^\alpha + (x_2 + k)^\alpha} - vx_1,$$

$$\frac{dx_2}{dt} = \varphi q_2 \frac{(x_2 + k)^\alpha}{(x_1 + k)^\alpha + (x_2 + k)^\alpha} - vx_2$$

Rozważmy rozwiązanie równowagowe dla dwóch jednakowych ścieżek:

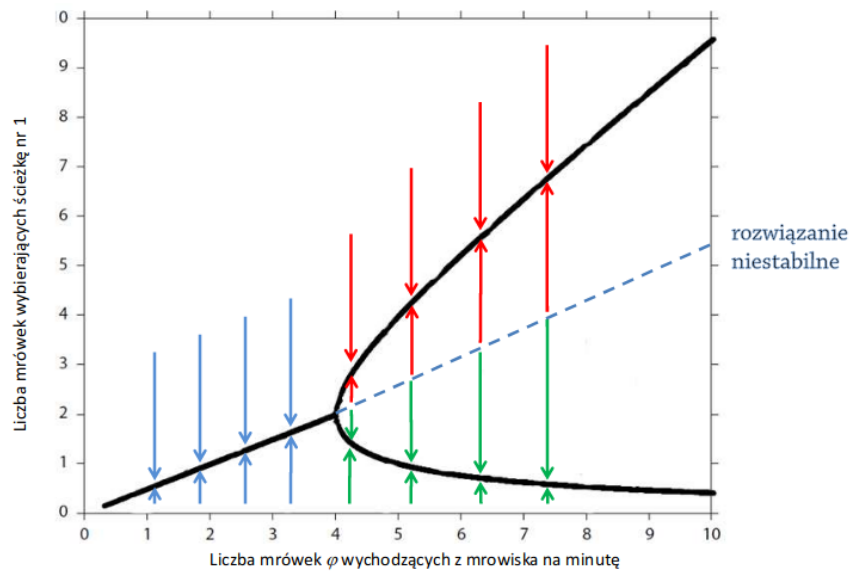
$$q_1 = q_2 = q \text{ oraz } \frac{dx_1}{dt} = \frac{dx_2}{dt} = 0:$$

$$vx_1 = \varphi qp_1,$$

$$vx_2 = \varphi qp_2,$$

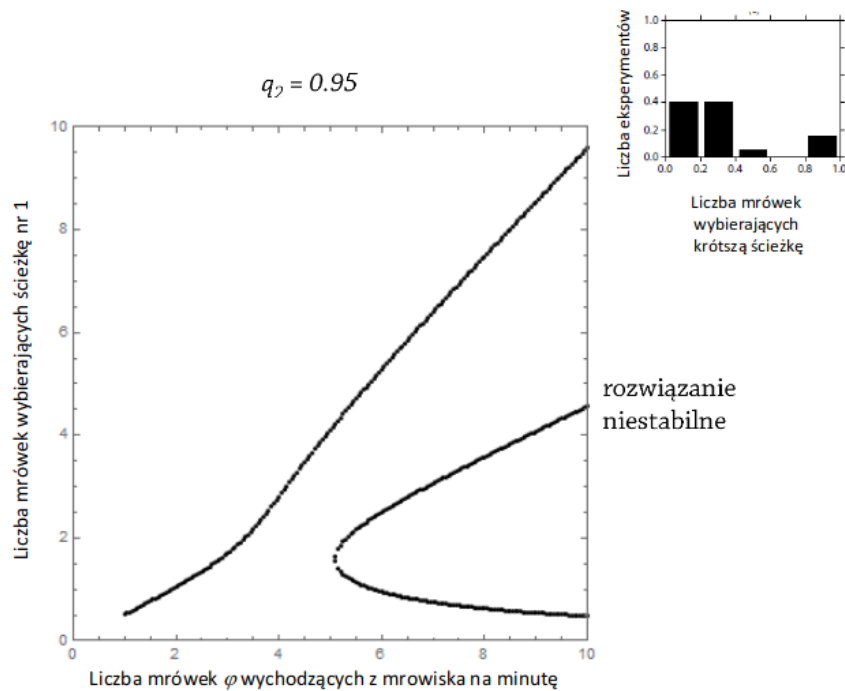
$$\frac{x_1}{p_1} = \frac{x_2}{p_2} \rightarrow \left(\frac{x_2 + k}{x_1 + k} \right)^\alpha = \frac{x_2}{x_1}$$

W przypadku możliwości wyboru między wieloma celami, załamanie symetrii zachodzi, gdy opłacalnym jest dla większości grupy, by podejmować tą samą decyzję (rys. 21).



Rysunek 21: Diagram bifurkacyjny dla $q = v = 1$, $k = 2$.

Dla różnych wartości q_i rozwiązanie łamiące symetrię jest niestabilne (rys. 22)



Rysunek 22: Diagram bifurkacyjny dla $q_1 = v = 1, k = 2$.

2.13 Informacja wzajemna w przetwarzaniu i transmisji danych.

Informacja

$I_i = -\log_2 p_i$, gdzie:

I_i - ilość informacji otrzymanej przy zdarzeniu x_i ,

p_i - prawdopodobieństwo zajścia zdarzenia x_i

Entropia

Przeciętna ilość informacji przypadająca na zajście zdarzenia z pewnego zbioru n zdarzeń (entropia bezwarunkowa tego zbioru, entropia przeciętna) jest średnią arytmetyczną ważoną ilości informacji otrzymywanej przy zajściu poszczególnych zdarzeń, gdzie wagami są prawdopodobieństwa tych zdarzeń:

$$H(X) = - \sum_{i=1}^n p(x_i) \log p(x_i)$$

Entropia względna jest miarą stosowaną w statystyce i teorii informacji do określenia rozbieżności między dwoma rozkładami prawdopodobieństwa p i q (wynosi zero dla identycznych rozkładów):

$$d_{KL}(P, Q) = \sum_i p(x_i) \log_2 \frac{p(x_i)}{q(x_i)}$$

Informacja wzajemna mierzy, ile informacji o X można poznać, znając Y , czyli o ile poznanie jednej z tych zmiennych zmniejsza niepewność o drugiej:

$$I(X; Y) = \sum_i \sum_j p(x_i, y_j) \log \frac{p(x_i, y_j)}{p(x_i)p(y_j)} = d_{KL}(P(X, Y), P(X)P(Y))$$

Jeśli zmienne X i Y są niezależne ($P(X, Y) = P(X)P(Y)$), to ich wzajemna informacja jest zerowa (znajomość jednej nie mówi niczego o drugiej). Jeśli X i Y są identyczne, to każda zawiera pełną wiedzę o drugiej. Wtedy informacja wzajemna jest równa entropii X (albo Y – skoro są identyczne, to ich entropia jest taka sama).

Informacja wzajemna to ilość informacji o zdarzeniach ze zbioru X (wartościach zmiennej losowej X), np. komunikatach nadanych (stanach źródła informacji), zawarta w zdarzeniach ze zbioru Y (wartościach zmiennej losowej Y), np. komunikatach odebranych (stanach odbiorcy). Definicja ta oznacza, że informacja wzajemna jest miarą efektywności transmisji.

Komunikacja jest procesem fizycznym - transmisja informacji może podlegać niekontrolowanym zniekształceniom - dwa różne komunikaty nadawane mogą być odebrane jako ten sam komunikat.

Kanał informacyjny jest to medium służące do wymiany informacji. Wyjście zależy probabilistycznie od wejścia: $X \rightarrow Y$. Przepustowość kanału informacyjnego:

$$C = \max_{P(X)} I(X, Y)$$

Informacja wzajemna jest średnią redukcją niepewności na temat jednego źródła przy uzyskaniu informacji na temat drugiego.

Kanał komunikacyjny składa się z dwóch różnych źródeł: wejściowego X i wyjściowego Y oraz zestawu prawdopodobieństw warunkowych: $P(Y = b | X = a)$.

Przepustowość kanału to maksymalna wartość informacji wzajemnej między wejściem, a wyjściem jaką możemy otrzymać, zmieniając rozkład prawdopodobieństw na wejściu.

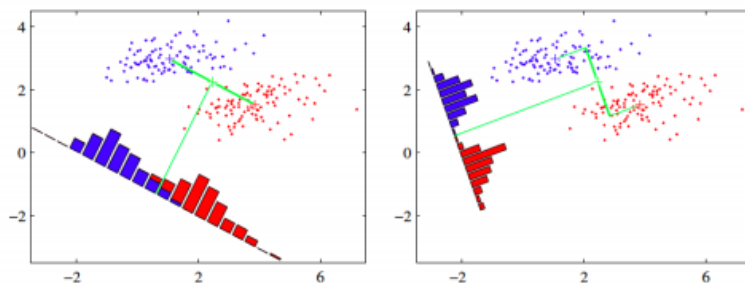
2.14 Metody redukcji wymiaru w eksploracji danych.

LDA - *Linear discriminant analysis* jest używana do znalezienia liniowej kombinacji cech, które najlepiej rozróżniają dwie lub więcej klas obiektów lub zdarzeń. Wynikowe kombinacje są używane jako klasyfikator liniowy lub, częściej, służą redukcji wymiarów do późniejszej klasyfikacji statystycznej.

Rozważmy zbiór obiektów z których każdy jest opisywany przez wektor cech x , przy czym dla każdego obiektu znana jest jego przynależność do klasy y . Podejście LDA bazuje na założeniu, że funkcje gęstości prawdopodobieństwa $\forall i \rightarrow p(x|y = i)$ mają rozkłady normalne i jednakowe macierze kowariancji:

$$S = \frac{1}{n-1} \sum_{i=1}^{i=n} (x_i - \bar{x})(x_i - \bar{x})^T$$

LDA próbuje znaleźć podprzestrzeń cech, która maksymalizuje separowalność klas. Dla przypadku dwóch klas oraz dwuwymiarowego wektora cech szukamy takiego kierunku, dla którego rozkłady rzutów zbiorów poszczególnych klas na ten kierunek są jak najbardziej rozdzielone od siebie (rys. 23).



Rysunek 23: Przykład LDA.

QDA - Quadratic discriminant analysis - zakładamy różne macierze kowariancji. QDA również próbuje znaleźć podprzestrzeń cech, która maksymalizuje separowalność klas. Dla powyższego przypadku będzie szukać najlepszego rzutu dla rozszerzonego dwuwymiarowego wektora cech o wartości 2-go rzędu tzn. $[x_1, x_2] \rightarrow [x_1, x_2, x_1^2, x_1x_2, x_2^2]$.

PCA - Principal components analysis

- zakładamy zbiór wielowymiarowych obserwacji, leżących w przestrzeni R^p ,
- najczęściej obserwacje nie są równomiernie rozrzucone wzdłuż wszystkich kierunków układu współrzędnych,
- koncentruj się w pewnych podprzestrzeniach przestrzeni R^p ,
- kierunki, wzdłuż których znajduje się większość obserwacji, nie muszą się pokrywać z osiami układu współrzędnych R^p ,
- próba losowa pochodzi z populacji o ciągłym rozkładzie w przestrzeni R^p z wektorem wartości oczekiwanych \mathbf{m} i macierzą kowariancji \mathbf{S}

Pierwszą składową główną szukamy poprzez znalezienie takiego kierunku $\gamma_1 \in R^p$, aby rzut ortogonalny wektora losowego \mathbf{x} na ten kierunek dawał zmienną losową o największej wariancji. Kolejne składowe szukamy w ten sam sposób biorąc pod uwagę warunek aby kolejne kierunki były prostopadłe do poprzednich.

Pierwszą składową główną wektora \mathbf{x} nazywamy rzut tego wektora na znaleziony kierunek przesunięty o wartość średnią:

$$y_1 = \gamma_1^T (\mathbf{x} - \mathbf{m})$$

Kolejne składowe główne odpowiadają kolejnym wektorom własnym macierzy kowariancji \mathbf{S} . Pierwsza składowa główna odpowiada wektorowi własnemu o największej wartości własnej (λ_1), kolejne wektory \rightarrow kolejne największe wartości własne.

Wielkość:

$$\frac{\lambda_1 + \dots + \lambda_k}{\lambda_1 + \dots + \lambda_p}$$

wyraża procent zmienności wektora losowego x wyjaśnionych przez k ($k \leq p$) pierwszych składowych głównych.

Skalowanie wielowymiarowe

Niech d_{ij} , $i, j = 1, \dots, n$ będą odległościami euklidesowymi między obserwacjami \mathbf{x}_i oraz \mathbf{x}_j w przestrzeni R^p . Zadanie polega na znalezienie takiej podprzestrzeni R^r o wymiarze r , by odległości euklidesowe d'_{ij} między rzutami obserwacji na tę podprzestrzeń minimalizowały sumę

$$V = \sum_{i=1}^n \sum_{j=1}^n (d_{ij}^2 - d'_{ij}^2)$$

Podobienstwo pomiędzy skalowaniem wielowymiarowym a analizą składowych głównych staje się równoważnością, gdy dana macierz d_{ij} jest macierzą odległości euklidesowych.

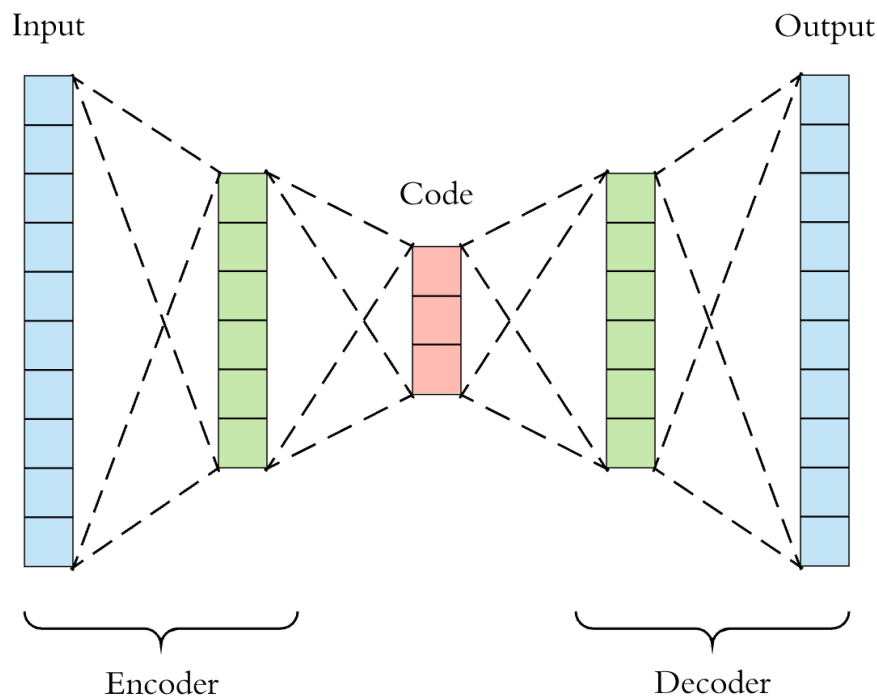
Autoencoder - jest rodzajem sztucznej sieci neuronowej wykorzystywanej do uczenia się wydajnego kodowania danych w sposób nienadzorowany. Celem *autoencodera* jest nauczenie się reprezentacji (kodowania) zestawu danych, zazwyczaj w celu zmniejszenia wymiarów, poprzez szkolenie sieci w zakresie ignorowania „szumu” sygnału.

Przykład

Powiedzmy że reprezentujemy zdanie poprzez wektor o długości liczby słów występujących w słowniku. Jeśli słowo występuje w zdaniu, wartość

elementu wektora odpowiadającego temu słowu wynosi 1, jeśli słowo występuje dwukrotnie - 2 itd... Słowniki składają się z tysięcy słów, więc nieoptymalna jest reprezentacja zdania w takiej przestrzeni, ponieważ tylko ułamek procenta składowych wektora reprezentującego jakiekolwiek zdanie będzie niezerowych.

Aby uzyskać reprezentację zdania w mniejszej przestrzeni możemy użyć *autoencodera*. Nasz zbiór uczący przepuszczamy przez sieć neuronową gdzie warstwa wejściowa składa się z N (ilość słów w słowniku) neuronów. Warstwa wyjściowa również składa się z N neuronów. Pomiedzy tymi warstwami jest nieparzysta liczba warstw ukrytych (zazwyczaj symetrycznych), a środkowa warstwa reprezentuje wektor kodujący (rys. 24). Podczas uczenia na wyjściu ustawiamy ten sam wektor co na wejściu.



Rysunek 24: Topologia sieci neuronowej typu *autoencoder*.

2.15 Metody wnioskowania statystycznego.

Wnioskowanie statystyczne polega na wyciąganiu wniosków dotyczących całej populacji na podstawie próby z tej populacji. Składowe próby powinny być niezależne oraz dobrze reprezentować populację.

Na podstawie próby $\{X_1, X_2, \dots, X_N\}$ chcemy opisać rozkład X . Podejścia:

1. Parametryczne - zakładamy, że X ma znany rozkład, nie znamy jednak parametrów tego rozkładu np. wiemy że X ma rozkład wykładniczy $e^{\lambda x}$, nie znamy λ .
2. Nieparametryczne - nic nie zakładamy o X .

Estymacja punktowa - polega na oszacowaniu wartości nieznanego parametru θ za pomocą funkcji mierzalnej, której argumentami są elementy próby losowej $\{X_1, X_2, \dots, X_N\}$.

Przykładem estymacji punktowej jest metoda największej wiarygodności: Niech $\{X_1, X_2, \dots, X_N\}$ będzie prostą próbą losową z rozkładu o gęstości

$f_{\theta}(x) =$, zaś $\{x_1, x_2, \dots, x_N\}$ jest jej realizacją. Wtedy funkcję:

$$L(\theta, x_1, x_2, \dots, x_N) = f_{\theta}(x_1)f_{\theta}(x_2)\dots f_{\theta}(x_N)$$

nazywamy funkcją wiarygodności. Estymatorem największej wiarygodności jest wartość parametru θ , która maksymalizuje funkcję wiarygodności.

Testy statystyczne - stawiamy hipotezę statystyczną (przypuszczenie dotyczące nieznanego rozkładu badanej cechy populacji o prawdziwości lub fałszywości którego wnioskujemy się na podstawie pobranej próby) np. wysuwamy hipotezę, że badana cecha ma rozkład normalny (hipoteza zerowa - H_0). Statystyką testową nazywamy funkcję próby $\delta(X_1, X_2, \dots, X_N)$, która służy do weryfikacji H_0 przeciwko H_1 . Zbiór wszystkich możliwych wartości funkcji δ dzielimy na dwa rozłączne zbiory W (zbiór krytyczny testu) oraz W' :

- jeśli $\delta(x_1, x_2, \dots, x_N) \in W$, to H_0 odrzucamy,
- jeśli $\delta(x_1, x_2, \dots, x_N) \in W'$, to H_0 przyjmujemy.

Wnioskujemy na podstawie próby, a nie całej populacji, jesteśmy więc narażeni na popełnienie błędu. Możemy:

- odrzucić H_0 gdy jest ona prawdziwa (błąd I-go rodzaju),
- przyjąć H_0 gdy jest ona fałszywa (błąd II-go rodzaju).

Test konstruuje się tak, aby maksymalne prawdopodobieństwo błędu I-go rodzaju wynosiło α (i nazywa się je poziomem istotności):

$$p(\delta(x_1, x_2, \dots, x_N) \in W | H_0) \leq \alpha$$

Najmniejszy poziom istotności, przy którym zaobserwowana wartość statystyki testowej prowadzi do odrzucenia H_0 nazywamy p wartością. Wartość p niższa od krytycznego poziomu istotności w podejściu częstościowym uprawnia jedynie do postępowania doraźnie tak, jakby hipoteza zerowa została odrzucona.

Moc testu parametrycznego to funkcja zmiennej θ :

$P(\theta) = p(\delta(x_1, x_2, \dots, x_N) \in W | \theta) =$ prawdopodobieństwo odrzucenia H_0 w sytuacji gdy nieznaną parametr wynosi θ .

Przykładowe testy:

Test Pearsona:

$$\chi^2 = \sum_{i=1}^n \left(\frac{O_i - E_i}{\sigma_i} \right)^2, \text{ gdzie:}$$

O_i - wartość mierzona,

E_i - wartość teoretyczna,

σ_i - odchylenie standardowe,

n - liczba pomiarów.

Otrzymaną wartość znormalizowanego testu ocenia się za pomocą rozkładu chi kwadrat. Polega to na obliczeniu prawdopodobieństwa otrzymania wartości testu równej lub większej, niż otrzymanej przez nas. Ponieważ takie obliczenia są dość skomplikowane, najczęściej korzysta się z danych zamieszczonych w odpowiednich tablicach. Mając odpowiednie prawdopodobieństwo, należy określić jego najmniejszą wartość, dla której jesteśmy w stanie zaakceptować hipotezę jako prawdę.

Test Kołmogorowa-Smirnowa - jest zwykle używany do sprawdzania, czy dana dystrybucja teoretyczna $F(x)$ opisuje rozkład populacji, z której wylosowano próbę o dystrybucji empirycznej $F_n(x)$.

Test t-Studenta - służy do:

- sprawdzenia hipotezy, czy na podstawie naszej próby możemy mamy podstawy do odrzucenia hipotezy, że populacja ma średnią równą danej średniej teoretycznej,
- sprawdzenie hipotezy, czy dwie niezależne próby pochodzą z populacji o tej samej średniej
- sprawdzenie hipotezy, czy dwie zależne próby (np. badanie przeprowadzone dwukrotnie na tej samej grupie osób) pochodzą z populacji o tej samej średniej.