GRZEGORZ BORUTA

PODSTAWY EKSPLOATACJI URZĄDZEŃ

PODSTAWY DIAGNOSTYKI TECHNICZNEJ

Każde urządzenie posiada **zdeterminowaną strukturę**, charakteryzowaną przez wymiary i wzajemne rozmieszczenie części, rodzaj ich połączeń, sposób współpracy i parametry opisujące ich cechy materiałowe – przez tzw. **parametry struktury**. Struktura ta determinuje charakterystyki ekonomiczne i techniczne urządzenia – jego **stan techniczny**.

W trakcie eksploatacji urządzenia ulegają fizycznemu starzeniu – następuje zmiana wartości ich parametrów struktury, powodowana procesami tarcia, korozji, erozji, zmęczenia itd. materiałów tworzących urządzenie – stan techniczny urządzenia zmienia się.

Aktualne wartości parametrów technicznych urządzenia określają jego aktualny stan techniczny.

Starzenie fizyczne urządzeń zmienia wartości ich charakterystyk eksploatacyjnych: zmniejsza niezawodność, gotowość i efektywność realizacji zadań, zwiększa jednostkowe koszty eksploatacji – pogarsza ich stan techniczny.

Na ogół nie jest możliwe ustalenie ilościowego wpływu poszczególnych czynników powodujących starzenie fizyczne urządzenia na jego stan techniczny →

→ starzenie fizyczne uważa się za proces w pewnej mierze nieregularny i przypadkowy →

→ w podobnych warunkach pracy urządzeń i dla takich samych wartości miary eksploatacji procesy starzenia różnych egzemplarzy mogą mieć różny przebieg →

→ istnieje konieczność okresowych kontroli stanu urządzenia pozwalających na określenie tego stanu, kolejnego terminu kontroli lub terminu i zakresu obsługi / naprawy.

Problematyką badania stanu urządzeń i ich elementów zajmuje się **diagnostyka techniczna**.

Diagnostyka – nauka o środkach i sposobach rozpoznawania stanu obiektów na podstawie charakterystycznych objawów (symptomów) tych stanów (wartości określonych wielkości charakteryzujących te obiekty).

Diagnostyka techniczna – dział nauki o eksploatacji urządzeń obejmujący problemy związane z rozpoznawaniem ich stanu technicznego, w ogólności bez ich demontażu lub przy częściowym ich demontażu nienaruszającym zasadniczych funkcjonalnych połączeń ich elementów.

Weryfikacja części – ocena stanu pojedynczych części (elementów) urządzeń.

Obiektem badań dla diagnostyki technicznej może być:
- urządzenie jako takie,
- jego zespół lub podzespół, a nawet pojedyncze skojarzenie – para kinematyczna,
- system eksploatacji, którego jest częścią.

Diagnozowanie – proces oceny stanu technicznego urządzenia, przyczyn jego wystąpienia oraz resursu dalszej poprawnej pracy. Efektem diagnozowania jest **diagnoza**.

Diagnozowanie może też dotyczyć procesu, w którym urządzenie bierze udział – ocenia się efektywność przebiegu tego procesu.

Diagnozowanie przeprowadza się w chwili uznanej za ważną.

Dziedziny diagnostyki technicznej:

- badanie rzeczywistych urządzeń, w tym:
  - poznanie prawidłowości właściwego funkcjonowania urządzeń,
  - systematyzacja elementów urządzeń i ustalenie związków między nimi,
  - wyróżnienie możliwych stanów technicznych urządzeń,
  - analiza technicznych możliwości kontroli parametrów charakteryzujących stan urządzeń,
  - gromadzenie i analiza informacji dotyczących kosztów badań diagnostycznych;
- badanie modeli diagnozowanych urządzeń i automatyzacja procesów diagnostycznych, w tym:
  - opracowanie metod budowy optymalnych testów diagnostycznych i programów diagnozowania,
  - opis i analiza istniejących systemów diagnostycznych, w tym ocena ich szybkości i niezawodności działania,
  - ocena zasadności i ekonomiczności automatyzacji procesów diagnozowania,
  - opracowanie zaleceń do projektowania urządzeń z uwzględnieniem wymagań diagnostyki.

Pomiędzy dziedzinami diagnostyki istnieją sprzężenia zwrotne. Analiza konkretnych rzeczywistych obiektów służy do budowy i oceny matematycznych modeli diagnostycznych. Rozwiązanie zadań teoretycznych, sformułowanych w odniesieniu do matematycznego modelu diagnostycznego, pozwala na doskonalenie konstrukcji konkretnego obiektu diagnozowania oraz metod jego diagnozowania.

Związane z dziedzinami diagnostyki technicznej jej podstawowe **zadania** to:

- badanie, ustalanie i klasyfikowanie uszkodzeń urządzeń oraz objawów tych uszkodzeń,
- opracowywanie metodyk, metod i aparatury do mierzenia, rejestrowania i odpowiedniego opracowywania wartości cech technicznych,
- ocena, na podstawie uzyskanych wartości cech technicznych, aktualnego stanu technicznego urządzenia, określenie zakresu koniecznych czynności odnawiających jego potencjał eksploatacyjny oraz określenie wartości miary eksploatacji jego dalszej poprawnej pracy.

Pełny proces diagnozowania obiektu obejmuje:

- **genezowanie** – określanie, na podstawie analizy wyników odpowiednich badań, warunków i przyczyn, które spowodowały, powodują lub mogą spowodować istnienie określonego stanu technicznego diagnozowanego obiektu lub zmiany tego stanu,

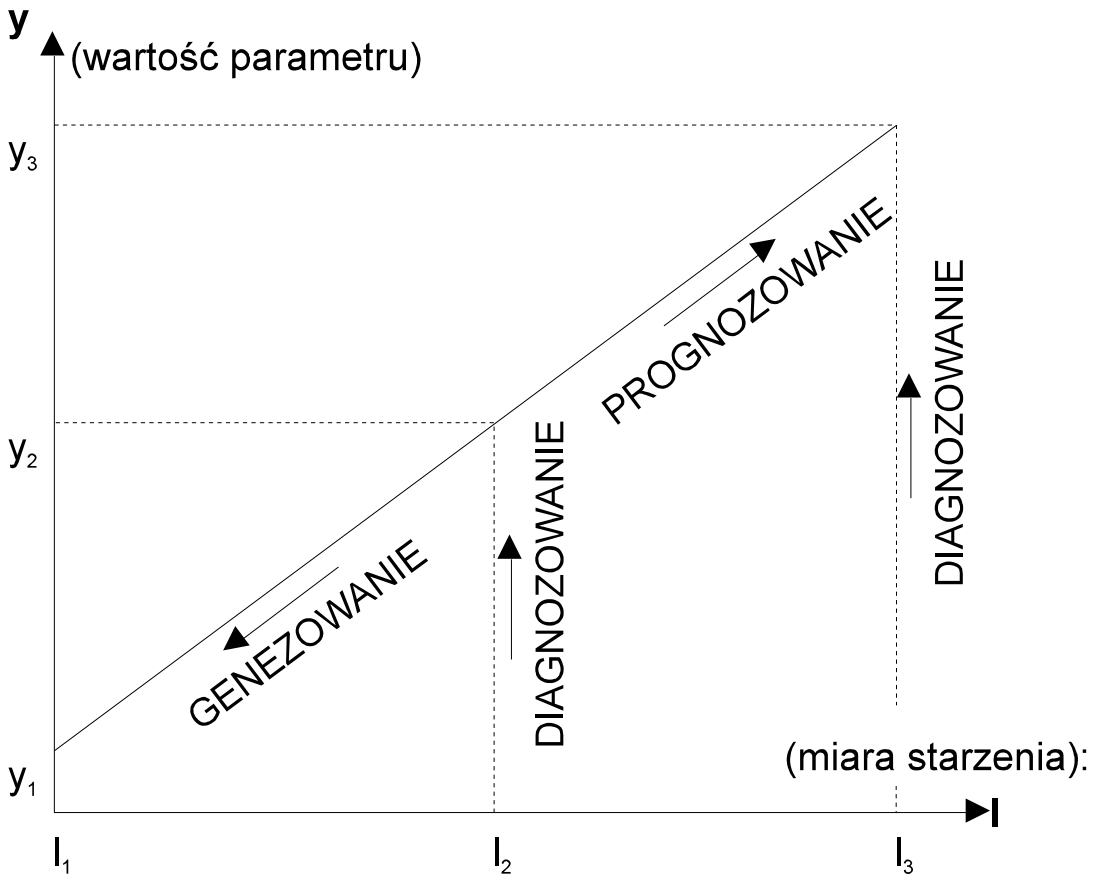
- **diagnozowanie** (właściwe) – wnioskowanie o bieżącym w danej chwili stanie technicznym diagnozowanego obiektu na podstawie analizy wyników badania dokonanego przy zastosowaniu odpowiednich metod oraz środków technicznych; prowadzone permanentnie nazywa się **dozorowaniem** lub **monitorowaniem**,

- **prognozowanie** – przewidywanie stanów technicznych diagnozowanego obiektu w przyszłości na podstawie wyników genezowania i diagnozowania, ustalające właściwości tego obiektu dla określonej przyszłej wartości miary eksploatacji poprzez wyznaczanie prognozowanych wartości cech technicznych i porównanie ich z pewnymi wartościami granicznymi określonymi dla wystąpienia określonego stanu technicznego, a także przewidywanie wartości miary eksploatacji poprawnej pracy tego obiektu do wystąpienia stanu granicznego uniemożliwiającego dalsze jego użytkowanie.

Diagnozowanie stanu technicznego obiektu może być wykonanie jedynie w trakcie jego eksploatacji, podczas której:

– następuje rozwój uszkodzeń i degradacja stanu technicznego tego obiektu w miarę upływu czasu jego życia lub zużywania resursu docelowego,
– rejestruje się wartości wybranych cech technicznych – wartości tzw. **parametrów diagnostycznych**.

Praktyczne zadania prognozowania lub genezowania rozwiązuje się poprzez określenie przebiegów zmian parametrów stanu technicznego diagnozowanego obiektu w funkcji miary eksploatacji, a następnie ich odpowiednią ekstrapolację.



Diagnozowanie składa się z trzech zasadniczych części:

1. pomiarów wartości parametrów diagnostycznych i porównanie ich z odpowiednimi wartościami nominalnymi (granicznymi, progowymi),
2. analizy charakteru i przyczyn powstawania ewentualnego odchylenia od stanu oczekiwanego i zakresu prac obsługowo-naprawczych niwelujących to odchylenie,
3. podania prawdopodobnej wartości miary eksploatacji dalszej poprawnej pracy urządzenia (pozostałego resursu).

Diagnozowanie urządzeń dostarcza systemowi eksploatacji informacji o ich stanie technicznym – jest częścią tego systemu.

Obiektem badań diagnostyki technicznej (jako nauki) i diagnozowania (jako procesu stawiania diagnozy) może być każdy obiekt techniczny (urządzenie), który w skrajnym przypadku może znajdować się w dwu różnych, wzajemnie wykluczających się stanach technicznych:
– w stanie zdolności do wykonywania pracy, czyli tzw. **stanie zdatności**,
– w stanie niezdolności do wykonywania pracy, czyli tzw. **stanie niezdatności**.

Jest to **dwuklasowy** podział stanów technicznych obiektu (**zdatny/niezdatny**).

Istnienie przynajmniej dwóch stanów technicznych wynika z potrzeby uzyskania jednoznacznej odpowiedzi na pytania:
- w jakim stanie znajduje się urządzenie w danej chwili?
- czy urządzenie wykona postawione jej zadanie czy nie?

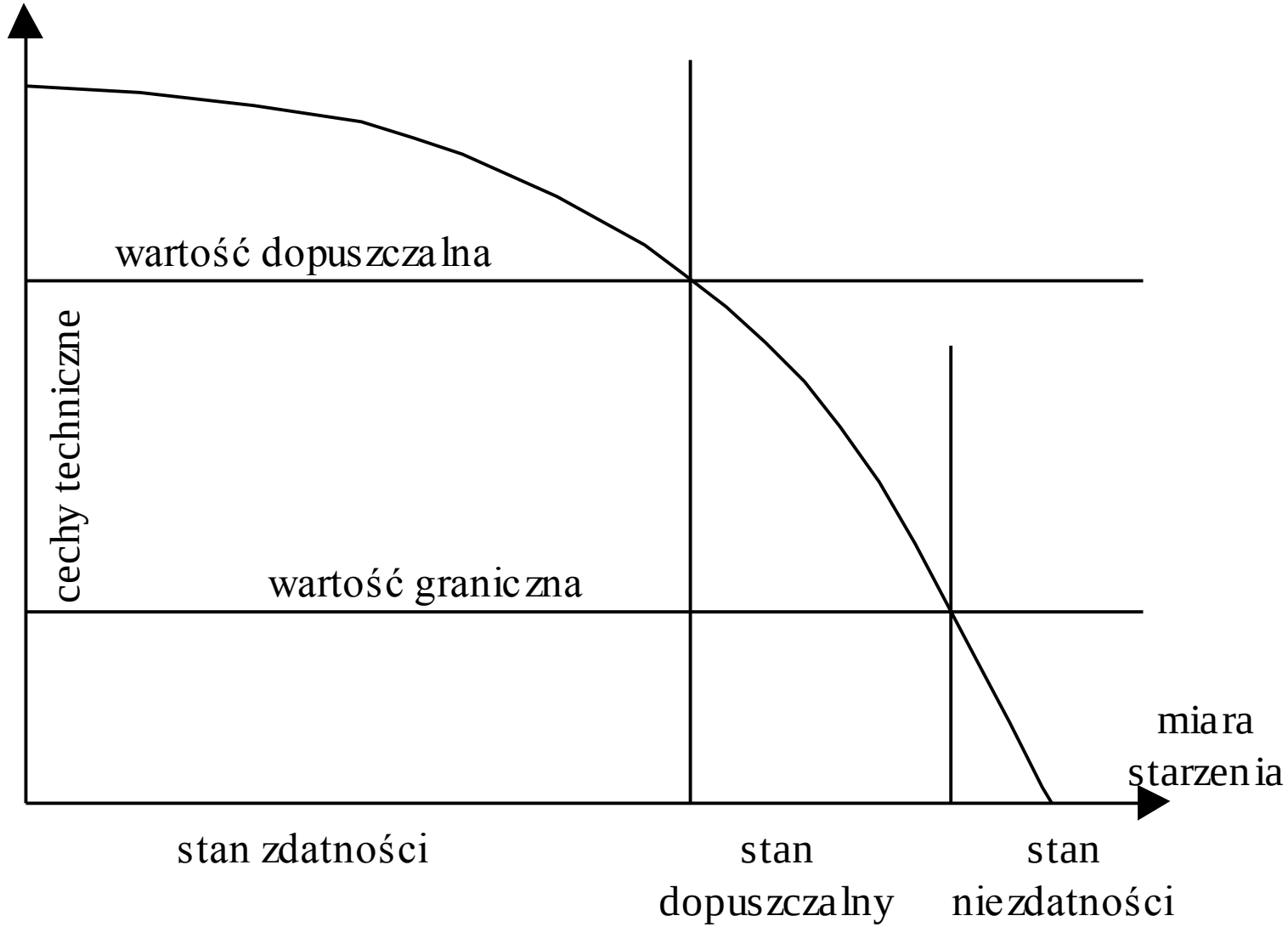
Wartość graniczna – progowa wartość cech technicznych dzieląca zakres możliwych wartości tych cech na przedziały odpowiednie dla stanu zdatności i stanu niezdatności.

Eksploatacja urządzenia i **procesy wymuszające jego starzenie fizyczne** → najczęściej **stopniowa zmiana wartości cech technicznych urządzenia** – **stopniowe pogarszanie się jej stanu technicznego** – stopniowa zmiana wartości cech technicznych od odpowiadających **stanowi zdatności** w stronę **wartości granicznej** i dalej do odpowiadających **stanowi niezdatności**.
Uszkodzenie (**awaria**) urządzenia – przejście wartości cech technicznych przez wartość graniczną ze strony stanu zdatności na stronę stanu niezdatności.
Naprawa urządzenia – przejście wartości cech technicznych przez wartość graniczną ze strony stanu niezdatności na stronę stanu zdatności.

Gdy wymagane jest uprzedzanie o zbliżającym się uszkodzeniu wprowadza się:
- **trójklasowy** podział stanów technicznych,
- dodatkową progową wartość cech technicznych – **wartość alarmową**,
- dodatkowy stan techniczny – **stan dopuszczalny**.

Wartość alarmowa powinna należeć do przedziału odpowiedniego dla stanu zdatności i być odpowiednio oddalona od wartości granicznej.

Stopniowa zmiana stanu technicznego urządzenia – zmiana wartości cech technicznych:
- najpierw urządzenie jest w **stanie zdatności,**
- potem stopniowo zmieniają się wartości cech technicznych i w pewnej chwili przekraczają one **wartość alarmową** – urządzenie znajduje się w **stanie dopuszczalnym** (zbliża się chwila wystąpienia uszkodzenia – sygnał dla kierownictwa eksploatacji o konieczności wykonania obsługi urządzenia w celu uniknięcia jego uszkodzenia),
- dalsza eksploatacja nieobsłużonego urządzenia i postępujące zmiany starzeniowe powodują dalszą zmianę wartości cech technicznych w kierunku **wartości granicznej** aż do jej osiągnięcia (przekroczenia) – urządzenie osiąga **stan niezdatności** (ulega uszkodzeniu – sygnał dla kierownictwa eksploatacji o konieczności wycofania urządzenia z systemu użytkowania).


wartość dopuszczalna
cechy techniczne
wartość graniczna
miara
starzenia
stan zdatności
stan
dopuszczalny
stan
niezdatności

Urządzenia w ogólności są obiektami technicznymi złożonymi z wielu zespołów, mechanizmów i części – tempo starzeniowych zmian cech technicznych różnych części może być różne → cechy techniczne pewnych części osiągają swoje wartości graniczne a innych swoje wartości alarmowe już wtedy, gdy cechy techniczne pozostałych części jeszcze nie.

Różne części spełniają różne funkcje. Na podstawie analizy wagi części urządzenia pod kątem wypełniania przez niego zadań założonych w fazie konstruowania i/lub zadań realizowanych w fazie eksploatacji, **elementy struktury urządzenia** dzieli się na:

– **zasadnicze**, zabezpieczające normalne wypełnianie podstawowych funkcji roboczych urządzenia,
– **drugorzędne**, zapewniające wygodę eksploatacji, estetykę itp.

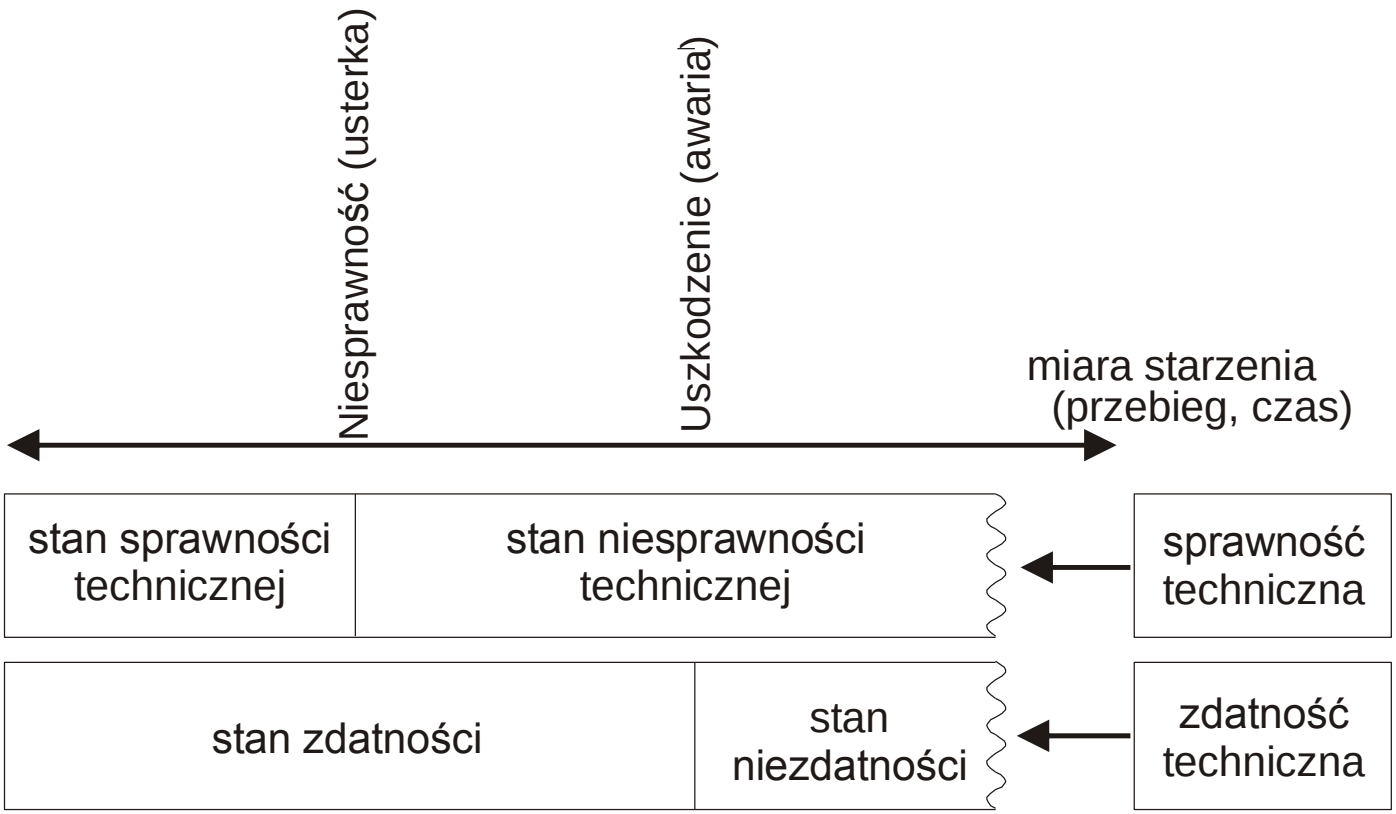
Dla urządzeń jako zbiorów części składowych przy dwuklasowym podziale stanów technicznych pojedynczych części wyróżnia się:

– **stan sprawności technicznej** – żadna cecha techniczna urządzenia nie osiągnęła jeszcze swojej wartości granicznej (*właściwości techniczno-eksploatacyjne urządzenia odpowiadają założonym podczas konstruowania i produkcji*),

– **stan niesprawności technicznej** – cechy techniczne wszystkich zasadniczych elementów struktury urządzenia nie osiągnęły jeszcze swoich wartości granicznych, ale przynajmniej jednego drugorzędnego elementu struktury już tak (*urządzenie może nadal wypełniać swoje funkcje robocze (jest zdatne) choć pewne właściwości techniczno-eksploatacyjne nie w pełni odpowiadają założonym podczas konstruowania i produkcji*) ***lub*** cechy techniczne przynajmniej jednego zasadniczego elementu struktury urządzenia już osiągnęły swoje wartości graniczne (*jest niezdatne*),

– **stan zdatności** – cechy techniczne wszystkich zasadniczych elementów struktury urządzenia nie osiągnęły jeszcze swoich wartości granicznych, a cechy techniczne drugorzędnych elementów struktury są nieistotne (*urządzenie może nadal wypełniać swoje funkcje robocze, choć pewne właściwości techniczno-eksploatacyjne mogą już nie w pełni odpowiadać założonym podczas konstruowania i produkcji*),

– **stan niezdatności** – cechy techniczne przynajmniej jednego zasadniczego elementu struktury urządzenia osiągnęły swoje wartości graniczne (*urządzenie nie może już wypełniać swoich funkcji roboczych*).



W przypadku konkretnych urządzeń zaliczenie poszczególnych stanów do odpowiednich klas może być **subiektywne** – elementy struktury urządzenia mają różne przeznaczenie i uszkodzenie jednego z nich w zależności od warunków eksploatacji (*używanie urządzenia w miejscach publicznych, na zamkniętych terenach prywatnych, na placach budowy, w warunkach pokojowych lub wojennych*) może powodować zaliczenie konkretnych elementów struktury jako zasadniczych lub drugorzędnych elementów struktury i w konsekwencji określenie stanu technicznego urządzenia raz jako stanu zdatności, a raz jako stanu niezdatności.

Bieżące wartości parametrów struktury (dla bieżącej wartości miary eksploatacji) określają stan techniczny urządzenia (są parametrami stanu technicznego). Strukturę urządzenia stanowi zbiór tworzących ją elementów konstrukcyjnych, uporządkowanych w ściśle ustalony sposób w celu wypełniania określonych funkcji:

$$\mathbf{U} = \{u_i\}, \quad i = \overline{1,n}$$ (n – liczba parametrów struktury).

Niestety na ogół nie jest możliwe zmierzenie wartości parametrów struktury urządzenia bez jego demontażu – **pomiar wartości parametrów struktury nie leży w dziedzinie diagnostyki technicznej**, tylko weryfikacji części.

Ale...

Podczas funkcjonowania urządzeń realizowane są różne procesy fizyczne i chemiczne przetwarzania energii i masy – **procesy wyjściowe**. Procesy te na ogół można obserwować i mierzyć z zewnątrz urządzenia.

Wyróżnia się procesy wyjściowe:

- **robocze** (**zasadnicze**), wynikające bezpośrednio z realizacji funkcji użytkowych urządzenia (dla tłokowego silnika spalinowego są to spalanie paliwa, wytwarzanie energii ruchu obrotowego wału korbowego, wymiana ciepła, wydalanie produktów spalania paliwa),
- **towarzyszące** (**resztkowe**), powstające jako wtórny efekt zachodzenia zasadniczych procesów roboczych (tarcie, procesy termiczne, drgania, hałas, pulsacje mediów roboczych, zjawiska świetlne, zapachy itp.),
- będące odpowiedzią na dostarczanie do urządzenia energii i masy wyłącznie w celu jego diagnozowania (np. tzw. nieniszczącego, wykorzystującego ultradźwięki, promieniowanie rentgenowskie, znaczenie substancjami promieniotwórczymi).

Procesy wyjściowe można opisać zbiorem **parametrów wyjściowych (symptomów)**:

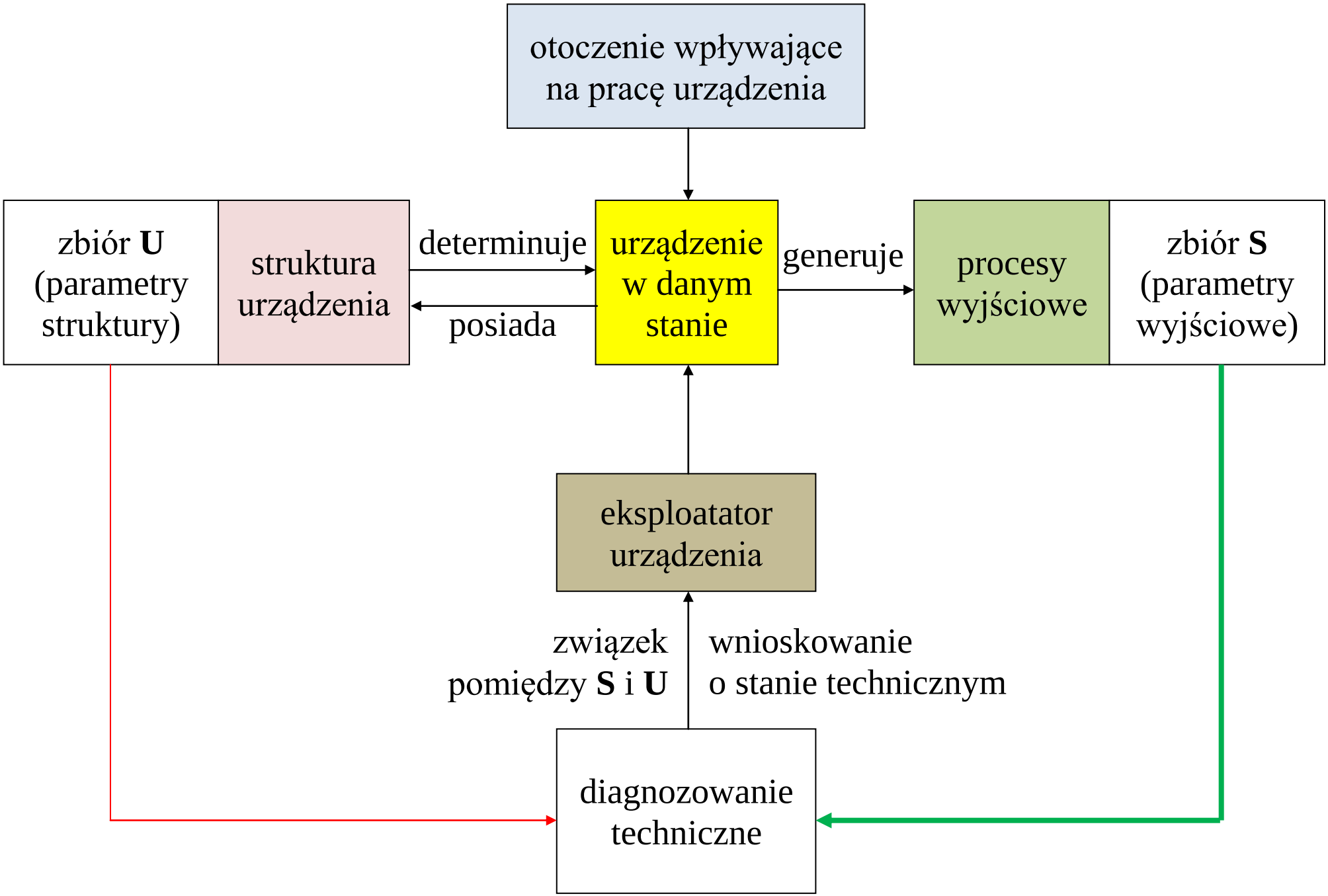
$$\mathbf{S} = \{s_j\}, \quad j = \overline{1, m} \qquad (m \text{ – liczba parametrów wyjściowych}).$$

Przebieg procesów wyjściowych jest uzależniony od stanu technicznego urządzenia (ich wartości są objawami (symptomami) stanu) → wartości parametrów wyjściowych powinny zmieniać się wraz z zmianą stanu technicznego urządzenia, podobnie jak parametry struktury.

Występuje więc następująca sytuacja: parametry struktury determinują stan techniczny urządzenia i w ogólności nie są bezpośrednio dostępne, ale stan ten jest odzwierciedlany w parametrach wyjściowych, w ogólności bezpośrednio dostępnych.

Wzajemny **związek parametrów struktury i parametrów wyjściowych** urządzenia pozwala w określonych warunkach **traktować parametry wyjściowe jako parametry stanu technicznego** urządzenia mierzone bez jej demontażu, czyli jako **parametry leżące w dziedzinie diagnostyki technicznej**.

Istota diagnostyki technicznej: **stan techniczny urządzenia** (wartości parametrów struktury) **ocenia się** nie na podstawie pomiarów parametrów struktury, ale **na podstawie pomiarów parametrów wyjściowych i znajomości wzajemnych zależności między parametrami struktury i parametrami wyjściowymi**.


otoczenie wpływające na pracę urządzenia
zbiór **U** (parametry struktury)
struktura urządzenia
determinuje
posiada
urządzenie w danym stanie
generuje
procesy wyjściowe
zbiór **S** (parametry wyjściowe)
eksploatator urządzenia
związek pomiędzy **S** i **U**
wnioskowanie o stanie technicznym
diagnozowanie techniczne

Generowanie procesów wyjściowych w wyniku funkcjonowania urządzenia
jest uzewnętrznieniem wzajemnych energetycznych oddziaływań
wewnątrz urządzenia i urządzenia z jej otoczeniem (urządzenie jest przetwornikiem energii).

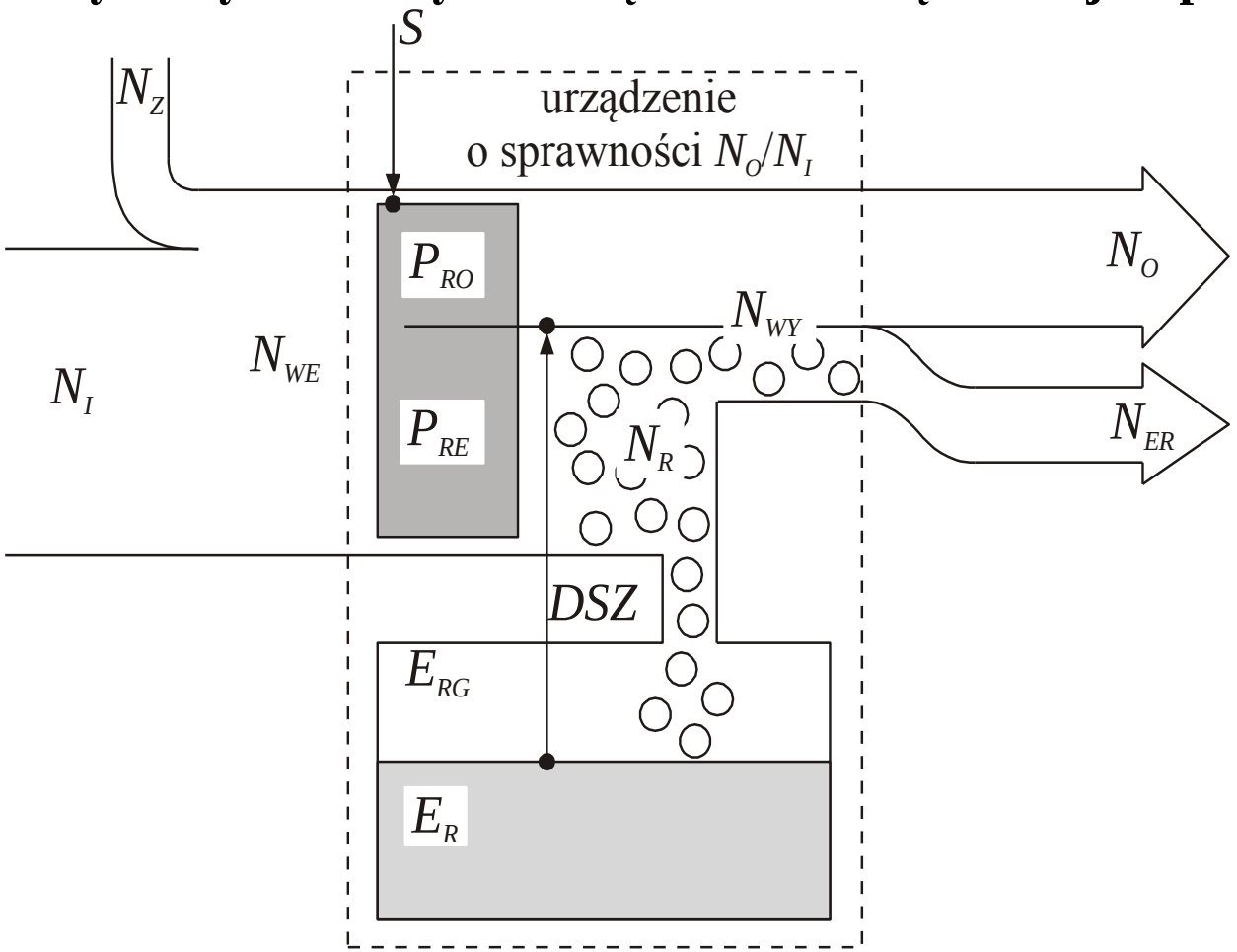
↓

Funkcjonowanie urządzenia można przedstawić jako
proces **kodowania informacji o stanie jego elementów**.
Informacje te przekazywane są w postaci **sygnałów diagnostycznych**.

↓

Diagnozowanie należy traktować jako
logiczny proces pozyskiwania i przetwarzania (rozkodowywania) informacji.

Zużyciowy model życia urządzenia - urządzenie jest przetwornikiem energii



S - sterowanie
N_I - moc zasilania
N_Z - moc zakłóceń
$N_{WE} = N_I + N_Z$ - moc wejściowa
P_{RO} - procesy robocze
P_{RE} - procesy resztkowe
N_O - moc użyteczna
N_{ER} - moc emisji procesów resztkowych do otoczenia (drgań, hałasu, ciepła, fal elektromagnetycznych itp.)
N_R - moc rozproszona w urządzeniu
$N_{WY} = N_O + N_{ER}$ - moc wyjściowa
E_R - energia rozproszona w urządzeniu
E_{RG} - energia graniczna, przy której następuje awaria urządzenia
DSZ - destrukcyjne sprzężenie zwrotne

Dla danej wartości Θ miary eksploatacji urządzenia, diagnozowanie techniczne urządzenia polega na ocenie jego stanu technicznego ściśle związanego z ilością energii $E_R(\Theta)$ pochłoniętej przez jego elementy na podstawie wartości parametrów wyjściowych, ściśle związanych z mocą procesów wyjściowych z urządzenia do środowiska $N_O(\Theta)$ i $N_{ER}(\Theta)$.

W procesie diagnozowania technicznego urządzeń nie można zastosować wszystkich parametrów wyjściowych.

Aby można było wykorzystać parametr wyjściowy do diagnozowania (aby można go było nazwać **parametrem diagnostycznym**) musi on być **jednoznaczny**, **dostępny** i **dostatecznie wrażliwy na zmianę stanu technicznego urządzenia**.

Jednoznaczność – jednej, ściśle określonej wartości parametru struktury (lub wartości miary eksploatacji) urządzenia musi odpowiadać jedna, ściśle określona wartość parametru diagnostycznego.

Ze względu na błędy, jakimi obarczone są zarejestrowane wartości parametrów diagnostycznych, warunek jednoznaczności nałożony jest na pewną funkcję aproksymującą zmiany wartości parametru diagnostycznego w funkcji parametru struktury (lub miary eksploatacji) urządzenia. Źródła błędów:

- błąd zastosowanej metody pomiarowej,
- błąd przyrządu pomiarowego,
- błędny odczyt wskazań przyrządu pomiarowego przez człowieka – operatora przyrządu,
- różne warunki zewnętrzne (np. atmosferyczne czy stan drogi podczas badań drogowych), w jakich prowadzone są badania diagnostyczne, tj. pracują badane urządzenie, przyrząd pomiarowy (diagnostyczny) i diagnosta.

Dostępność – parametr diagnostyczny powinien być dostępny, łatwy do zmierzenia.

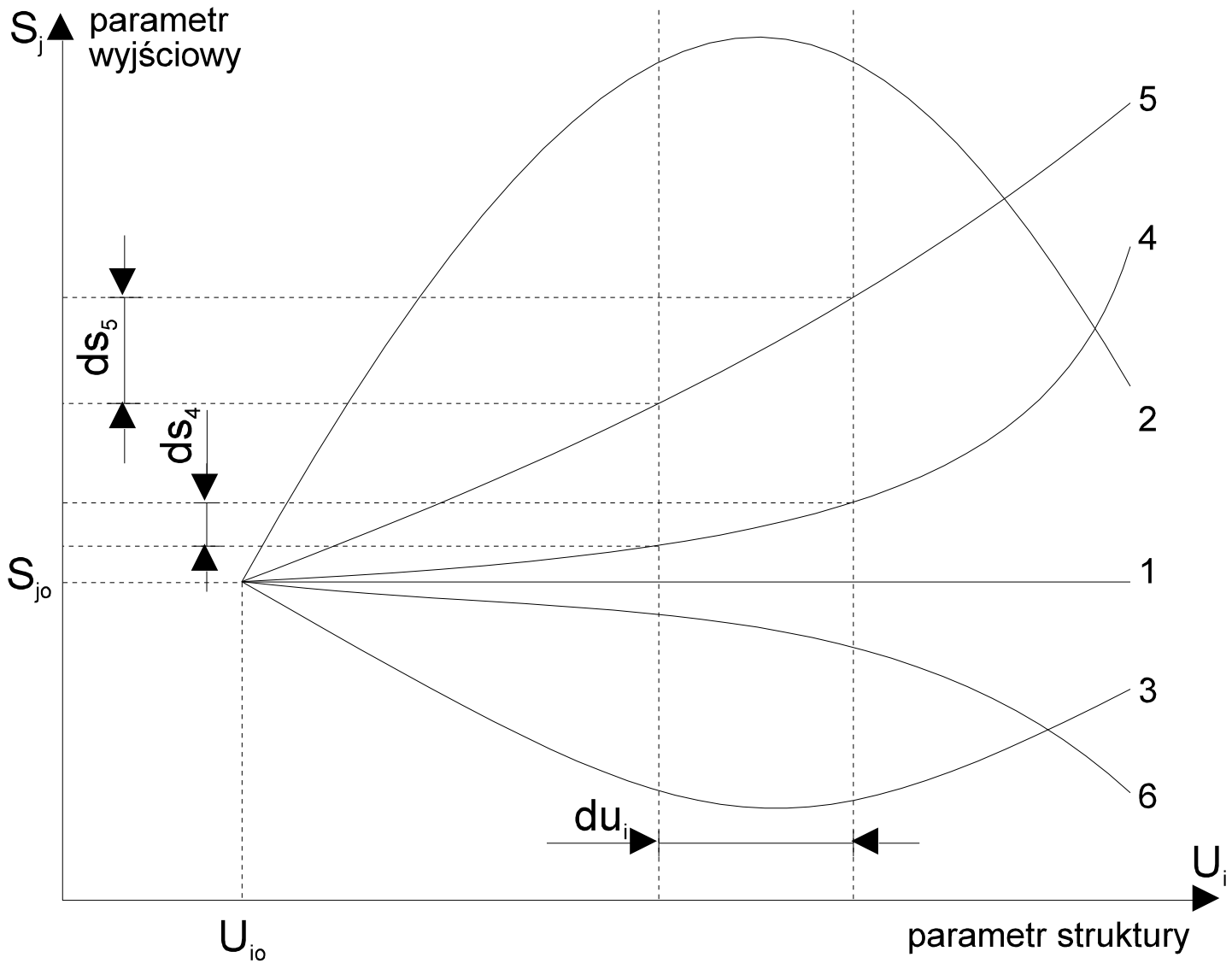
Parametr diagnostyczny ma być dostępny w sensie „widoczny”, „łatwy do wyodrębnienia” spośród wszystkich parametrów wyjściowych.

Musi istnieć przyrząd kontrolno–pomiarowy, za pomocą którego będzie można zarejestrować ten parametr wyjściowy oraz podać jego konkretną wartość w określonych jednostkach miary.

Względność dostępności – dla diagnostów niedysponujących odpowiednim sprzętem diagnostyczno-pomiarowym niedostępne są pewne parametry, dostępne innym diagnostom, którzy odpowiedni sprzęt diagnostyczno-pomiarowy posiadają i się nim posługują.

Dostateczna wrażliwość na zmianę stanu technicznego urządzenia – najlepszym parametrem diagnostycznym jest ten, którego zmiany w miarę zmian stanu technicznego są najszybsze.

Można wtedy dość szybko i łatwo oraz stosunkowo tanio (stosując tańsze urządzenia diagnostyczne o stosunkowo małej rozdzielczości) zauważyć niewielkie zmiany stanu technicznego i w efekcie podejmować odpowiednie decyzje eksploatacyjne.


S_j
parametr
wyjściowy
5
4
2
ds_5
ds_4
S_jo
1
3
6
du_i
U_i
U_io
parametr struktury

Podział parametrów diagnostycznych wg charakteru związków między nimi:

- **niezależne** (niezależnie od innych wskazują na zmianę stanu technicznego konkretnego elementu diagnozowanego urządzenia),
- **zależne** (zmianę stanu technicznego urządzenia można określić dopiero za pomocą kilku parametrów),

oraz wg pojemności i charakteru informacji na:

- **szczegółowe** (sygnalizujące zmianę stanu technicznego konkretnego pojedynczego elementu urządzenia),
- **ogólne (redundantne)** (charakteryzujące stan techniczny urządzenia jako całości i informacje o stanie konkretnych elementów trzeba wyodrębniać z zarejestrowanego sygnału).

W zależności od grupy, do której można zaliczyć konkretny parametr diagnostyczny należy stosować mniej lub bardziej skomplikowane przyrządy diagnostyczne, metody pomiarowe i algorytmy decyzyjne o faktycznym stanie technicznym urządzenia.

Stany techniczne, w których może znaleźć się urządzenie, tworzą zbiór

$$\mathbf{W} = \{w_l\}, \quad l = \overline{1,k}$$ (k – liczba możliwych stanów technicznych urządzenia).

Konkretny stan techniczny w_l urządzenia jest zdeterminowany przez n parametrów struktury u_i:

$$w_l = \varphi(\mathbf{U}_l) = \varphi(\{u_{li}\})$$

i objawia się w m parametrach wyjściowych s_j.

Diagnozowanie techniczne nie polega na mierzeniu parametrów struktury, ale na mierzeniu parametrów diagnostycznych (odpowiednio wybranych parametrów wyjściowych) i zadanie identyfikacji stanu technicznego trzeba rozwiązać zastępując wielkości u_{li} wielkościami s_{lj}:

$$w_l = \psi(\mathbf{S}_l) = \psi(\{s_{lj}\}).$$

Należy więc znać relacje pomiędzy parametrami struktury i parametrami diagnostycznymi:

$$s_{lj} = f(\{u_{li}\}) \quad \text{i} \quad u_{li} = f^{-1}(\{s_{lj}\}).$$

Postać funkcji $u_{li} = f^{-1}(\{s_{lj}\})$ próbuje się ustalać podczas cechowania metod diagnostycznych. Uzyskany układ n równań tego typu można rozpatrywać jako odwzorowanie przestrzeni parametrów diagnostycznych o współrzędnych s_1, s_2, ... s_m w przestrzeni parametrów struktury opisanej współrzędnymi u_1, u_2, ... u_n. Odwzorowanie to obrazuje proces stawiania diagnozy – odwzorowanie odwrotne $s_{lj} = f(\{u_{li}\})$ – pracę urządzenia.

Praktyczne posługiwanie się takimi układami równań jest bardzo skomplikowane. Jeżeli jest to możliwe, obiekt dzieli się umownie na zespoły i równania tworzy się oddzielnie dla każdego z nich. Największe uproszczenie uzyska się wtedy, gdy każdy parametr struktury można związać tylko z jednym parametrem diagnostycznym. Wtedy układ równań rozpada się na n niezależnych równań typu $u_{li} = f^{-1}(s_{lj})$.

Eksperymenty diagnostyczne stosowane podczas cechowania metod diagnostycznych polegają na obserwacji parametrów wyjściowych przy zadanych parametrach struktury. Wyznacza się wtedy relacje $s_{lj} = f(\{u_{li}\})$ (**proste modele diagnostyczne**) i dopiero później próbuje się znaleźć postać wymaganych relacji $u_{li} = f^{-1}(\{s_{lj}\})$ (**odwrotnych modeli diagnostycznych**), co udaje się dość rzadko.

W efekcie, w procesie diagnozowania najczęściej nie wyznacza się wartości parametrów struktury u_{li} i **stan techniczny** w_l **obiektu jest określany bezpośrednio na podstawie zmierzonych wartości parametrów diagnostycznych** s_{lj} za pomocą zależności

$$w_l = \psi(\mathbf{S}_l) = \psi(\{s_{lj}\}).$$

Model:

– uproszczony obraz rzeczywistości przedstawiony w sformalizowanej formie,
– dający się pomyśleć lub materialnie wykonać układ odtwarzający przedmiot badania i zastępujący go tak, że badanie modelu dostarcza nowej informacji o przedmiocie modelowanym – narzędzie do opisywania i badania modelowanego przedmiotu badania w różnych warunkach.

Modele:

- **sformalizowane**, tworzone z zachowaniem ścisłych i jednoznacznych reguł:
  - **analogowe**, w których odpowiednie elementy modelowanego przedmiotu badania reprezentowane są przez inne wielkości fizyczne (np. symulacja systemu sieci hydraulicznej za pomocą adekwatnego układu elektronicznego),
  - **symboliczne**, w których poszczególne parametry i właściwości modelowanego przedmiotu badania są reprezentowane przez symbole. Do modeli symbolicznych należą **modele słowne**, **modele graficzne** i **modele matematyczne**,
- **intuicyjne** (np. ekspertowe), opierające się na dedukcjach i ocenach myślowych, które zawierają zawsze dużą dozę niejednoznaczności (z wiedzą niepełną i rozmytą).

Największe znaczenie mają modele matematyczne, które opisują rzeczywistość za pomocą równań i nierówności matematycznych, związków logicznych itp..

Modele diagnostyczne urządzenia opisują wzajemne związki między zmiennymi:
- parametrami diagnostycznymi i parametrami struktury,
- parametrami diagnostycznymi i stanami technicznymi.

Cele tworzenia modeli diagnostycznych:

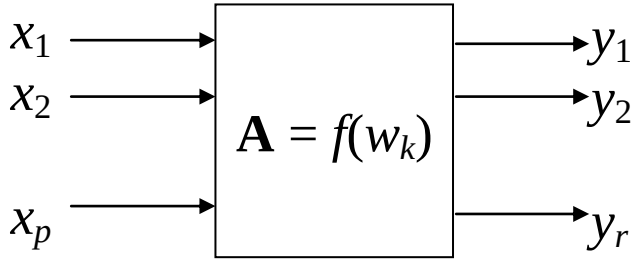
– diagnozowanie – model jest podstawą ustalenia algorytmu diagnozowania i określenia stanu diagnozowanego urządzenia (systemu),
– prognozowanie – model służy do przewidywania rozwoju zmian stanu,
– genezowanie – model służy do ustalenia przyczyn stanu zaistniałego,
– projektowanie – model służy do optymalizacji struktury i parametrów projektowanego systemu (np. obsługiwania / naprawiania zawierającego diagnozowanie),
– sterowanie – model służy do podejmowania decyzji w działającym systemie.

Model funkcjonalny – najprostszy model diagnostyczny – graficzne przedstawienie urządzenia jako zbioru bloków odpowiadających jego elementom funkcjonalnym.

Wejścia (też zakłócenia) to bodźce.

Wyjścia (też parametry diagnostyczne) to reakcje.

Funkcja przejścia **A** określa związki pomiędzy bodźcami i reakcjami: $\mathbf{Y} = \mathbf{A} \cdot \mathbf{X}$.
Działanie funkcji **A** zależy od stanu technicznego urządzenia – obserwując reakcje na zadane bodźce można określić sposób działania bloku modelu (stan techniczny elementu urządzenia).



Głównym kryterium doboru bloków funkcjonalnych są wymagania dotyczące rozróżniania elementów podczas diagnozowania – są one uzależnione od przyjętej techniki i aparatury diagnostycznej.

Model topologiczny – bardziej ogólny od funkcjonalnego – pewien abstrakcyjny opis urządzenia dokonany za pomocą kategorii pojęciowych stosowanych w topologii (jednej z dziedzin matematyki). Najczęściej ma postać skończonego grafu zorientowanego, którego wierzchołki stanowią zbiór właściwości urządzenia istotnych dla jego normalnego funkcjonowania, a łuki („topologie") wynikają ze związków pomiędzy tymi właściwościami. Wierzchołkami mogą być stany techniczne (parametry struktury) urządzenia i parametry wyjściowe, w tym diagnostyczne, a łukami wzajemne ich powiązania. Dysponując pełnym modelem topologicznym wykonanym w fazie projektowania urządzenia można określić zbiór rozpoznawalnych usterek urządzenia i sposób ich objawiania się w parametrach wyjściowych oraz ich oddziaływanie na inne właściwości urządzenia i procesy w nim zachodzące.

Model analityczny – urządzenie w stanie technicznym określonym przez zbiór parametrów struktury **U** jest przekształtnikiem bodźców (zbiór **X**) na reakcje (zbiór **Y** (w tym na parametry wyjściowe **S**: $\mathbf{Y} \supset \mathbf{S}$)) i poszukuje się **jawnych** postaci funkcji przekształcających. Obserwując urządzenie dla określonych wartości miar eksploatacji tworzących uporządkowany zbiór $\mathbf{T} = t_e, e = \overline{0, aw}$ (t_0 – „chwila” wyprodukowania urządzenia, t_{aw} – „chwila” wystąpienia niezdatności) wyróżnia się:

1. ***model w przestrzeni stanu***: $\mathbf{X} \times \mathbf{T} \rightarrow \mathbf{U}$ – odwzorowanie bodźców **X** w parametry struktury **U** urządzenia. Podstawowe równanie tego modelu uzależnia prędkość zmian stanu technicznego od aktualnego stanu technicznego i aktualnej wartości bodźców, np. $\frac{\mathrm{d}u(t)}{\mathrm{d}t} = au(t) + bx(t)$, gdzie a i b są pewnymi współczynnikami, zależnymi od stanu technicznego elementów struktury urządzenia. Dla jednoznacznego ustalenia funkcji $u(t)$ należy znać warunki początkowe, tzn. wartość i pochodną funkcji w chwili $t = 0$, określające stan początkowy urządzenia.

2. ***model wejściowo-wyjściowy*** **(*„czarnej skrzynki”*)**: $\mathbf{X \times T \rightarrow Y}$ – badanie reakcji **Y** (np. parametrów wyjściowych **S**) urządzenia na bodźce **X**. W opisie tym nie wprowadza się w sposób jawny parametrów struktury, lecz formułuje równania uzależniające od siebie bodźce i reakcje urządzenia. Działanie urządzenia może być opisane np. równaniami różniczkowymi

$$a_n \frac{\mathrm{d}^n y(t)}{\mathrm{d}t^n} + \ldots + a_2 \frac{\mathrm{d}^2 y(t)}{\mathrm{d}t^2} + a_1 \frac{\mathrm{d}\, y(t)}{\mathrm{d}t} + a_0 y(t) = b_0 x(t) + b_1 \frac{\mathrm{d}\, x(t)}{\mathrm{d}t} + b_2 \frac{\mathrm{d}^2 x(t)}{\mathrm{d}t^2} + \ldots + b_m \frac{\mathrm{d}^m x(t)}{\mathrm{d}t^m}$$

gdzie a i b są pewnymi współczynnikami, zależnymi od stanu technicznego elementów struktury urządzenia.

3. ***kompleksowe ujęcie diagnozowania*****: $\mathbf{X \times (U \times T) \rightarrow Y}$** – badanie reakcji **Y** (np. parametrów wyjściowych **S**) urządzenia na bodźce **X** w warunkach bieżąco zmieniającego się stanu **U** tego urządzenia. Praca urządzenia jest opisana równaniem wyjścia, na przykład: $y(t) = ax(t) + bu(t)$, gdzie a i b są pewnymi współczynnikami, zależnymi od stanu technicznego elementów struktury urządzenia.

4. ***typowe zadanie diagnostyki*****: $\mathbf{U \times T \rightarrow Y}$** – badanie reakcji **Y** (np. parametrów wyjściowych **S**) urządzenia na jego stan **U** w warunkach niezmiennych bodźców (dla zachowania jednoznaczności diagnozowania i zredukowania liczby niewiadomych). Jest to już opisany prosty model diagnostyczny.

Część bodźców stanowi zewnętrzne zakłócenia pracy urządzenia.
Wszystkie badania diagnostyczne odbywają się w pewnych losowych warunkach,
których dokładne ustalenie nie jest możliwe.
Sygnał diagnostyczny podlega też zakłóceniom w urządzeniach pomiarowych.

↓

Określenie jawnych postaci przedstawionych powyżej funkcji jest w zasadzie niemożliwe
i częściej stosuje się inne modele matematyczne:
regresyjne, niejawne typu obrazu i probabilistyczne.

Model regresyjny – otrzymany na drodze obróbki statystycznej (regresji) zmian wartości parametrów wyjściowych w funkcji miary eksploatacji albo zmian parametrów struktury uzyskanych w wyniku przeprowadzenia eksperymentu diagnostycznego. Wnioskowanie diagnostyczne polega na porównywaniu bieżącej wartości parametru diagnostycznego z jego wartością alarmową i graniczną oraz na ocenie zapasu zdatności na podstawie położenia pomiaru diagnostycznego na krzywej regresji.
Podstawowe modele regresyjne urządzeń:

1. **prosty model diagnostyczny**: $s_j = f_j(u_1,...,u_n)$,
2. **odwrotny model diagnostyczny** $u_i = g_i(s_1,...,s_m)$,
3. **krzywe życia** $s_j = h_j(l)$, l - miara eksploatacji (czas, motogodziny, przebieg itp.).

Postać tych funkcji może być różna. Najlepsza jest ta, dla której wartość współczynnika korelacji z danymi z eksperymentu diagnostycznego jest najwyższa.

Przykładowe modele regresyjne typu prostego modelu diagnostycznego:

- potęgowy: $s^p = s_0 + \alpha^p u^{\beta^p}$, s_0 - wartość początkowa, α^p - dowolne, $\beta^p > 0$, przy czym dla: $\beta^p < 1$ - model pierwiastkowy, $\beta^p = 1$ - model liniowy, $\beta^p > 1$ - model potęgowy,
- hiperboliczny: $s^h = \left(s_0 + \alpha^h u\right)^{-\beta^h}$, α^h, $\beta^h > 0$,
- wykładniczy: $s^w = s_0 \exp\left(\alpha^w u^{\beta^w}\right)$, α^w - dowolne, $\beta^w > 0$, przy czym dla: $\alpha^w < 0$ - model Weibulla, $\alpha^w > 0$ - model wykładniczy,
- Frecheta rosnący: $s^{F1} = s_0 + \exp\left(\alpha^{F1} u^{\beta^{F1}}\right)$, α^{F1}, $\beta^{F1} < 0$,
- Frecheta malejący: $s^{F2} = s_0 - \exp\left(\alpha^{F2} u^{\beta^{F2}}\right)$, α^{F2}, $\beta^{F2} < 0$,
- logarytmiczny: $s^l = s_0 + \alpha^l \ln\left(1 + \beta^l u\right)$, α^l, β^l - dowolne, przy czym dla $\beta^l < 0$ trend istnieje tylko dla $u < \frac{-1}{\beta^l}$.

Najpopularniejsza metoda wyznaczania funkcji regresji to **metoda najmniejszych kwadratów**, stosowana najczęściej dla modeli liniowych. Aby z niej skorzystać dla modeli nieliniowych można je wcześniej „uliniowić” poprzez odpowiednie przekształcenia, np. model potęgowy można przekształcić do postaci $\ln\left(s^p - s_0\right) = \beta^p \cdot \ln u + \ln\left|\alpha^p\right|$, czyli postaci typu $y = ax + b$.

Rozpoznanie obrazu – dla urządzeń złożonych, o dużej liczbie możliwych stanów technicznych i dużej liczbie stosowanych parametrów diagnostycznych (m), często tylko się przypuszcza, że w danym stanie technicznym wartości parametrów diagnostycznych powinny mieć jakąś wartość, lecz nie wiadomo dlaczego (jakie jest powiązanie z parametrami struktury) ani kiedy (jakie jest powiązanie z miarą eksploatacji) to nastąpi – wiadomo tylko jak powinno „wyglądać" urządzenie w stanie zdatności lub niezdatności. Każde urządzenie opisane jest w chwili diagnozowania zbiorem parametrów diagnostycznych, stanowiących punkt lub obraz w m-wymiarowej przestrzeni parametrów diagnostycznych. Diagnozowanie polega na ocenie „wyglądu" (położenia) urządzenia w tej przestrzeni.

Metody rozpoznawania można podzielić na:

- **metody odległościowe**, budujące funkcję przynależności na różnych miarach odległości (Euklidesa, Hamminga itp.) rozpoznawanego obrazu od obrazów ze zbiorów wzorcowych,
- **metody dzielenia przestrzeni parametrów diagnostycznych** na podprzestrzeń stanu zdatnego, dopuszczalnego i niezdatnego za pomocą hiperpowierzchni decyzyjnych (dla przypadków jednowymiarowych jest jednoznaczne ze znajdowaniem wartości alarmowych i granicznych).

Probabilistyczny model diagnostyczny opiera się na określeniu prawdopodobieństw warunkowych dwu wykluczających się zdarzeń:

– prawdopodobieństwa $P(s/Z)$ zdatności urządzenia dla wartości *s* parametru diagnostycznego,

– prawdopodobieństwa $P(s/N)$ niezdatności urządzenia dla tej samej wartości parametru diagnostycznego,

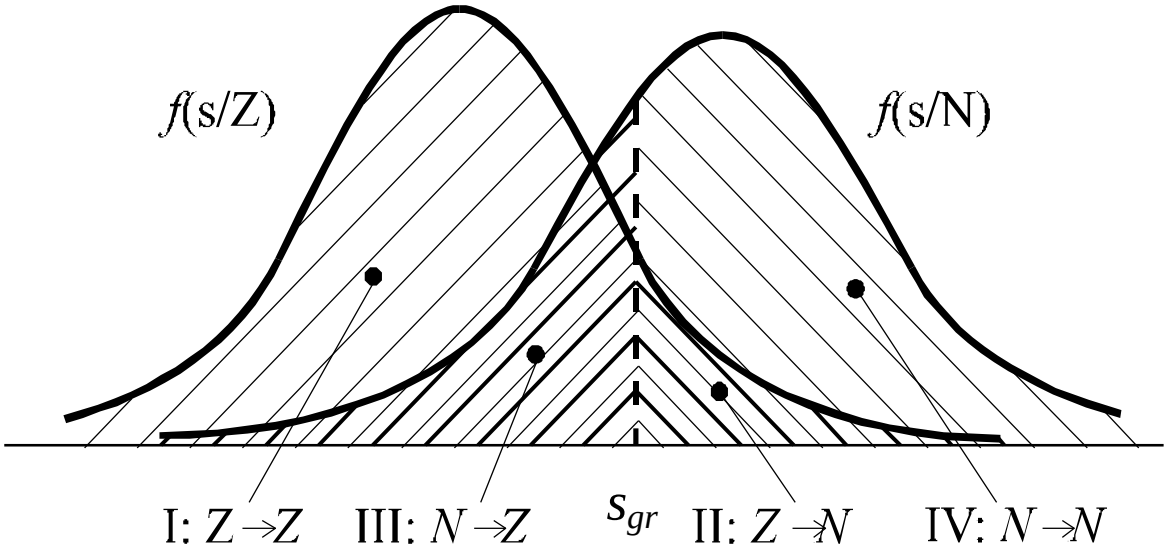
przy prawdopodobieństwie całkowitym wystąpienia wartości *s* parametru diagnostycznego:

$$P(s) = P(s/Z)P(Z) + P(s/N)P(N).$$

$P(Z)$ i $P(N)$ są wzajemnie uzupełniającymi się prawdopodobieństwami zdatności i niezdatności urządzenia.

Zgodnie z teorią błędów diagnozy Bayesa dla dwóch klas stanów technicznych oddzielonych wartością graniczną s_{gr} można wyróżnić:

- I: prawdopodobieństwo trafnej oceny stanu zdatności: $Z \rightarrow Z$: $P(Z)P(s \in \mathbf{S}_Z / Z)$,
- II: prawdopodobieństwo fałszywego alarmu i zbędnej naprawy, wynikającej z błędnej oceny stanu zdatności jako stanu niezdatności: $Z \rightarrow N$: $P(Z)P(s \in \mathbf{S}_N / Z)$,
- III: prawdopodobieństwo niewykrycia uszkodzenia – niewykrycia stanu niezdatności: $N \rightarrow Z$: $P(N)P(s \in \mathbf{S}_Z / N)$,
- IV: prawdopodobieństwo wykrycia uszkodzenia – trafnej oceny stanu niezdatności: $N \rightarrow N$: $P(N)P(s \in \mathbf{S}_N / N)$.



Na bazie teorii błędów diagnozy Bayesa można wyznaczać wartości graniczne.

Metoda minimalnego ryzyka decyzji diagnostycznej

Rozważa się następujące prawdopodobieństwa możliwych zdarzeń:

- prawdopodobieństwo trafnej oceny stanu zdatności: I: $Z \rightarrow Z$:

$$P(Z)P\left(s < s_{gr} / Z\right) = P(Z)\int_{-\infty}^{s_{gr}} f(s/Z)\mathrm{d}s,$$

- prawdopodobieństwo fałszywego alarmu i zbędnej naprawy: II: $Z \rightarrow N$:

$$P(Z)P\left(s \geq s_{gr} / Z\right) = P(Z)\int_{s_{gr}}^{\infty} f(s/Z)\mathrm{d}s,$$

- prawdopodobieństwo niewykrycia uszkodzenia – niewykrycia stanu niezdatności: III: $N \rightarrow Z$:

$$P(N)P\left(s < s_{gr} / N\right) = P(N)\int_{-\infty}^{s_{gr}} f(s/N)\mathrm{d}s,$$

- prawdopodobieństwo wykrycia uszkodzenia – trafnej oceny stanu niezdatności: IV: $N \rightarrow N$:

$$P(N)P\left(s \geq s_{gr} / N\right) = P(N)\int_{s_{gr}}^{\infty} f(s/N)\mathrm{d}s.$$

Przyporządkowując każdemu rodzajowi decyzji diagnostycznej odpowiednie uogólnione koszty $c_{w_m \to w_m}$ (np. kary za złą diagnozę i nagrody za dobrą diagnozę) ryzyko decyzji diagnostycznej przedstawia się w postaci

$$R = c_{Z\to Z} P(Z) \int_{-\infty}^{s_{gr}} f(s/Z)\mathrm{d}s + c_{Z\to N} P(Z) \int_{s_{gr}}^{\infty} f(s/Z)\mathrm{d}s +$$

$$+ c_{N\to Z} P(N) \int_{-\infty}^{s_{gr}} f(s/N)\mathrm{d}s + c_{N\to N} P(N) \int_{s_{gr}}^{\infty} f(s/N)\mathrm{d}s$$

Poszukiwana wartość graniczna minimalizuje to ryzyko ($\frac{\mathrm{d}R}{\mathrm{d}s} = 0$), czyli

$$c_{Z\to Z} P(Z) f(s_{gr}/Z) - c_{Z\to N} P(Z) f(s_{gr}/Z) + c_{N\to Z} P(N) f(s_{gr}/N) - c_{N\to N} P(N) f(s_{gr}/N) = 0$$

$$\frac{f(s_{gr}/Z)}{f(s_{gr}/N)} = \frac{(c_{N\to N} - c_{N\to Z})P(N)}{(c_{Z\to Z} - c_{Z\to N})P(Z)}$$

Znając postacie gęstości prawdopodobieństw dla stanu zdatności *Z*: $f(s/Z)$ i stanu niezdatności *N*: $f(s/N)$ oraz koszty decyzji diagnostycznych i prawdopodobieństwa wystąpienia stanów *Z* i *N* można z wyznaczyć wartości graniczne s_{gr} parametrów diagnostycznych.

Metoda minimalnej liczby błędnych decyzji diagnostycznych (minimalnego błędu diagnozy) Uproszczenie metody minimalnego ryzyka polegające na opuszczeniu skrajnych członów w prawej stronie równania ryzyka decyzji diagnostycznej odpowiadających poprawnym decyzjom diagnostycznym. Do rozwiązania otrzymuje się równanie

$$\frac{f\left(s_{gr}/Z\right)}{f\left(s_{gr}/N\right)} = \frac{c_{N\to Z}P(N)}{c_{Z\to N}P(Z)} = \frac{1}{\frac{c_{Z\to N}}{c_{N\to Z}}}\frac{P(N)}{P(Z)} = \frac{1}{\lambda}\frac{P(N)}{P(Z)},$$

gdzie: $c_{Z\to N}$ - koszt fałszywego alarmu i zbędnej naprawy,

$c_{N\to Z}$ - koszt niewykrycia uszkodzenia.

Zwykle wzajemny stosunek tych kosztów λ przyjmuje się na poziomie 1/3.

Znając postacie gęstości prawdopodobieństw dla stanu zdatności *Z*: $f\left(s/Z\right)$ i stanu niezdatności *N*: $f\left(s/N\right)$ oraz koszty błędnych decyzji diagnostycznych i prawdopodobieństwa wystąpienia stanów *Z* i *N* można z wyznaczyć wartości graniczne s_{gr} parametrów diagnostycznych.

Metoda Neymana-Pearsona

Minimalizowane jest prawdopodobieństwo niewykrycia uszkodzenia przy założonym poziomie zbędnych napraw, na jakie zezwala polityka remontowo-eksploatacyjna prowadzona w systemie eksploatacji. Błędna decyzja o zbędnej naprawie nastąpi gdy wartość parametru diagnostycznego urządzenia będącego w stanie zdatności Z przekroczy wartość graniczną s_{gr}. Prawdopodobieństwo tego zdarzenia dla przypadku $s_Z < s_N$ opisane jest relacją II: $Z \rightarrow N$:

$$P(Z)P\left(s \geq s_{gr}/Z\right) = P(Z)\int_{s_{gr}}^{\infty} f(s/Z)\mathrm{d}s.$$

Przyrównanie go do zadanego prawdopodobieństwa zbędnych napraw A daje relację Neymana-Pearsona: $P(Z)P\left(s \geq s_{gr}/Z\right) = P(Z)\int_{s_{gr}}^{\infty} f(s/Z)\mathrm{d}s = A.$

s_{gr} jest kwantylem rzędu $1-\frac{A}{P(Z)}$ (rzędu $\frac{A}{P(Z)}$ dla $s_Z > s_N$) rozkładu przyjętego do opisu prawdopodobieństwa występowania wartości parametru diagnostycznego dla zdatnego urządzenia.

Dla gęstości prawdopodobieństwa $f(s/Z)$ ciągłych, o rozkładach z jednym maksimum i symetrycznych (niekoniecznie gaussowskich), relację Neymana-Pearsona można zastąpić oszacowaniem wynikającym z nierówności Gaussa, co daje:

$$s_{gr} \le s_{ZN} \pm \frac{2}{3} os_{ZN} \sqrt{\frac{P(Z)}{A}}$$ („+” dla $s_Z < s_N$, „–” dla $s_Z > s_N$),

gdzie: s_{ZN} - wartość średnia wszystkich wartości parametru diagnostycznego,
os_{ZN} - odchylenie standardowe wszystkich wartości parametru diagnostycznego,
$P(Z)$ - prawdopodobieństwo zdatności urządzenia,
A - dopuszczalne w systemie eksploatacji prawdopodobieństwo zbędnych napraw (fałszywych alarmów); $A = 0{,}05$ oznacza, że 5 % wykonywanych napraw może być zbędne (wywołane fałszywymi alarmami).

W przypadku niemożliwości wyznaczenia gęstości prawdopodobieństwa $f(s/Z)$ – może to być dowolny rozkład, relację Neymana-Pearsona można zastąpić oszacowaniem wynikającym z nierówności Czebyszewa, co daje $s_{gr} \le s_{ZN} \pm os_{ZN} \sqrt{\frac{P(Z)}{A}}$ („+” dla $s_Z < s_N$, „–” dla $s_Z > s_N$).

Prawdopodobieństwo zdatności $P(Z)$ można zastąpić współczynnikiem gotowości eksploatacyjnej $k_g = \frac{t_u}{t_u + t_o}$

t_u - łączny czas użytkowania urządzenia w systemie eksploatacji,

t_o - łączny czas obsługiwania i napraw urządzenia w systemie eksploatacji.

Wtedy wartość graniczna parametru diagnostycznego jest uzależniona nie tylko od właściwości eksploatowanych urządzenia, ale także od parametrów charakteryzujących system ich eksploatacji.

Eksploatując nowe urządzenia można przyjąć dużą wartość współczynnika gotowości eksploatacyjnej i mały poziom zbędnych napraw, co da wartość graniczną parametru diagnostycznego bardziej różniącą się od wartości średniej.

Dla urządzeń mocno wyeksploatowanych można obniżyć wartość współczynnika gotowości eksploatacyjnej i jednocześnie podwyższyć poziom zbędnych napraw. Da to w efekcie wartość graniczną parametru bardziej zbliżoną do wartości średniej, czyli wcześniej będzie można kierować urządzenia do obsługi lub naprawy.

Przy diagnozowaniu urządzenia złożonego, gdy istnieje wiele (k) stanów technicznych w (zdatność/niezdatność poszczególnych elementów składowych urządzenia) i wiele (r) reakcji urządzenia y (w tym m parametrów wyjściowych s), do podjęcia decyzji wykorzystuje się tzw. **probabilistyczną macierz diagnostyczną**:

$$\mathbf{M}_p^d = \begin{array}{c} \\ w_1 \\ w_2 \\ \dots \\ w_i \\ \dots \\ w_k \end{array} \begin{array}{c} \begin{array}{ccc} y_1 & \dots & y_r \end{array} \\ \begin{bmatrix} P(y_1 / w_1) & \dots & P(y_r / w_1) \\ P(y_1 / w_2) & \dots & P(y_r / w_2) \\ \dots & \dots & \dots \\ P(y_1 / w_i) & \dots & P(y_r / w_i) \\ \dots & \dots & \dots \\ P(y_r / w_k) & \dots & P(y_r / w_k \end{bmatrix} \end{array}$$

gdzie $P(y_j / w_i)$ są prawdopodobieństwami warunkowymi zaobserwowania wartości y_j reakcji urządzenia znajdującego się w stanie technicznym w_i

oraz **probabilistyczną macierz diagnostyczną zdarzeń jednoczesnych** ($\mathbf{M}_{pr}^{d}$), której elementami są prawdopodobieństwa jednoczesnego zajścia zdarzenia pojawienia się wartości y_j reakcji urządzenia i zdarzenia znajdowania się tego urządzenia w stanie technicznym w_i, czyli prawdopodobieństwo $p(w_i, y_j)$:

$$\mathbf{M}_{pr}^{d} = \begin{array}{c} \\ w_1 \\ w_2 \\ \dots \\ w_i \\ \dots \\ w_k \end{array} \begin{array}{c} \begin{array}{cccc} y_1 & y_2 & \dots & y_r \end{array} \\ \begin{bmatrix} p(w_1, y_1) & p(w_1, y_2) & \dots & p(w_1, y_r) \\ p(w_2, y_1) & p(w_2, y_2) & \dots & p(w_2, y_r) \\ \dots & \dots & \dots & \dots \\ p(w_i, y_1) & p(w_i, y_2) & \dots & p(w_i, y_r) \\ \dots & \dots & \dots & \dots \\ p(w_k, y_1) & p(w_k, y_2) & \dots & p(w_k, y_r) \end{bmatrix} \end{array}$$

Elementy probabilistycznej macierzy diagnostycznej i probabilistycznej macierzy diagnostycznej zdarzeń jednoczesnych są ze sobą powiązane zależnościami:

$$\begin{cases} p(w_i, y_j) = p(w_i)\,p(y_j / w_i) \\ p(w_i, y_j) = p(y_j)\,p(w_i / y_j) \end{cases}$$

gdzie: $p(w_i)$ - prawdopodobieństwo wystąpienia stanu technicznego w_i urządzenia,

$p(y_j)$ - prawdopodobieństwo pojawienia się wartości y_j reakcji urządzenia,

$p(y_j / w_i)$ - prawdopodobieństwo warunkowe pojawienia się wartości y_j reakcji urządzenia znajdującego się stanie technicznym w_i,

$p(w_i, y_j)$ - prawdopodobieństwo zaistnienia stanu technicznego w_i urządzenia i zaobserwowania wartości y_j reakcji tego urządzenia.

Do oceny ogólności (reduntantności) parametrów diagnostycznych i podatności diagnostycznej urządzeń (ogólnie pojawienia się wartości y_j reakcji urządzenia w sytuacji wystąpienia stanu w_i) można wykorzystać **binarną macierz diagnostyczną**:

$$\mathbf{M}_b^d = \begin{array}{c} \\ w_1 \\ w_2 \\ \dots \\ w_i \\ \dots \\ w_k \end{array} \begin{array}{c} \begin{array}{ccc} y_1 & \dots & y_r \end{array} \\ \begin{bmatrix} 0 \vee 1 & \dots & 0 \vee 1 \\ 0 \vee 1 & \dots & 0 \vee 1 \\ \dots & \dots & \dots \\ 0 \vee 1 & \dots & 0 \vee 1 \\ \dots & \dots & \dots \\ 0 \vee 1 & \dots & 0 \vee 1 \end{bmatrix} \end{array}$$

1 dla przypadków, gdy wystąpienie stanu w_i urządzenia powoduje zmianę parametru diagnostycznego y_j (parametr y_j odzwierciedla wystąpienie stanu w_i),

0 dla przypadków, gdy wystąpienie stanu w_i urządzenia nie powoduje zmiany parametru diagnostycznego y_j (parametr y_j nie odzwierciedla wystąpienia stanu w_i).

Budowa testów diagnostycznych

Identyfikacja stanu technicznego urządzenia polega na jego kontroli i na lokalizacji uszkodzenia.

Pojedynczą operację wykonywaną w celu takiej identyfikacji nazywa się **sprawdzeniem**.

Sprawdzeniem jest wykonanie pomiaru wartości parametru wyjściowego (diagnostycznego) i przypisanie mu określonego stanu technicznego.

Wykonanie sprawdzenia dzieli zbiór elementów struktury urządzenia na trzy podzbiory:

- niezawierający elementów uszkodzonych,
- zawierający elementy uszkodzone,
- zawierający elementy w nieznanym stanie technicznym.

Rzeczywistą liczbę sprawdzeń możliwych do wykonania w urządzeniu można wyznaczyć analizując jego modele diagnostyczne.

Zbiór sprawdzeń, umożliwiający rozróżnianie wszystkich stanów technicznych urządzenia nazywa się **testem diagnostycznym**.

Dąży się do tego, aby testy diagnostyczne zawierały jak najmniej sprawdzeń (były jak najmniej pracochłonne).

Metoda „dziecięca” albo **kolejnego wyboru sprawdzeń** – stosowana gdy nie są znane wskaźniki niezawodnościowe urządzenia, czasy wykonywania sprawdzeń i ich koszty – sprawdzane są kolejno hipotezy o możliwych uszkodzeniach elementów urządzenia, przy czym sprawdzenia ostatniego elementu się nie wykonuje.

Metoda maksymalnej ilości informacji – stosowana gdy znana jest ilość informacji dostarczana przez różne sprawdzenia różnych cech urządzenia oraz znane są prawdopodobieństwa znalezienia się urządzenia w różnych stanach – sprawdzenia wykonuje się w kolejności malejącej warunkowej ilości informacji (tj. odniesionej do prawdopodobieństwa wystąpienia uszkodzeń).

Metoda spadku skuteczności informatycznej – stosowana gdy znane są koszty wykonania sprawdzeń i ilość uzyskanej w wyniku ich stosowania informacji – sprawdzenia wykonuje się w kolejności spadku wartości stosunku ilości informacji, której dostarczają, do nakładów jakie trzeba ponieść na ich zastosowanie.

Metoda różnych prawdopodobieństw (kontroli „słabych ogniw”) – stosowana gdy znane są wartości prawdopodobieństw wystąpienia niezdatności poszczególnych elementów urządzenia – najpierw sprawdza się te elementy, o których z dotychczasowej eksploatacji wiadomo, że uszkadzają się najczęściej, następnie bada się stan elementów urządzenia w kolejności malenia prawdopodobieństw uszkodzenia. Metoda ta daje tym lepsze efekty, im większa jest ilość elementów w urządzeniu i im bardziej różnią się od siebie wartościami prawdopodobieństw wystąpienia ich niezdatności.

Metoda spadku skuteczności probabilistycznej – stosowana gdy znane są wartości prawdopodobieństw wystąpienia niezdatności poszczególnych elementów urządzenia oraz wartości czasów wykonania sprawdzeń stanu tych elementów (ich pracochłonność) – w pierwszej kolejności wyznacza się dla wszystkich możliwych rodzajów sprawdzeń wszystkich elementów urządzenia stosunki prawdopodobieństw wystąpienia niezdatności elementów urządzenia do pracochłonności wykonania sprawdzeń, potrzebnych do ich wykrycia, następnie wykonuje się sprawdzenia w kolejności spadku wartości tak wyznaczonych stosunków.

Metoda najmniejszych kosztów diagnozowania – stosowana gdy znane są tylko wartości kosztów wykonania sprawdzeń stanu poszczególnych elementów urządzenia – sprawdzenia wykonuje się w kolejności od najtańszych do najdroższych.

Metoda najkrótszego czasu diagnozowania – stosowana gdy znane są tylko wartości czasów wykonania sprawdzeń stanu poszczególnych elementów urządzenia (ich pracochłonność) – sprawdzenia wykonuje się w kolejności od najkrócej do najdłużej trwających.

Metoda kontroli grupowej – za pomocą pierwszego sprawdzenia dokonuje się podziału elementów urządzenia na grupę elementów zdatnych i grupę elementów uszkodzonych. Następne sprawdzenia wykonuje się na grupie elementów uszkodzonych, co prowadzi do dalszego podziału na podgrupy aż do wykrycia uszkodzonych pojedynczych elementów. Zaletą kontroli grupowej jest jej zmienność, ponieważ wybór następnego sprawdzenia w trakcie badania zależy od wyniku sprawdzenia poprzedniego. Taki zmienny test dostarcza najwięcej informacji o urządzeniu przy stosunkowo najmniejszych nakładach z tym związanych.

Metoda podziału połówkowego – stosowana gdy nie są znane wartości prawdopodobieństw występowania niezdatności oraz gdy nie bierze się pod uwagę kosztów sprawdzeń (bo są nie znane albo takie same) – podczas pierwszego sprawdzenia bada się połowę elementów urządzenia, następnie bada się połowę elementów z tej części urządzenia, która okazała się niezdatna i tak dalej aż do wykrycia uszkodzonego elementu.

Metoda równych prawdopodobieństw – stosowana gdy znane są wartości prawdopodobieństw wystąpienia niezdatności elementów urządzenia – zbiór elementów urządzenia dzieli się tak aby sumy wartości tych prawdopodobieństw były równe lub prawie równe – podczas pierwszego sprawdzenia bada się pierwszą grupę elementów urządzenia, następnie bada się grupę elementów z tej części urządzenia, która okazała się niezdatna i tak dalej aż do wykrycia uszkodzonego elementu.

Metoda „drzewa defektów” – stosowana gdy znany jest topologiczny model-graf urządzenia, tzw. „drzewo defektów”. Na najwyższym poziomie tego drzewa umieszcza się uszkodzenia powodujące niezdatność całej urządzenia. Niżej grupowane są defekty związane z cechami coraz mniej wpływającymi na funkcjonowanie urządzenia, aż do cech w zasadzie nieistotnych dla funkcjonowania (ale też ważnych z innych powodów, np. estetycznych). Sprawdzenia stanu urządzenia wykonuje się kolejnymi poziomami opracowanego „drzewa defektów”. W przypadku stwierdzenia niezdatności na którymś z poziomów do lokalizacji uszkodzeń stosuje się odpowiednie inne metody diagnozowania.

Metody i techniki diagnozowania

- **bezprzyrządowe** (**organoleptyczne**) – diagnosta ocenia stan urządzenia posługując się tylko swoimi zmysłami, czyli wzrokiem, słuchem, dotykiem, węchem i smakiem,
- **przyrządowe** – diagnosta posługuje się przyrządami diagnostycznymi. Wyróżnia się metody:
  - **bezpośrednie** – bez demontowania elementów urządzenia bezpośrednio mierzy się wartości parametrów struktury, wykorzystując wzorce, szablony (np. szczelinomierze), przyrządy kreskowe (np. linijki, metrówki, suwmiarki, śruby mikrometryczne), przyrządy optyczne, indukcyjne, pojemnościowe, rezystancyjne (np. tensometry) itp.,
  - **pośrednie** – stan urządzenia określa się na podstawie pomiarów parametrów wyjściowych. Wyróżnia się metody diagnozowania oparte o wykorzystywanie:
    - **procesów roboczych** – obserwacje wzajemnego przetwarzania i przenoszenia energii w zależności od przeznaczenia i konstrukcji diagnozowanej urządzenia,
    - **procesów towarzyszących** – obserwacje procesów termicznych, wibracyjnych, akustycznych, elektromagnetycznych, tarciowych itd.,
    - **badań nieniszczących** – sonografia, rentgenografia, znakowanie substancjami promieniotwórczymi itp..

W zależności od sposobów i warunków realizacji wyróżnia się:

– **badania ruchowe** (dla pojazdów: drogowe) – stosunkowo mało dokładne i mało powtarzalne bo silnie uzależnienie od warunków pracy urządzenia i atmosferycznych; pracochłonne i niewygodne, często do ich przeprowadzenia wymagany jest specjalny pomiarowy odcinek drogowy, wykonywane są w rzeczywistych warunkach użytkowania urządzenia; badania:
  - **ciągłe** – wykonywane podczas bieżącego użytkowania urządzenia za pomocą wbudowanych urządzeń kontrolno-pomiarowych „organoleptycznie” przez człowieka obsługującego urządzenie lub automatycznie przez pokładowy system dozorujący (monitorujący),
  - **okresowe** – wykonywane okresowo za pomocą przenośnych przyrządów kontrolno-pomiarowych podczas prób drogowych,

– **badania stacjonarne** (**stanowiskowe**) – stosunkowo dokładne i powtarzalne; nie są uzależnione od warunków drogowych i atmosferycznych; umożliwiają wykorzystanie różnorodnej aparatury kontrolno-pomiarowej oraz wykonywanie czynności regulacyjnych podczas pomiarów; nie uwzględniają dla urządzeń mobilnych oporów hydro- i aerodynamicznych i oporów wleczonych, toczonych i ciągnionych elementów nienapędzanych; badania:

  • **ogólne** (**kompleksowe**) – wykonywane okresowo w cyklach określonych przez instrukcję eksploatacji urządzenia lub przepisy określające dopuszczenie urządzenia do eksploatacji na specjalnych stanowiskach diagnostycznych lub za pomocą przenośnej aparatury kontrolno-pomiarowej w celu stwierdzenia zdatności (przydatności) obiektu,

  • **szczegółowe** – wykonywane w razie potrzeby (np. po negatywnym wyniku badań ogólnych) na specjalizowanych stanowiskach diagnostycznych lub na stanowiskach obsługowych w celu lokalizacji uszkodzeń.

W zależności od posiadanego wyposażenia stosowane są często mieszane sposoby realizacji badań.

Ze względu na stopień automatyzacji diagnozowania wyróżnia się metody:

- **ręczne**, wykonywane przez człowieka – diagnostę, w których kolejność i jakość wykonywanych czynności diagnostycznych zależą od diagnosty,
- **półautomatyczne**, w których udział człowieka – diagnosty ogranicza się do sterowania diagnozowaniem i wnioskowaniem diagnostycznym,
- **automatyczne**, w których diagnozowanie odbywa się układzie decyzyjnym diagnoskopu, bez ingerencji człowieka; zwykle tego typu metody diagnozowania stosuje się w pokładowych systemach diagnostyczno-sterujących (mechatronicznych).

W zależności od szczegółowości oraz od celu i roli diagnozowania urządzenia w systemie jego eksploatacji wyróżnia się:

1. **Diagnozowanie ogólne** całego urządzenia sprowadzające się do stwierdzenia, czy urządzenie jest zdatne (przydatne) czy niezdatne (nieprzydatne).

   a) Pierwszy etap to **diagnozowanie ciągłe (lub eksploatacyjne lub dozorowanie lub monitorowanie**). Zbiór informacji jest ograniczony i różny dla różnych urządzeń oraz zależy od doświadczenia użytkownika. Jest podstawą do skierowania urządzenia do systemu obsługiwania w celu dokładniejszego sprawdzenia. Niektóre urządzenia posiadają wbudowane czujniki progowe współpracujące z sygnalizatorami alarmowymi, co ogranicza możliwość błędów.

   b) **Właściwe diagnozowanie ogólne** odbywa się po skierowaniu urządzenia do systemu obsługiwania poprzez wyznaczanie wskaźników ogólnotechnicznych (np. skład spalin i hałas pojazdów), efektywności i ekonomii pracy (np. moc i zużycie paliwa i oleju w silniku). W efekcie tych badań urządzenie jest kierowane albo do użytkowania albo do obsługiwania / naprawiania i diagnozowanie szczegółowe. Diagnozowanie to spełnia też funkcję kontroli jakości wykonywanych czynności obsługowo-naprawczych urządzenia.

2. **Diagnozowanie szczegółowe** przeprowadza się w celu identyfikacji stanu urządzenia prowadzącej do ewentualnej lokalizacji uszkodzonych elementów urządzenia. W wyniku tego diagnozowania ustala się potrzeby i zakres czynności obsługowo-naprawczych, które należy wykonać aby urządzenie mogło być skierowane do użytkowania.

Integralnym elementem procesu diagnozowania urządzeń jest genezowanie i prognozowanie.

Genezowanie – nie istnieją metody (poza metodami związanymi z systemami ekspertowymi) ustalania stanów urządzeń, które zaistniały w przeszłości.

Metody prognozowania stanu technicznego:

- **metody matematyczne** – wszelkie subiektywne przesłanki dotyczące badania zmian stanów urządzeń opierają się na modelowaniu matematycznym:
  - **klasyczna ekstrapolacja trendu** – dobór odpowiedniej funkcji aproksymującej historię obserwowanego zjawiska, którą następnie ekstrapoluje się na prognozowane okresy. Wybór konkretnej postaci funkcji aproksymującej wynika z oceny dopasowania uzyskiwanych według niej wyników do danych empirycznych. Zakłada się utrzymywanie przyjętej tendencji rozwojowej w prognozowanych okresach – tworzy się różne modyfikacje tworzenia modelu, np. ogranicza proces estymacji funkcji trendu do zawężonego zbioru danych (pomija się najstarsze obserwacje);
  - **adaptacyjny model trendu**, np. metoda wyrównywania wykładniczego, która omija etap wyznaczenia jawnej postaci trendu obserwowanego procesu, tworząc od razu oszacowania jego liczbowych wartości – procedury te w sposób iteracyjny przy każdej nowej obserwacji analizowanego procesu modyfikują przyjęty model trendu poprzez zmianę parametrów funkcji prognozującej (predykatora) lub jego postaci. Umożliwia to szybkie dopasowanie modelu do zmian badanego procesu;
- **metody intuicyjne,** obejmujące rozległy obszar rozciągający się od wypowiedzi poszczególnych fachowców do ekspertyz zbiorowych opracowanych metodą głosowania, metodą dyskusji panelowych lub metodą delficką. Zebrane w ten sposób opinie będą tak dobre, jak dobrymi specjalistami są wytypowani do współpracy eksperci. Termin następnego diagnozowania ustalają eksperci – diagności.

Podatność diagnostyczna – właściwość urządzenia charakteryzująca jego przystosowanie do realizacji jego diagnozowania.

Podatność diagnostyczna razem z podatnościami użytkową, obsługową / naprawczą i likwidacyjną stanowią podatność eksploatacyjną urządzenia. Wszystkie składowe podatności eksploatacyjnej urządzeń można kształtować zarówno na etapie konstruowania jak i samej eksploatacji.

Właściwie skonstruowane urządzenie powinno być przystosowane do szybkiej i pewnej oceny stanu technicznego. Wymaga to prowadzenia między innymi:

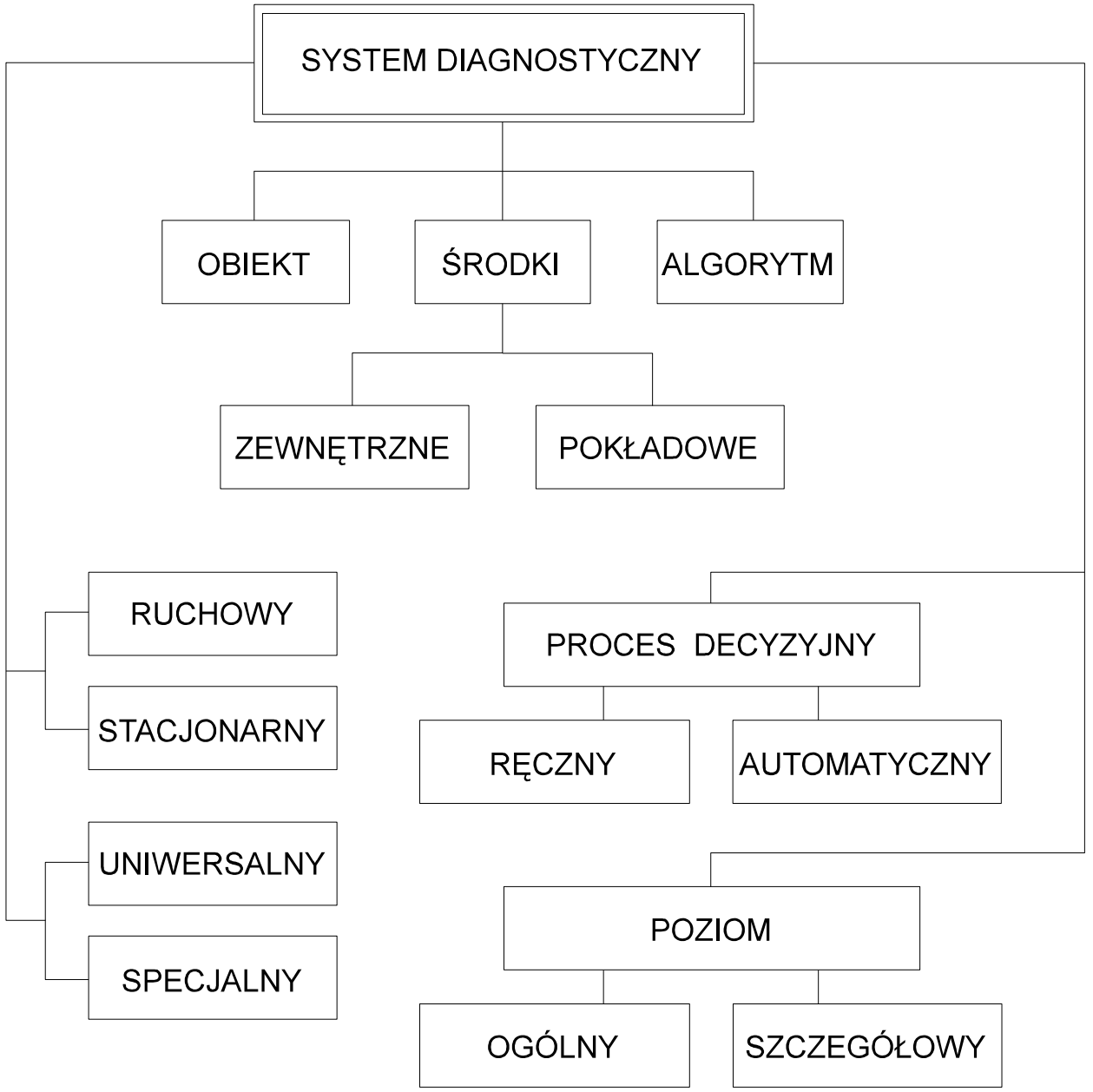
- analizy procesów starzenia i powstawania uszkodzeń konstruowanego urządzenia,
- analizy istniejących metod diagnozowania i urządzeń diagnostycznych oraz możliwości ich zastosowania do określania stanu technicznego konstruowanej urządzenia,
- określenia ekonomicznej i technicznej celowości diagnozowania poszczególnych elementów urządzenia.

Na podatność diagnostyczną urządzenia składają się następujące właściwości:

- dostępność miejsc diagnozowania (miejsc, do których mocuje się czujniki aparatury diagnostycznej),
- łatwość podłączenia aparatury diagnostycznej do urządzenia,
- możliwość diagnozowania za pomocą wielu przyrządów diagnostycznych,
- wyposażenie urządzenia w pokładowe systemy diagnostyczne (czujniki, okablowanie, procesory diagnostyczne, złącza diagnostyczne i układy wizualizacji stanu technicznego),
- jakość uzyskiwanych informacji diagnostycznych,
- pracochłonność i koszt diagnozowania,
- wygoda diagnozowania.

Zbiór składający się z diagnozowanego urządzenia, diagnostów oraz środków i algorytmów diagnozowania tworzy **system diagnostyczny**.

Środkami diagnostycznymi są odpowiednie przyrządy pomiarowe i stanowiska kontrolne (techniczne środki pomiarowe), a także subiektywne możliwości człowieka, organy jego zmysłów, doświadczenie, wiedza i wprawa.


SYSTEM DIAGNOSTYCZNY
OBIEKT
ŚRODKI
ALGORYTM
ZEWNĘTRZNE
POKŁADOWE
RUCHOWY
STACJONARNY
UNIWERSALNY
SPECJALNY
PROCES DECYZYJNY
RĘCZNY
AUTOMATYCZNY
POZIOM
OGÓLNY
SZCZEGÓŁOWY

Stwierdzono, że diagnozowanie należy traktować jako logiczny proces pozyskiwania i przetwarzania informacji przekazywanych są w postaci sygnałów diagnostycznych.

Informacja w języku potocznym ma sens wiadomości, zbioru faktów, idei, rzeczy realnie istniejących, pojęć i właściwości przedmiotów. W tym sensie „pełne" informacji są encyklopedie, słowniki, leksykony, podręczniki szkolne i akademickie, czasopisma, gazety itd. W tym przypadku istotna jest treść i wartość informacji – ***wymaga się aby informacje były prawdziwe i obiektywne***.

Informacja w teorii informacji, jako dziedzinie nauki związanej z telekomunikacją, jest związana z ilością wiadomości przenoszonych przez kanał telekomunikacyjny. W tym ***przypadku treść i wartość informacji nie jest istotna*** – wiadomość, nawet bezsensowna czy fałszywa może być z powodzeniem przesyłana.

Nośniki informacji – energia lub masa przepływające między danymi punktami czasoprzestrzeni i niosące wszelkie informacje w otaczającym nas świecie. Informacje są zawarte w zmianach stanu przepływających energii lub masy.

Modulacja – zmiana stanu nośnika informacji.

Sygnał – zmieniający się (modulowany) nośnik przekazujący informację.

S**ygnał** – wielkość fizyczna, której wartość lub czasowa zmienność przenosi informację.

Zakłócenie – sygnał zawierający informacje nieużyteczne lub niepożądane.
Szum – sygnał nie zawierający informacji użytecznych lub pożądanych.

Często szum i zakłócenie są rozumiane tożsamo, gdyż oba po nałożeniu się na sygnał zawierający informacje użyteczne i pożądane powodują zniekształcanie (fałszowanie) lub maskowanie tych użytecznych i pożądanych informacji.

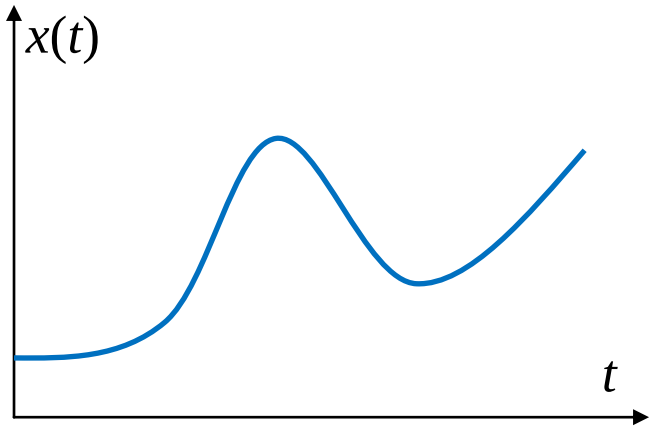
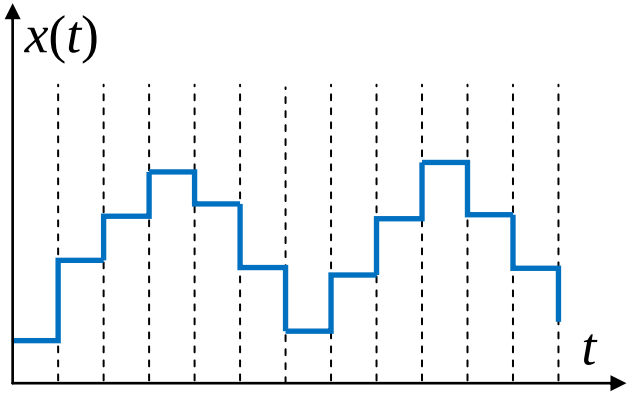
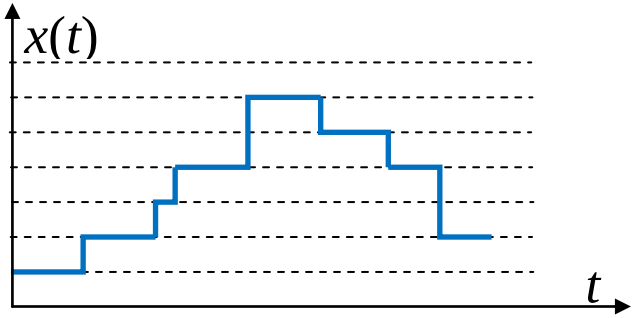
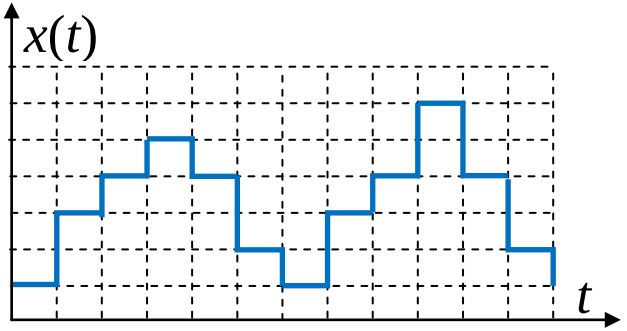
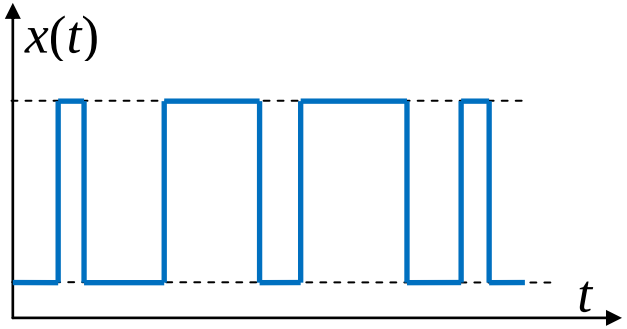
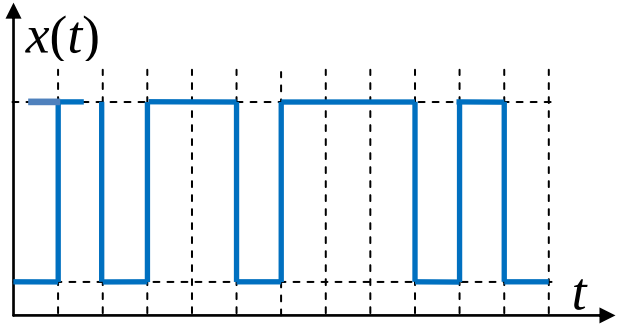
Sygnały ze względu na możliwe **wartości**:

- **analogowe**, bez kwantowania parametru informacyjnego:
  - parametr informacyjny może przyjąć ***dowolną*** wartość z przedziału swojej zmienności,
  - parametr informacyjny może przyjmować ***nieskończenie*** wiele wartości w ramach przedziału swojej zmienności różniących się od siebie pomijalnie mało,
- **dyskretne**, z kwantowaniem parametru informacyjnego:
  - parametr informacyjny może przyjąć ***określoną*** wartość z przedziału swojej zmienności,
  - parametr informacyjny może przyjmować ***skończenie*** wiele ustalonych wartości w ramach przedziału swojej zmienności różniących się od siebie o ustalone przedziały wartości (przedziały te zawierają wartości, których parametr informacyjny nie może przyjąć).

  Sygnały dyskretne klasyfikowane są ze względu na liczbę dopuszczalnych wartości. Szczególnie ważnym sygnałem dyskretnym jest sygnał **binarny**, w którym parametr informacyjny może przyjmować tylko dwie wartości, umownie przyjmowane za 0 (poziom niski) i 1 (poziom wysoki).

Sygnały ze względu na **czasową ciągłość** przepływu nośnika informacji:

- **ciągłe** – zmiana wartości parametru informacyjnego może nastąpić w ***dowolnej*** chwili czasu (czas nie jest kwantowany, jest ciągły),
- **nieciągłe** – zmiana wartości parametru informacyjnego może nastąpić w ***określonych*** chwilach czasu (czas jest kwantowany, nie jest ciągły).

sygnał	ciągły	nieciągły
analogowy		
dyskretny wielopoziomowy		
binarny		

Sygnały ze względu na **znajomość** całkowitej **postaci** sygnału:

– **zdeterminowane** – znana jest „całkowita” postać sygnału, włącznie z jego przeszłym i przyszłym przebiegiem, a informacja zawarta w tych sygnałach jest niezmienna i z góry znana; wyróżnia się sygnały:
  - **okresowe**, dla których $x(t) = x(t + T)$ – wartości parametru informacyjnego powtarzają się co stałą, niezerową i skończoną wartość czasu T, zwaną okresem sygnału; wyróżnia się sygnały:
    - **harmoniczne** (**sinusoidalne**) $x(t) = X\sin(\omega_x t + \varphi_x)$, gdzie:
      – X - amplituda sygnału w jego jednostkach,
      – $\omega_x = 2\pi f_x$ - częstość (pulsacja) sygnału w rad/s,
      – $f_x = \frac{1}{T}$ - częstotliwość sygnału w Hz,
      – φ_x - faza początkowa sygnału w rad, zwykle przyjmowana za 0 rad;

- **odkształcone (poliharmoniczne)** $x(t)=\sum_{k\in\mathbf{N}} X_k \sin\left(\omega_{x,k}t+\varphi_{x,k}\right)$, gdzie:
  - X_k - amplituda k-tej składowej harmonicznej sygnału w jego jednostkach,
  - $\omega_{x,k} = k\omega_x = 2\pi k f_x = 2\pi f_{x,k}$ - częstość (pulsacja) k-tej składowej harmonicznej sygnału w rad/s,
  - $f_{x,k} = kf_x = \frac{k}{T} = \frac{1}{T_k}$ - częstotliwość k-tej składowej harmonicznej sygnału w Hz,
  - $T_k = \frac{T}{k}$ - okres k-tej składowej harmonicznej sygnału w s,
  - $\varphi_{x,k}$ - faza początkowa k-tej składowej harmonicznej sygnału w rad,

- okresy składowych harmonicznych sygnału są całkowitymi podwielokrotnościami **niezerowego i skończonego okresu podstawowego**,
- składowe harmoniczne sygnału mają częstości (częstotliwości) równe tylko całkowitym wielokrotnościom **niezerowej i skończonej częstości** (**częstotliwości**) **podstawowej**,
- widmo częstotliwościowe sygnału jest dyskretne – niezerowe wartości amplitudy X_k składowych harmonicznych sygnału mogą występować tylko dla dyskretnych wartości częstotliwości, będących całkowitymi wielokrotnościami częstotliwości podstawowej (graficznie są to wykresy funkcji impulsowych o wysokościach X_k występujących dla częstości $\omega_{x,k} = k\omega_x$, a więc tak jakby **wyraźnie oddzielonych od siebie** tzw. ***prążków***),

- **odkształcone (poliharmoniczne)** $x(t) = \sum_{k \in \mathbf{N}} X_k \sin\left(\omega_{x,k} t + \varphi_{x,k}\right)$, gdzie:
  - X_k - amplituda k-tej składowej harmonicznej sygnału w jego jednostkach,
  - $\omega_{x,k} = k\omega_x = 2\pi k f_x = 2\pi f_{x,k}$ - częstość (pulsacja) k-tej składowej harmonicznej sygnału w rad/s,
  - $f_{x,k} = kf_x = \frac{k}{T} = \frac{1}{T_k}$ - częstotliwość k-tej składowej harmonicznej sygnału w Hz,
  - $T_k = \frac{T}{k}$ - okres k-tej składowej harmonicznej sygnału w s,
  - $\varphi_{x,k}$ - faza początkowa k-tej składowej harmonicznej sygnału w rad,

- wyróżnia się sygnały:
  - o skończonym widmie: $k = 1, 2, 3, ..., K$,
  - o nieskończonym widmie: $k = 1, 2, 3, ..., \infty$,

  im mniejsze jest K tym sygnał jest bardziej ***wąskopasmowy***, a im większe jest K tym sygnał jest bardziej ***szerokopasmowy***,

- czasem dodaje się jeszcze jedną składową dla $k = 0$, tzw. ***składową stałą*** (trudno ją nazwać składową harmoniczną (sinusoidalną), bo jej częstość (częstotliwość) wg definicji wynosi $0\omega_x = 0$ rad/s ($0f_x = 0$ Hz)); występowanie składowej stałej powoduje okresową zmianę wartości parametru informacyjnego nie wokół wartości 0, ale wokół wartości X_0;

– **zdeterminowane** – znana jest „całkowita” postać sygnału:
  - **nieokresowe**, wśród których wyróżnia się sygnały:
    - **rosnące** lub **zanikające**, np. opisywane funkcją skoku $x(t)=\begin{cases} X_1, & t \in (-\infty, t_0] \\ X_2, & t \in (t_0, \infty) \end{cases}$ lub

      innymi funkcjami monotonicznymi typu $x(t) \to \begin{cases} X_1, & t \to -\infty \\ (X_1 \mapsto X_2), & t \in (-\infty, \infty); \\ X_2, & t \to \infty \end{cases}$

    - **impulsowe** $x(t)=\begin{cases} X_1, & t \in (-\infty, t_1] \\ X_2, & t \in (t_1, t_2] \\ X_1, & t \in (t_2, \infty) \end{cases}$, $t_2 - t_1 \to 0$;

    - **trwałe** (**stałe**, **niezmienne**) $x(t) = \text{const}$;

- **prawie okresowe** w ujęciu pierwszym $x(t)=\sum_{k\in\mathbf{N}} X_k \sin\left(\omega_{p,k}t+\varphi_{x,k}\right)$, gdzie:
  - X_k - amplituda k-tej składowej harmonicznej sygnału w jego jednostkach,
  - $\omega_{p,k} = k\omega_p = 2\pi k f_p = 2\pi f_{p,k}$ - częstość (pulsacja) k-tej składowej harmonicznej sygnału w rad/s, przy czym pulsacja podstawowa $\omega_p \rightarrow 0$ rad/s,
  - $f_{p,k} = kf_p = \frac{k}{T_p} = \frac{1}{T_{p,k}}$ - częstotliwość k-tej składowej harmonicznej sygnału w Hz, przy czym częstotliwość podstawowa $f_p \rightarrow 0$ Hz,
  - $T_{p,k} = \frac{T_p}{k}$ - okres k-tej składowej harmonicznej sygnału w s, przy czym okres podstawowy $T_p \rightarrow \infty$ s,
  - $\varphi_{x,k}$ - faza początkowa k-tej składowej harmonicznej sygnału w rad,

- okresy składowych harmonicznych sygnału są całkowitymi podwielokrotnościami **nieskończenie długiego okresu podstawowego**,
- składowe harmoniczne sygnału mają częstości (częstotliwości) równe tylko całkowitym wielokrotnościom **nieskończenie małej częstości** (**częstotliwości**) **podstawowej**,
- widmo częstotliwościowe sygnału jest nieskończenie gęste – niezerowe wartości amplitudy X_k składowych harmonicznych sygnału mogą występować dla dowolnych wartości częstotliwości (graficznie ***prążki*** o wysokościach X_k występujące dla częstości $\omega_{p,k} = k\omega_p$ **mogą leżeć nieskończenie gęsto**), czasem także ciągłe,

- **prawie okresowe** w ujęciu pierwszym $x(t)=\sum_{k\in\mathbf{N}} X_k \sin\left(\omega_{p,k}t+\varphi_{x,k}\right)$, gdzie:
  - X_k - amplituda k-tej składowej harmonicznej sygnału w jego jednostkach,
  - $\omega_{p,k} = k\omega_p = 2\pi k f_p = 2\pi f_{p,k}$ - częstość (pulsacja) k-tej składowej harmonicznej sygnału w rad/s, przy czym pulsacja podstawowa $\omega_p \rightarrow 0$ rad/s,
  - $f_{p,k} = kf_p = \frac{k}{T_p} = \frac{1}{T_{p,k}}$ - częstotliwość k-tej składowej harmonicznej sygnału w Hz, przy czym częstotliwość podstawowa $f_p \rightarrow 0$ Hz,
  - $T_{p,k} = \frac{T_p}{k}$ - okres k-tej składowej harmonicznej sygnału w s, przy czym okres podstawowy $T_p \rightarrow \infty$ s,
  - $\varphi_{x,k}$ - faza początkowa k-tej składowej harmonicznej sygnału w rad,

- wyróżnia się sygnały:
  - o skończonym widmie: $k = 1, 2, 3, ..., K$,
  - o nieskończonym widmie: $k = 1, 2, 3, ..., \infty$,

  im mniejsze jest K tym sygnał jest bardziej ***wąskopasmowy***, a im większe jest K tym sygnał jest bardziej ***szerokopasmowy***,
- czasem dodaje się jeszcze jedną składową dla $k = 0$, tzw. ***składową stałą*** (tak samo jak dla sygnałów poliharmonicznych); występowanie składowej stałej powoduje okresową zmianę wartości parametru informacyjnego nie wokół wartości 0, ale wokół wartości X_0;

- **prawie okresowe** w ujęciu drugim $x(t) = \sum_{i \in \mathbf{R}^+} X_i \sin(\omega_i t + \varphi_i)$, gdzie:
  - X_i - amplituda *i*-tej składowej harmonicznej sygnału w jego jednostkach,
  - $\omega_i = 2\pi f_i$ - częstość (pulsacja) *i*-tej składowej harmonicznej sygnału w rad/s,
  - $f_i = \frac{1}{T_i}$ - częstotliwość *i*-tej składowej harmonicznej sygnału w Hz,
  - T_i - okres *i*-tej składowej harmonicznej sygnału w s,
  - φ_i - faza początkowa *i*-tej składowej harmonicznej sygnału w rad,

- składowe harmoniczne sygnału mają **dowolne okresy i częstości** (**częstotliwości**),
- widmo częstotliwościowe sygnału jest nieskończenie gęste – niezerowe wartości amplitudy X_i składowych harmonicznych sygnału mogą występować dla dowolnych wartości częstotliwości (graficznie ***prążki*** o wysokościach X_i występujące dla częstości ω_i **mogą leżeć nieskończenie gęsto**), czasem także ciągłe,
- wyróżnia się sygnały:
  - o skończonym widmie: $i \in (0, I]$,
  - o nieskończonym widmie: $i \in (0, \infty)$,

  im mniejsze jest I tym sygnał jest bardziej ***wąskopasmowy***, a im większe jest I tym sygnał jest bardziej ***szerokopasmowy***,

- **prawie okresowe** w ujęciu drugim $x(t) = \sum_{i \in \mathbf{R}^+} X_i \sin(\omega_i t + \varphi_i)$, gdzie:
  - X_i - amplituda *i*-tej składowej harmonicznej sygnału w jego jednostkach,
  - $\omega_i = 2\pi f_i$ - częstość (pulsacja) *i*-tej składowej harmonicznej sygnału w rad/s,
  - $f_i = \frac{1}{T_i}$ - częstotliwość *i*-tej składowej harmonicznej sygnału w Hz,
  - T_i - okres *i*-tej składowej harmonicznej sygnału w s,
  - φ_i - faza początkowa *i*-tej składowej harmonicznej sygnału w rad,

- czasem dodaje się jeszcze jedną składową dla $i = 0$, tzw. ***składową stałą*** (tak samo jak dla sygnałów poliharmonicznych); występowanie składowej stałej powoduje okresową zmianę wartości parametru informacyjnego nie wokół wartości 0, ale wokół wartości X_0;

- **niezdeterminowane** – znana jest tylko postać zarejestrowanej (obserwowanej) części sygnału, a informacja zawarta w tych sygnałach może być zmienna i z góry nie jest znana; widma tych sygnałów mogą być nieskończenie gęste, także ciągłe; wyróżnia się sygnały:
  - **stochastyczne (losowe)** – wartości parametrów opisujących sygnał wyznaczone na podstawie ich dotychczasowej obserwacji można opisać rozkładami prawdopodobieństwa (np. Gaussa, Poissona itd.) i na tej podstawie, z określonym prawdopodobieństwem powodzenia, przewidywać ich przyszłe wartości; wyróżnia się sygnały:
    - **stacjonarne** – wartości parametrów opisujących sygnał nie zależą od położenia przedziału czasu obserwacji sygnału na osi czasu;
    - **ergodyczne** – wszystkie wystarczająco długie obserwowane fragmenty sygnału są „typowe” w takim sensie, że znajomość pojedynczego wystarczająco długiego fragmentu sygnału z jego nieskończenie długiego (lub w praktyce dostatecznie długiego) przebiegu pozwala na wyznaczenie wartości parametrów opisujących sygnał prawdziwych również dla innych, hipotetycznie identycznych (obserwowanych w tych samych warunkach) fragmentów sygnału;
  - **całkowicie przypadkowe**.

Sygnały w dziedzinie czasu i częstości (częstotliwości)

Dzięki całkowym przekształceniom (prostej i odwrotnej transformacji) Fouriera przyjmuje się wzajemną jednoznaczność zapisu sygnału w dziedzinie czasu t i w dziedzinie częstości $\omega = 2\pi f$ (częstotliwości f): **przebieg czasowy (waveform) ↔ widmo zespolone (complex spectrum).**
Posiadając pełny sygnał jako rzeczywistą funkcję czasu $x(t)$ (nieskończenie długi lub przynajmniej wystarczająco długi fragment sygnału) dzięki przekształceniu prostemu $X(\omega = 2\pi f) = \int_{-\infty}^{\infty} x(t) e^{-j\omega t}\,\mathrm{d}t = \int_{-\infty}^{\infty} x(t) e^{-j2\pi f t}\,\mathrm{d}t$ otrzymuje się zespoloną funkcję przedstawiającą zespolone widmo sygnału.
Dysponując z kolei zespolonym widmem sygnału dzięki przekształceniu odwrotnemu $x(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} X(\omega) e^{j\omega t}\,\mathrm{d}\omega = \int_{-\infty}^{\infty} X(f) e^{j2\pi f t}\,\mathrm{d}f$ otrzymuje się sygnał jako funkcję rzeczywistą.

Każda z zespolonych wartości $X(\omega)$ jest ***zespolonym „prążkiem" wartości sygnału*** w jego ***zespolonym widmie*** i reprezentuje „ilość sygnału" przypadającą na daną częstość ω (lub nieskończenie mały przedział częstości $\mathrm{d}\omega$) i stąd bywa nazywana ***gęstością widmową wartości (amplitudy) sygnału***.

Ma ona wymiar ***ilorazu*** **jednostki opisującej wartość sygnału** (m dla długości, V dla potencjału elektrycznego, A dla natężenia prądu itd.) i **rad/s**, gdy w widmie używa się skali częstości ω, lub **Hz**, gdy w widmie używa się skali częstotliwości *f*.

Gęstość widmowa wartości sygnału $X(\omega)$ zawiera

pełną informację o sygnale

i na jej podstawie można wiernie odtworzyć

rzeczywisty przebieg czasowy sygnału $x(t)$.

Zespolony „prążek” wartości sygnału można zapisać w postaci

$$X(\omega) = \mathrm{Re}\, X(\omega) + j\, \mathrm{Im}\, X(\omega) = |X(\omega)| e^{j\psi(\omega)},$$ gdzie:

- moduł $|X(\omega)| = \sqrt{(\mathrm{Re}\, X(\omega))^2 + (\mathrm{Im}\, X(\omega))^2}$ - rzeczywista amplituda składowej sygnału o częstości ω (wszystkie te wartości (graficznie prążki) tworzą rzeczywiste ***widmo amplitudowe***),
- argument $\psi(\omega) = \arg X(\omega) = \mathrm{arctg} \dfrac{\mathrm{Im}\, X(\omega)}{\mathrm{Re}\, X(\omega)}$ - rzeczywista faza składowej sygnału o częstości ω (wszystkie te wartości (graficznie prążki) tworzą rzeczywiste ***widmo fazowe***).

Widma amplitudowe $|X(\omega)|$ i fazowe $\psi(\omega)$
zawierają częściowe informacje o sygnale
i nie można na ich podstawie, rozpatrując je oddzielnie, pojedynczo,
odtworzyć rzeczywistego przebiegu czasowego sygnału $x(t)$.

Odtworzenie to jest możliwe tylko poprzez łączne rozpatrywane obu widm odpowiednio powiązanych funkcją eksponens.

Interpretacja fizyczna tożsamości Parsevala dla transformacji Fouriera

$$\int_{-\infty}^{\infty} x^2(t)\,\mathrm{d}t = \frac{1}{2\pi}\int_{-\infty}^{\infty} |X(\omega)|^2\,\mathrm{d}\omega$$

– sumując kwadraty wartości wszystkich składowych (graficznie prążków) widma amplitudowego otrzymuje się wartość energii sygnału wg definicji energii (pracy) procesu dynamicznie zmiennego opartej na wartości skutecznej tego procesu: $W = \int_{-\infty}^{\infty} x^2(t)\,\mathrm{d}t$.

W zasadzie wyznaczenie gęstości widmowej wartości sygnałów nieokresowych nie jest możliwe (ze względu na ich teoretycznie nieskończenie długie okresy i w związku z tym nieskończenie długie czasy rejestracji potrzebne do uzyskania całego sygnału do całkowania od $-\infty$ do $+\infty$ w prostej transformacie Fouriera).

Dzięki możliwościom pomiarowym, np. prostej filtracji wąskopasmowej i rejestracji wartości skutecznej dowolnego sygnału przez odpowiednio długi ale skończony czas i jej uśrednianiu po tym skończonym czasie pomiaru, można otrzymać wielkość nazywaną ***gęstością widmową mocy sygnału*** oznaczaną symbolem $G_{xx}(\omega)$, przy czym moc sygnału jest definiowana w oparciu o wartość skuteczną sygnału jako $P = \overline{x^2(t)} = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} x^2(t)\,\mathrm{d}t$.

Wymóg nieskończenie długiego czasu rejestracji sygnału dotyczy w zasadzie tylko sygnałów nieokresowych. Dla sygnałów okresowych czas rejestracji można ograniczyć do jednego okresu sygnału, a dla sygnałów ergodycznych do odpowiednio długiego ale skończonego czasu zawierającego „typowy” fragment sygnału itd.. Na mocy tożsamości Parsevala otrzymuje się

$$P = \int_{-\infty}^{\infty} \lim_{T\to\infty} \frac{1}{2T} \frac{|X(\omega)|^2}{2\pi}\,\mathrm{d}\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} \lim_{T\to\infty} \frac{|X(\omega)|^2}{2T}\,\mathrm{d}\omega.$$

Gęstość widmowa mocy sygnału przedstawia ilość mocy tego sygnału przypadającą na daną częstość ω (lub nieskończenie mały przedział częstości dω): $G_{xx}(\omega) = \frac{\mathrm{d}P}{\mathrm{d}\omega}$ (definicja).

Odwracając tę zależność dostaje się $P = \int G_{xx}(\omega)\mathrm{d}\omega$, a z porównania obu podejść do mocy sygnału (porównania funkcji podcałkowych) można zapisać, że $G_{xx}(\omega) = \frac{1}{2\pi}\lim_{T\to\infty}\frac{|X(\omega)|^2}{2T}$ – dana wartość (graficznie prążek dla częstości ω) gęstości widmowej mocy sygnału jest uśrednionym po czasie pomiaru kwadratem odpowiedniej wartości (graficznie prążka) widma amplitudowego tego sygnału.

Ponieważ gęstość widmowa mocy jest związana tylko z widmem amplitudowym, to na jej podstawie nie można odtworzyć rzeczywistego przebiegu czasowego sygnału *x*(*t*).

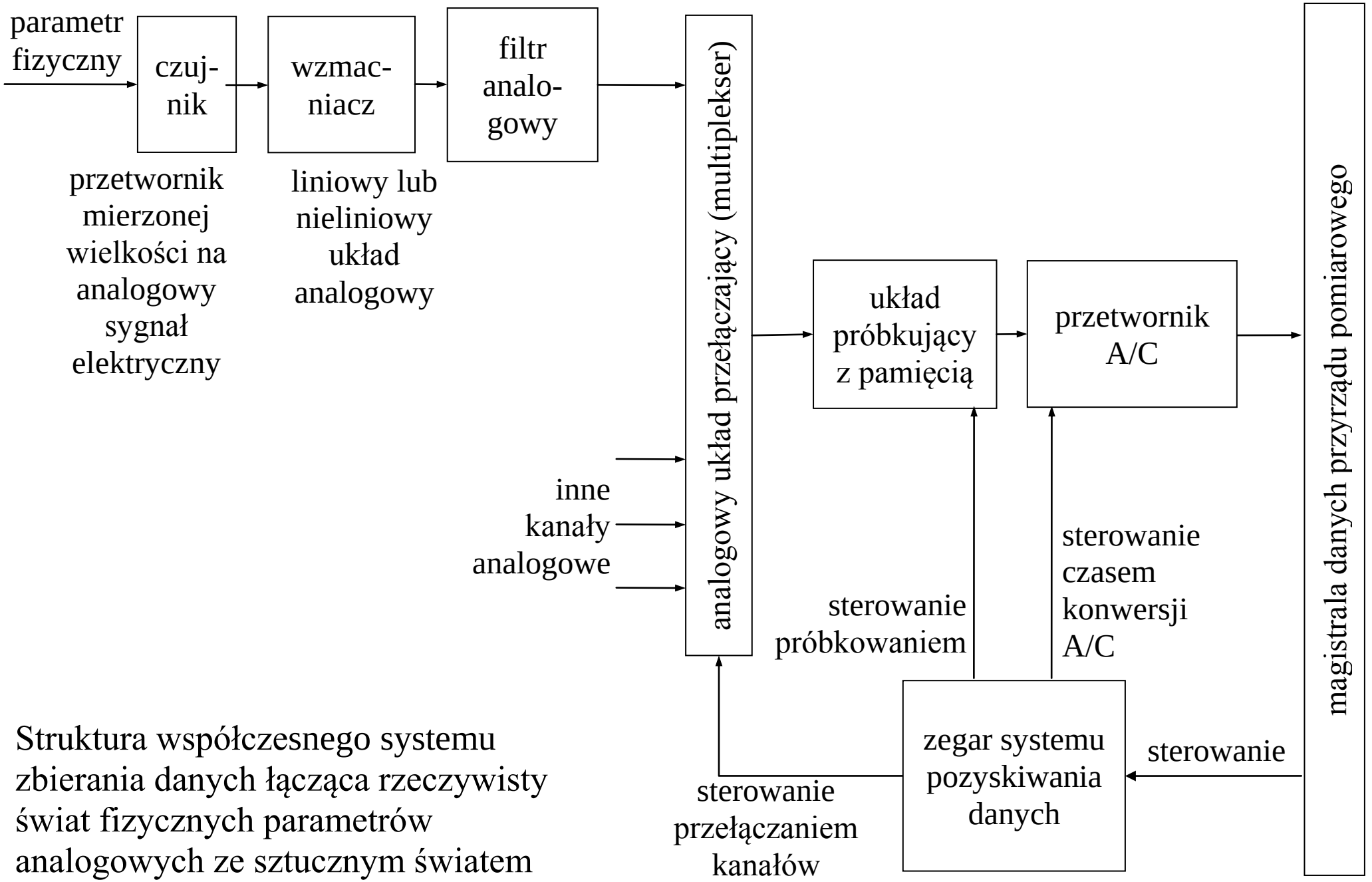
Dokonując odwrotnej transformaty Fouriera gęstości widmowej mocy sygnału uzyskuje się $R_{xx}(\tau)=\frac{1}{2\pi}\int_{-\infty}^{\infty}G_{xx}(\omega)e^{j\omega\tau}\mathrm{d}\omega$ (jedna zależności Wienera-Chinczyna).

Funkcja $R_{xx}(\tau)$ nazywa się ***funkcją autokorelacji sygnału*** i w dziedzinie czasu jest definiowana w postaci $R_{xx}(\tau)=\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}x(t)x(t+\tau)\mathrm{d}t=\overline{x(t)x(t+\tau)}$ – jest uogólnioną wartością średniokwadratową sygnału (uśrednioną po czasie pomiaru sumą iloczynów wartości fragmentów tego sygnału przesuniętych o pewien odcinek czasu) i podaje współzależność między dwoma fragmentami tego sygnału opóźnionymi o odcinek czasu τ.

Druga zależność Wienera-Chinczyna – prosta transformata Fouriera funkcji autokorelacji sygnału: $G_{xx}(\omega)=\int_{-\infty}^{\infty}R_{xx}(\tau)e^{-j\omega\tau}\mathrm{d}\tau$.

sygnał w dziedzinie czasu $x(t)$ $x(t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}X(\omega)e^{j\omega t}\,\mathrm{d}\,\omega=\int_{-\infty}^{\infty}X(f)e^{j2\pi ft}\,\mathrm{d}\,f$ funkcja rzeczywista przebieg czasowy sygnału	sygnał w dziedzinie częstości (częstotliwości) $X(\omega=2\pi f)$ $X(\omega=2\pi f)=\int_{-\infty}^{\infty}x(t)e^{-j\omega t}\,\mathrm{d}\,t=\int_{-\infty}^{\infty}x(t)e^{-j2\pi ft}\,\mathrm{d}\,t$ funkcja zespolona gęstość widmowa wartości sygnału $X(\omega)=\mathrm{Re}\,X(\omega)+j\,\mathrm{Im}\,X(\omega)=\lvert X(\omega)\rvert e^{j\psi(\omega)}$ $\lvert X(\omega)\rvert=\sqrt{(\mathrm{Re}\,X(\omega))^2+(\mathrm{Im}\,X(\omega))^2}$ funkcja rzeczywista – widmo amplitudowe $\psi(\omega)=\mathrm{arctg}\,\frac{\mathrm{Im}\,X(\omega)}{\mathrm{Re}\,X(\omega)}$ funkcja rzeczywista – widmo fazowe
operacja uśredniania sygnału po czasie jego pomiaru $R_{xx}(\tau)=\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}x(t)x(t+\tau)\mathrm{d}t=\overline{x(t)x(t+\tau)}$ funkcja rzeczywista – autokorelacja sygnału $R_{xx}(\tau)=\frac{1}{2\pi}\int_{-\infty}^{\infty}G_{xx}(\omega)e^{j\omega\tau}\mathrm{d}\omega=\int_{-\infty}^{\infty}G_{xx}(f)e^{j2\pi f\tau}\mathrm{d}f$	operacja uśredniania widma amplitudowego sygnału po czasie jego pomiaru T $G_{xx}(\omega)=\frac{1}{2\pi}\lim_{T\to\infty}\frac{\lvert X(\omega)\rvert^2}{2T}$ funkcja rzeczywista – gęstość widmowa mocy sygnału $G_{xx}(\omega)=\int_{-\infty}^{\infty}R_{xx}(\tau)e^{-j\omega\tau}\mathrm{d}\tau$

Większość mierzonych wielkości ma postać analogową (z zastrzeżeniem kwantowego charakteru zjawisk w skali atomowej). Procesory sygnałów mogą działać zarówno w sposób analogowy jak i cyfrowy. Ze względu na szerokie rozpowszechnienie elektronicznych układów binarnych, przetwarzanie sygnałów odbywa się obecnie raczej przy użyciu ich postaci cyfrowej. Sprzęganie stopni analogowych (czujniki, wyświetlacze) z cyfrowymi (cyfrowe procesory sygnałów) stanowi istotny problem układów pomiarowych (diagnostycznych), a układami umożliwiającymi odpowiednie ich dopasowanie są przetworniki analogowo-cyfrowe (A/C) oraz cyfrowo-analogowe (C/A). Przetworniki te znajdują się w zasadzie we wszystkich produkowanych współcześnie elektronicznych przyrządach pomiarowych (diagnostycznych).

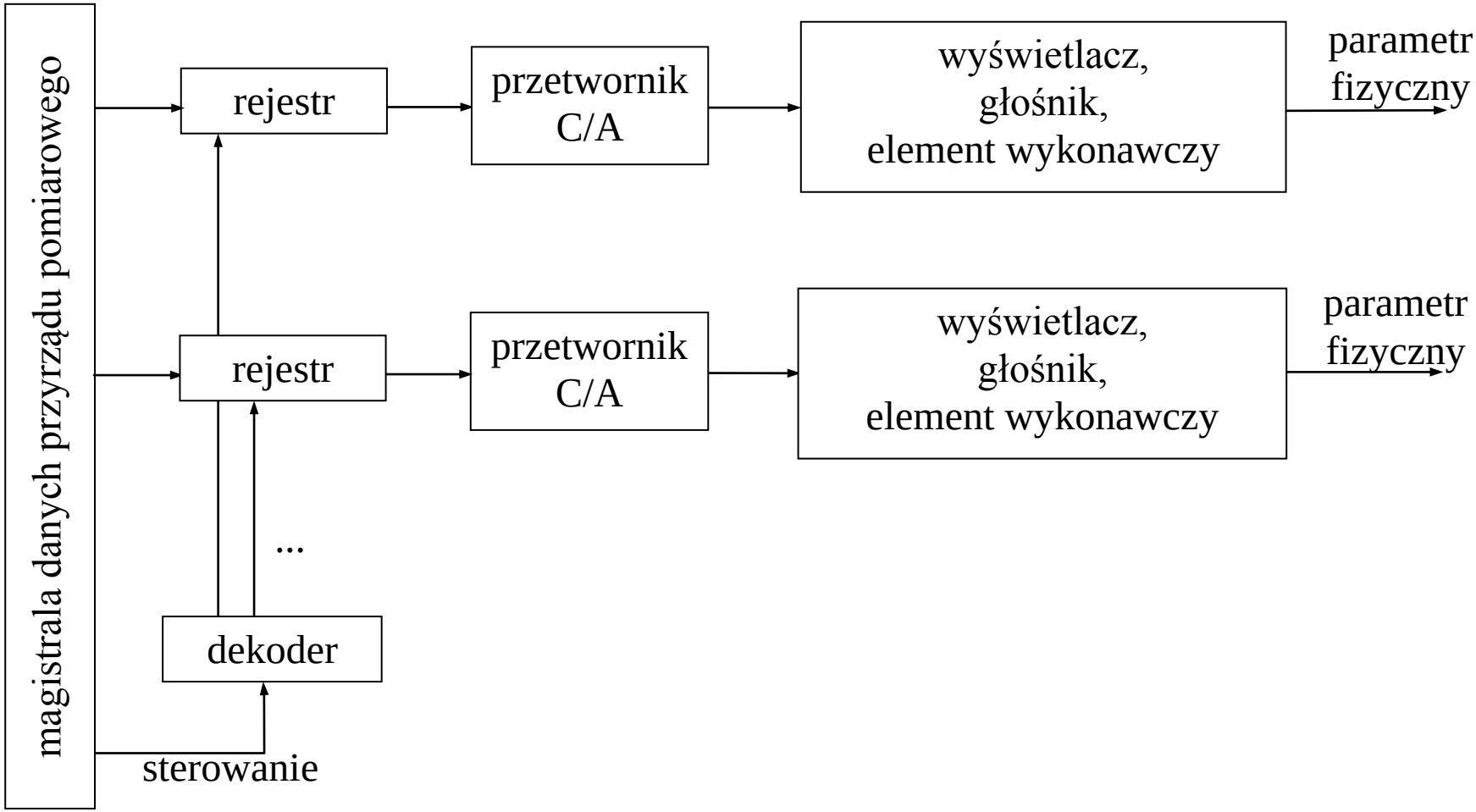


Struktura współczesnego systemu zbierania danych łącząca rzeczywisty świat fizycznych parametrów analogowych ze sztucznym światem cyfrowego przetwarzania danych

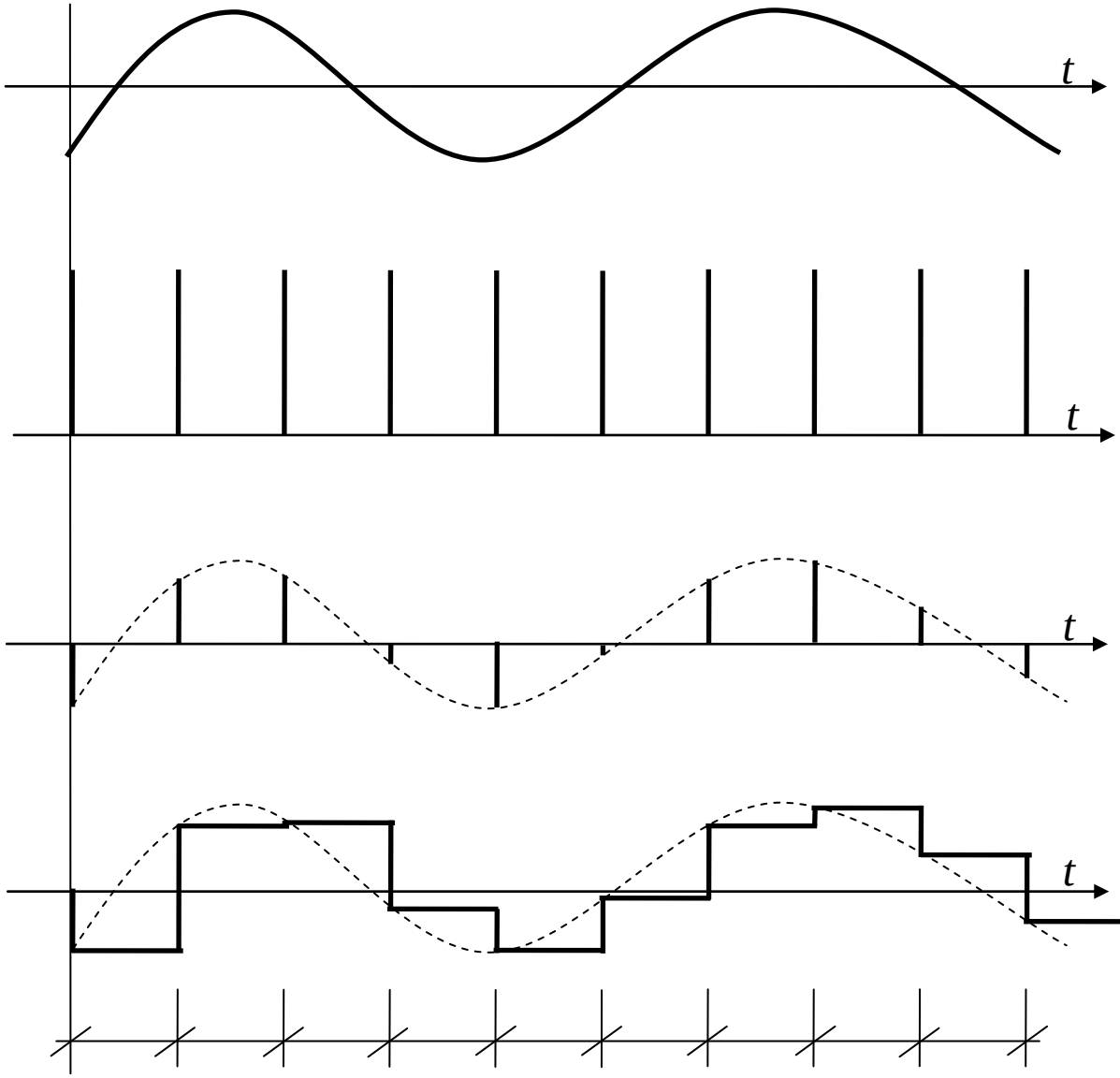
czujnik czujnik ... czujnik	→ → →	multiplekser	→	wzmacniacz i filtr	→	próbkowanie i A/C	→	magistrala danych		każdy kanał daje informację raz na cykl przełączania multipleksera, tym rzadziej, im więcej jest kanałów; wspólny wzmacniacz i filtr muszą być przestrajane odpowiednio do sygnału z podłączonego czujnika
czujnik czujnik ... czujnik	→ → →	wzmacniacz i filtr wzmacniacz i filtr wzmacniacz i filtr	→ → →	multiplekser	→	próbkowanie i A/C	→	magistrala danych		każdy kanał daje informację raz na cykl przełączania multipleksera, tym rzadziej, im więcej jest kanałów; czujnik, wzmacniacz i filtr mogą stanowić jedną całość
czujnik czujnik ... czujnik	→ → →	wzmacniacz i filtr wzmacniacz i filtr wzmacniacz i filtr	→ → →	próbkowanie i A/C próbkowanie i A/C ... próbkowanie i A/C	→ → →	magistrala danych				każdy kanał daje informację co każde próbkowanie

Sygnały cyfrowe przesyłane są do **elementów obliczeniowych** (**procesorów**), **pamiętających** (pamięci stałych albo ulotnych) i innych, określonych strukturą przyrządu pomiarowego (a umieszczone na magistrali danych przyrządu pomiarowego są pobierane przez te elementy zgodnie z programem (algorytmem działania) **systemu operacyjnego** i innych programów sterujących pracą przyrządu pomiarowego).

Po odpowiednim (wynikającym z działania tych elementów) przetworzeniu, przetworzone sygnały pomiarowe (cyfrowe wyniki pomiarów) są przekazywane kolejno dalej, odpowiednio do struktury przyrządu pomiarowego (lub znowu umieszczane na magistrali danych). Sygnały te mają postać cyfrową, najczęściej binarną, i aby były czytelne dla człowieka muszą zostać przekonwertowane na postać dla niego czytelną. Konwersja ta odbywa się w **systemie dystrybucji danych**.



Próbkowanie



wejściowy zmienny sygnał analogowy poddawany próbkowaniu

ciąg impulsów próbkujących o stałej (jednostkowej) amplitudzie powtarzających się co stały czas T_s

spróbkowany sygnał analogowy – zmodyfikowany ciąg impulsów próbkujących o wartościach równych wartości próbkowanego sygnału analogowego w chwilach występowania impulsów próbkujących

sygnał analogowy na wyjściu układu próbkująco-pamiętającego

czasy pomiędzy kolejnymi pobraniami próbek wartości mierzonego sygnału – równe sobie czasy, w trakcie których stała, zapamiętana, analogowa wartość sygnału mierzonego jest kwantowana i kodowana

Na pytanie o gęstość próbkowania w czasie (częstotliwość próbkowania) odpowiada **twierdzenie o próbkowaniu Shannona-Kotielnikowa**:

Jeżeli sygnał ciągły ma ograniczone widmo (ograniczone pasmo częstotliwości), a więc nie zawiera składowych sinusoidalnych o częstotliwościach większych od częstotliwości f_c, to sygnał ten może być odtworzony bez zniekształceń (jest jednoznacznie określony) tylko z próbek pobieranych z częstotliwością f_s, nie mniejszą niż dwukrotność częstotliwości f_c, czyli pobieranych w równo oddalonych chwilach czasu następujących po sobie nie rzadziej niż co $T_s = \frac{1}{f_s} = \frac{1}{2 f_c}$ ($f_s \geq 2f_c$; $f_N = 2f_c$ nazywa się częstotliwością Nyquista).

Twierdzenie to łatwo zilustrować korzystając z transformaty Fouriera iloczynu $x(t)\cdot\delta_{T_s}(t)=x(t)\cdot\sum_{n=0}^{\infty}\delta(t-nT_s)$ modelującego próbkowanie przebiegu czasowego sygnału $x(t)$ o widmie ograniczonym do częstotliwości f_c ciągiem funkcji impulsowych $\delta_{T_s}(t)=\sum_{n=0}^{\infty}\delta(t-nT_s)$ powtarzanych co stały czas T_s, będący odwrotnością częstotliwości próbkowania $f_s\left(T_s=\frac{1}{f_s}\right)$.

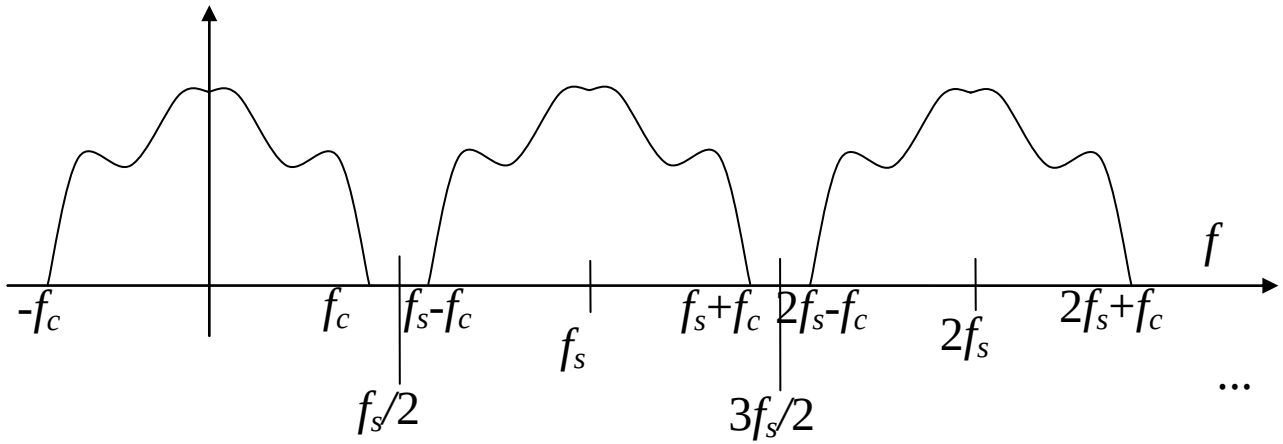
Transformatą tą jest splot $X(f)*f_s\delta_{f_s}(f)=f_sX(f)*\sum_{n=-\infty}^{\infty}\delta(f-nf_s)$, gdzie:

$X(f)$ - transformata Fouriera przebiegu czasowego analizowanego sygnału (widmo sygnału – zespolona funkcja częstotliwości o niezerowych wartościach tylko do częstotliwości f_c),

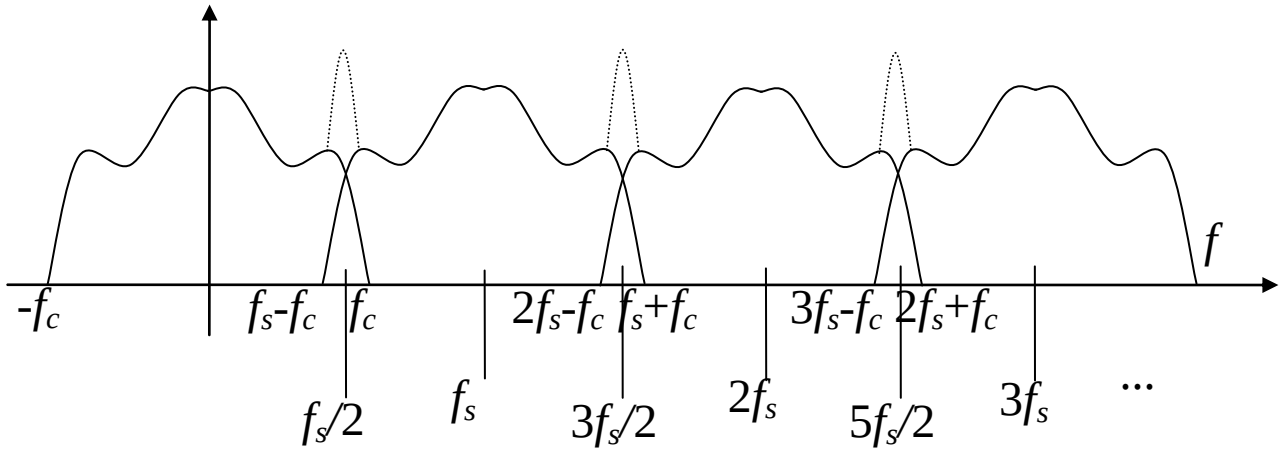
$f_s\delta_{f_s}(f)=f_s\sum_{n=-\infty}^{\infty}\delta(f-nf_s)$ - transformata Fouriera ciągu funkcji impulsowych powtarzanych co stały czas T_s (ciąg funkcji impulsowych powtarzanych co stałą częstotliwość, będącą częstotliwością próbkowania f_s, o wysokości uwzględniającej tę częstotliwość próbkowania).

Efektem splotu dowolnej funkcji z ciągiem funkcji impulsowych oddalonych o daną wartość ich argumentu jest suma powieleń (kopii) tej funkcji wokół każdej funkcji impulsowej w ich ciągu.

Widmo próbkowanego sygnału mierzonego składa się z sumy wielu powtarzających się widm sygnału mierzonego powielonych wokół częstotliwości f_s, $2f_s$, $3f_s$, $4f_s$ itd.. Z widma tego, po odfiltrowaniu wszystkich składowych o częstotliwościach większych od $\frac{f_s}{2}$ (po ograniczeniu widma otrzymanego przez powielania do jego pierwotnego zakresu) można odtworzyć, na drodze odwrotnej transformaty Fouriera, bez zniekształceń przebieg czasowy mierzonego sygnału tylko wtedy, gdy kolejne, sąsiednie powielenia jego widma nie zachodzą na siebie (czyli z widma identycznego z oryginalnym widmem tego sygnału), czego warunkiem jest spełnienie nierówności $f_c \leq \frac{f_s}{2}$, czyli $f_s \geq 2f_c$.



z takiego widma sygnału próbkowanego można otrzymać widmo sygnału oryginalnego, czyli informację pierwotną sprzed próbkowania, stosując filtr dolnoprzepustowy o częstotliwości granicznej $f_s/2$ (sąsiednie fragmenty widma są rozdzielone)



z takiego widma sygnału próbkowanego nie można otrzymać widma sygnału oryginalnego, czyli informacji pierwotnej sprzed próbkowania, niezależnie od stosowanego filtru dolnoprzepustowego (wokół częstotliwości $f_s/2$ sąsiednie widma nakładają się)

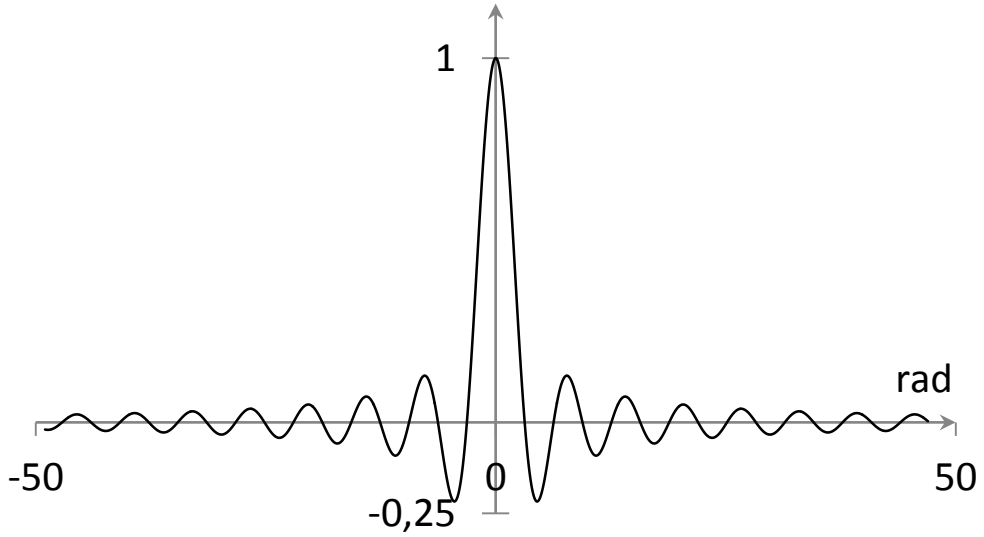
Nakładanie się widm przy próbkowaniu przebiegu czasowego sygnału można wyeliminować dwoma sposobami:
– albo zwiększyć częstotliwość próbkowania przynajmniej do wartości $2f_c$,
– albo ograniczyć widmo rejestrowanego sygnału do co najwyżej wartości $\frac{f_s}{2} = \frac{1}{2T_s}$.

Pierwszy sposób jest związany ze związkiem między wartością częstotliwości określającą okresowość widm próbkowanych przebiegów czasowych sygnałów f_s i zastosowanym czasem próbkowania dla konkretnego przypadku próbkowania T_s: $f_s = \frac{1}{T_s}$. Należy po prostu odpowiednio skrócić czas próbkowania, a więc zastosować odpowiednio szybsze układy próbkujące z pamięcią i przetworniki A/C, co może być przynajmniej kosztowne, o ile w ogóle fizycznie niewykonalne.

Drugi sposób jest związany ze świadomym zrezygnowaniem z informacji zawartej w tłumionych wysokoczęstotliwościowych składowych rejestrowanego sygnału (w efekcie nie mierzy się sygnału wyjściowego z czujnika, ale powstały z niego inny sygnał, pozbawiony składowych widmowych o częstotliwościach większych od $\frac{f_s}{2}$).

Kolejnym problemem związany z próbkowaniem – fakt jego ograniczonego trwania w czasie: kwantowane i kodowane są wartości pobierane z sygnału mierzonego co czas T_s przez skończony czas pomiaru T_w.

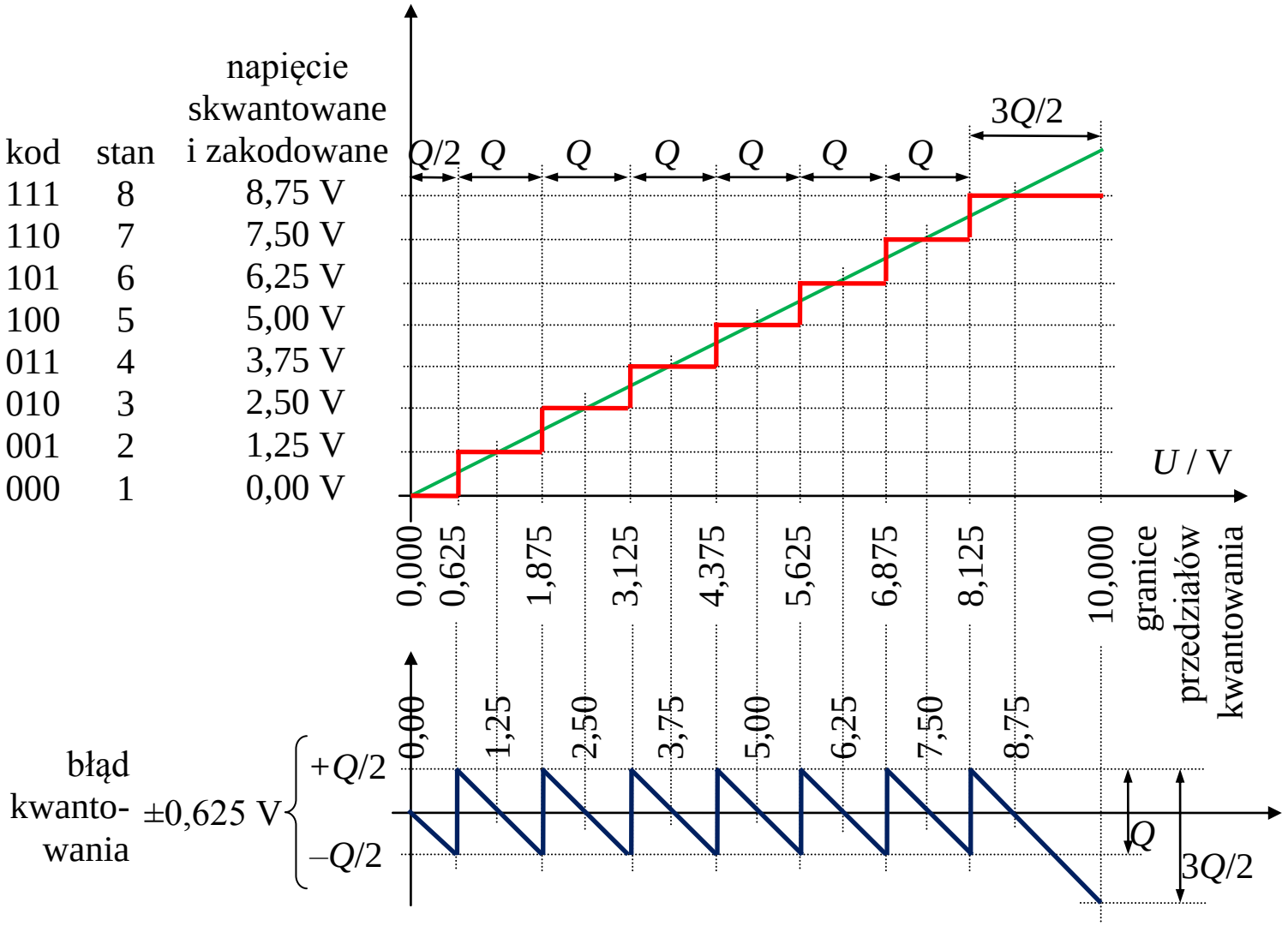
Przebieg czasowy $x(t)$ sygnału mierzonego jest mnożony przez ciąg funkcji impulsowych $\delta_{T_s}(t)$ powtarzających się co T_s oraz przez funkcję prostokątnego okna czasowego $w_{T_w}(t)=\begin{cases}1, & t\in[0,T_w]\\ 0, & t\notin[0,T_w]\end{cases}$ o długości T_w. Transformatą Fouriera iloczynu $x(t)\cdot\delta_{T_s}(t)\cdot w_{T_w}(t)$ w dziedzinie czasu jest w dziedzinie częstotliwości splot transformat $X(f)*f_s\delta_{f_s}(f)*W_{f_w}(f)$ – splot widma sygnału z ciągiem funkcji impulsowych o odpowiedniej wysokości i z funkcją $W_{f_w}(f)\approx\frac{\sin\pi fT_w}{\pi fT_w}$ – transformatą Fouriera funkcji okna $w_{T_w}(t)$. Miejsca zerowe funkcji $W_{f_w}(f)$ występują dla częstotliwości równych całkowitym wielokrotnościom odwrotności długości okna czasowego $f_w=\frac{1}{T_w}$, a wartość maksymalna dla częstotliwości 0 Hz.



Próbkowaniu przez czas T_w dużo dłuższy od czasu T_s odpowiada splot relatywnie bardzo „wąskiej”, „ostrej” funkcji typu $\frac{\sin \pi f T_w}{\pi f T_w}$ (bardzo podobnej do funkcji impulsowej) z ciągiem funkcji impulsowych powtarzających się co $f_s >> f_w$ i w efekcie splot ten może być przybliżony samym tylko ciągiem funkcji impulsowych jako jednym z czynników tego splotu – wpływ okna czasowego próbkowania może być pominięty, z tym mniejszym błędem analitycznym, im czas pobierania próbek jest większy od czasu próbkowania. W praktyce częstotliwość próbkowania bywa znacznie większa niż wartość minimalna, wynikająca z twierdzenia o próbkowaniu i ograniczona tylko szybkością działania elektronicznych układów próbkujących z pamięcią i przetworników A/C oraz pojemnością pamięci przyrządu pomiarowego.

Innym sposobem ograniczającym niekorzystny wpływ stosowania okna prostokątnego przy próbkowaniu jest zastosowanie innych okien, np. Hamminga, Hanna, Blackamana, Bartletta, Parzena, Welcha, Nuttalla, Bohmana, Gaussa, Kaisera, Tukey’a, Plancka, Streita, Poissona, Lanczosa itd., o nieliniowym kształcie w dziedzinie czasu i „niefalujących” charakterystykach częstotliwościowych. Możliwość stosowania tych okien jest często uwzględniana przez producentów przyrządów pomiarowych (częścią oprogramowania jest możliwość zmiany wartości pobieranych próbek odpowiednio do okna wybranego przez mierzącego spośród oferowanych w danym przyrządzie, przy czym zwykle liczba dostępnych okien jest związana z ceną przyrządu).

Kwantowanie i kodowanie w przetworniku A/C



Charakterystyka przetwarzania idealnego układu kwantowania o ośmiu stanach wyjściowych (funkcja nieliniowa (schodkowa)), za pomocą której kwantuje się sygnał napięciowy o zakresie zmienności od 0 do 10 V (unipolarny dodatni).

Rozdzielczość binarnego *n*-bitowego przetwornika A/C jako układu kwantowania – liczba stanów wyjściowych wyrażona w bitach.

Dla kodowania binarnego o *n* bitach (*n*-bitowego przetwornika binarnego) liczba stanów wyjściowych wynosi 2^n.

Liczba stanów wyjściowych dla kwantowania z kodowaniem binarnym wynosi
dla przetwornika 3-bitowego $2^3 = 8$,
dla przetwornika 8-bitowego $2^8 = 256$,
dla przetwornika 12-bitowego $2^{12} = 4\,096$,
dla przetwornika 16-bitowego $2^{16} = 65\,536$,
dla przetwornika 32-bitowego $2^{32} = 4\,294\,967\,296$.

Dla pokazanego przykładu funkcja przetwarzania ma $2^{n=3} - 1 = 8 - 1 = 7$ punktów decyzyjnych (poziomów progowych), wynoszących kolejno 0,625, 1,875, 3,125, 4,375, 5,625, 6,875 i 8,125 V.

Przedział kwantowania Q – różnica pomiędzy dwoma sąsiednimi poziomami progowymi (kwantowania) – najmniejsza różnica sygnału analogowego możliwa do rozróżnienia przez układ kwantowania.

Ogólny przypadek binarnego kodowania efektów kwantowania przetwornikiem *n*-bitowym: $Q = \frac{FSR}{2^n}$, $FSR = FSR_H - FSR_L$ (*Full Scale Range*) - pełny zakres przetwarzania – różnica między górną i dolną wartością ograniczającą zakres pomiarowy.

Dla podanego przykładu FSR = 10 V – 0 V = 10 V i $Q = \frac{10\text{V}}{2^3} = \frac{10\text{V}}{8} = 1{,}25\text{V}$.

Odpowiednio dla przetwornika 8-bitowego $Q = \frac{10\text{V}}{2^8} = \frac{10\text{V}}{256} \approx 39\,\text{mV}$,

dla przetwornika 12-bitowego $Q = \frac{10\text{V}}{2^{12}} = \frac{10\text{V}}{4096} \approx 2{,}44\,\text{mV}$,

dla przetwornika 16-bitowego $Q = \frac{10\text{V}}{2^{16}} = \frac{10\text{V}}{65536} \approx 152{,}5\,\mu\text{V}$,

dla przetwornika 32-bitowego $Q = \frac{10\text{V}}{2^{32}} = \frac{10\text{V}}{4294967296} \approx 2{,}33\,\text{pV}$.

Z kwantowaniem związany jest zawsze piłokształtny sygnał błędu.

Błąd kwantowania może być zmniejszony tylko przez zwiększenie liczby stanów, czyli rozdzielczości przetwornika, czyli przez zwiększenie liczby bitów składających się na jedno słowo kodowe.

Błąd kwantowania przyjmuje wartość zero tylko w tych przypadkach, gdy wartości sygnału analogowego są równe wartościom skwantowanym.

Sygnał błędu nazywa się czasem **nieoznaczonością** lub **szumem kwantowania**.

Do kodowania skwantowanych wartości mierzonego sygnału może służyć wiele kodów, w praktyce wykorzystuje się tylko kilka, z których najpopularniejszy jest **naturalny kod dwójkowy** (**zwykły binarny**), w którym liczba z przedziału [0,1] jest przedstawiana w postaci $N = \frac{a_1}{2^1} + \frac{a_2}{2^2} + \frac{a_3}{2^3} + ... + \frac{a_n}{2^n}$, gdzie n jest liczbą znaków kodu (bitów), a współczynniki a_1, a_2, a_3, ... a_n przyjmują wartości 0 lub 1.

Np. słowo 3-bitowe 011 przedstawia liczbę $N = \frac{0}{2^1} + \frac{1}{2^2} + \frac{1}{2^3} = 0 + 0{,}25 + 0{,}125 = 0{,}375$.

Przedział [0,1] jest tu przedziałem unormowanym, dla którego 0 odpowiada dolnej granicy zakresu przetwarzania, a 1 odpowiada górnej granicy zakresu przetwarzania.

Gdyby 3-bitowy przetwornik binarny kodował wartości sygnału w zakresie przetwarzania FSR = (0÷10) V = 10 V, to słowo kodowe 011 odpowiadało by wartości napięcia $U = FSR_d + N \cdot FSR$ = 0 V + 0,375 × 10 V = 3,75 V, co zgodne jest z wykresem charakterystyki przetwarzania idealnego układu kwantowania o ośmiu stanach wyjściowych kodującego 3-bitowymi słowami binarnymi wartości napięć z zakresu (0÷10) V.

Naturalny kod dwójkowy należy do klasy kodów ważonych – kodów z dodatnią wagą – każdy *k*-ty element kodu ma określoną wagę $\frac{1}{2^k}$ i żadna z tych wag nie jest ujemna.

Największą wagę ma bit pierwszy z lewej (aż $\frac{1}{2^1} = 0{,}5$ pełnego zakresu pomiarowego) i nazywa się **bitem najbardziej znaczącym** (**MSB** – most significant bit).
Najmniejszą wagę ma bit pierwszy z prawej (tylko $\frac{1}{2^n}$, np. dla przetwornika 3-bitowego $\frac{1}{2^3} = 0{,}125$ pełnego zakresu pomiarowego) i nazywa się **bitem najmniej znaczącym** (**LSB** – least significant bit).

Jak łatwo zauważyć:
- LSB ma wagę równą zdefiniowanemu przedziałowi kwantowania podzielonemu przez pełny zakres przetwarzania: $waga(\mathrm{LSB}) = \frac{Q}{FSR} = \frac{\frac{FSR}{2^n}}{FSR} = \frac{1}{2^n}$,
- przedział kwantowania jest iloczynem zakresu przetwarzania i wagi LSB: $Q = FSR \cdot waga(\mathrm{LSB})$.

Binarne słowo kodowe złożone z samych jedynek nie odpowiada wartości pełnego zakresu pomiarowego, lecz wartości mniejszej o wagę LSB.

Np. dla przetwornika 3-bitowego binarne słowo 111 daje wartość $N = \frac{1}{2^1} + \frac{1}{2^2} + \frac{1}{2^3} = 0{,}5 + 0{,}25 + 0{,}125 = 0{,}875$, co przy FSR = (0÷10) V = 10 V odpowiada wartości napięcia $U = FSR_d + N \cdot FSR$ = 0 V + 0,875 × 10 V = 8,75 V = 10 V – 1,25 V = 10 V – 10 V×0,125 V = $= FSR - FSR \cdot waga(\text{LSB}) = FSR(1 - waga(\text{LSB})) = FSR\left(1 - \frac{1}{2^n}\right)$.

Maksymalne napięcie zakodowane przez binarny przetwornik n-bitowy, równe $FSR\left(1 - \frac{1}{2^n}\right)$, jest zawsze mniejsze od wartości określonej jako zakres przetwarzania przetwornika A/C o wartość wagi LSB.

Dynamika (*DR* - dynamic range) *n*-bitowego przetwornika o kodowaniu binarnym - zwyczajowo - wyrażona w decybelach liczba stanów możliwych do przedstawienia za pomocą *n*-bitowych słów binarnych odniesiona do jednego stanu:

$$DR/\mathrm{dB} = 20\log_{10}\frac{2^n}{1} = 20\log_{10}2^n = 20n\log_{10}2 \approx 20n\times 0{,}301 = 6{,}02n.$$

Lub $DR/\mathrm{dB} = 20\log_{10}\frac{1}{\frac{1}{2^n}} = 20\log_{10}\frac{1}{waga(\mathrm{LSB})} = 20\log_{10}\frac{FSR}{Q}$. Stąd dla $DR/\mathrm{dB} = X$: $Q = \frac{FSR}{10^{\frac{X}{20}}}$.

przetwornik binarny	rozdzielczość przetwornika w bitach: *n*	liczba stanów (rozróżnialnych wartości): 2^n		waga LSB: $\frac{1}{2^n}$	*DR* / dB
1-bitowy	1	2^1 =	2	0,5	6,0
2-bitowy	2	2^2 =	4	0,25	12,0
3-bitowy	3	2^3 =	8	0,125	18,1
4-bitowy	4	2^4 =	16	0,062 5	24,1
8-bitowy	8	2^8 =	256	0,003 906 25	48,2
12-bitowy	12	2^{12} =	4 096	0,000 244 140 625	72,2
16-bitowy	16	2^{16} =	65 536	0,000 015 258 789 062 5	96,3
32-bitowy	32	2^{32} =	4 294 967 296	0,000 000 000 232 830 643 653...	192,7
64-bitowy	64	2^{64} =	$1{,}844\,674\times10^{19}$	0,000 000 000 000 000 000 054...	385,3

Np. dla binarnego przetwornika 16-bitowego w zakresie przetwarzania rozróżnia się 65536 poziomów. Oznacza to, że nierozróżnialne (szumem) są wartości zmieniające się o mniej niż 1/65536 (wartość wagi LSB) zakresu pomiarowego i

$$DR = 20\log_{10} 2^{16} = 20 \times 16 \log_{10} 2 \approx 320 \times 0{,}301 \approx 96{,}3\,\text{dB},$$

czyli że zakres dynamiczny tego przetwornika w stosunku do nierozróżnialnego szumu to 96,3 dB. Zmniejszając te różnice tylko do 1/10000 (ponad 6-krotnie) zawyża się poziom szumów do poziomu 10000 razy mniejszego niż rozpiętość zakresu pomiarowego, co daje $DR = 20\log_{10} 10^4 = 20 \times 4 \times 1 = 80\,\text{dB}$ – zakres dynamiczny binarnego przetwornika 16-bitowego to przynajmniej 80 dB (taką informację spotyka się w dokumentacji technicznej cyfrowych przyrządów pomiarowych z takimi przetwornikami). Odpowiednio $Q = \frac{FSR}{10^{\frac{80}{20}}} = \frac{FSR}{10^4} = \frac{FSR}{10000}$ – dla FSR = 10 V: Q = 1 mV.

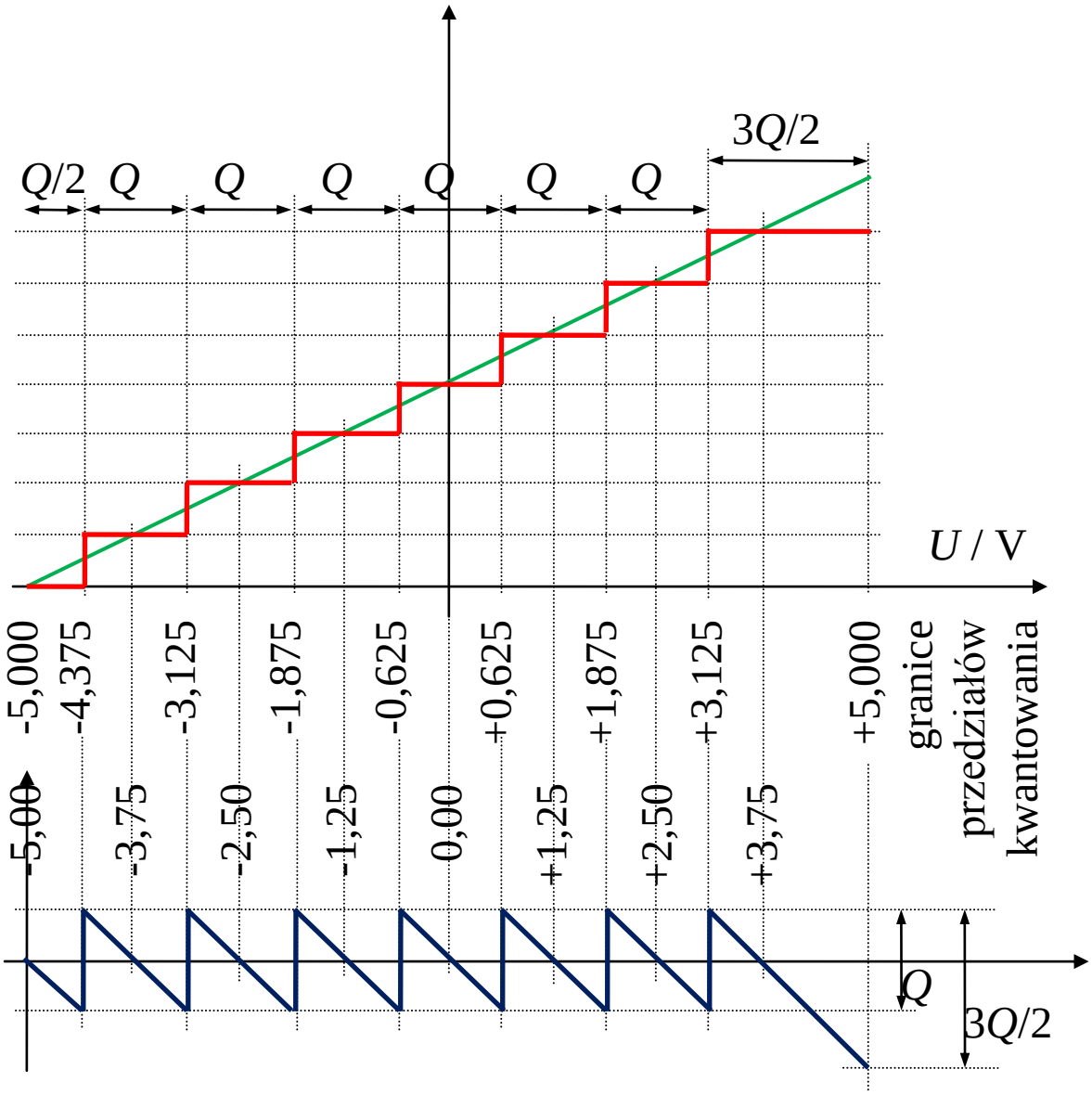
Oprócz naturalnego kodu dwójkowego szczególnie ważny dla systemów zbierania danych jest kod binarny (dwójkowy) przesunięty, powstający przez przesunięcie wykresu charakterystyki przetwornika o zakresie unipolarnym (jednoznakowym, np. (0÷10) V) tak, aby przetwornik miał zakres bipolarny (dwuznakowy symetryczny, np. (–5 ÷ +5) V) przy zachowaniu pełnego zakresu przetwornika *FSR* (w obu tych przypadkach *FSR* = (10 – 0) V = (+5 – (–5)) V = (5 + 5) V = 10V).

Dla symetrycznego zakresu bipolarnego pierwszy bit (MSB) określa znak sygnału:
– 0 odpowiada wartościom ujemnym,
– 1 odpowiada wartościom nieujemnym.

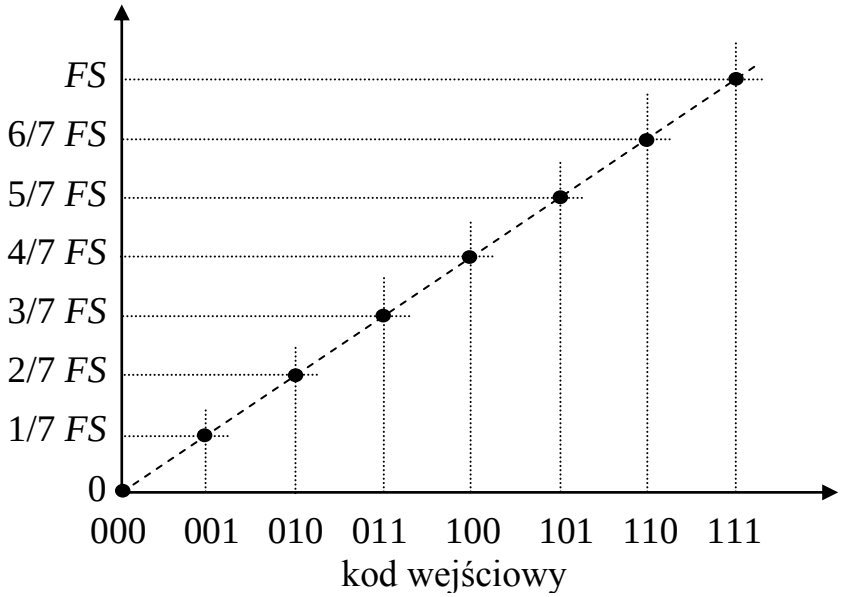
Np. dla zakresu (–5 ÷ +5) V i przetwornika 3-bitowego, MSB = 0 przypada dla napięć od –5 do –0,625 V kwantowanych kolejno jako –5, –3,75, –2,5 i –1,25 V, a MSB = 1 przypada dla napięć od –0,625 do +5 V kwantowanych kolejno jako 0, +1,25, +2,5 i +3,75 V.

Pierwszy bit (MSB) dla przetworników bipolarnych symetrycznych jest nazywany **bitem znaku** (bitem określającym zakodowany znak napięcia bipolarnego).

kod	stan	napięcie skwantowane i zakodowane
111	8	+3,75 V
110	7	+2,50 V
101	6	+1,25 V
100	5	0,00 V
011	4	-1,25 V
010	3	-2,50 V
001	2	-3,75 V
000	1	-5,00 V



Przetworniki C/A służą do zamiany cyfrowych słów kodowych na sygnały analogowe oddziałujące na świat fizyczny (do zobrazowania wyników pomiarów dla ludzi – mierniczych – badaczy, na sygnały sterujące pracą elementów wykonawczych w automatycznych systemach sterowania procesami). Przetworniki te czasem nazywa się dekoderami.



Charakterystyka przetwarzania idealnego, liniowego, unipolarnego, 3-bitowego przetwornika C/A

Każdemu słowu kodowemu odpowiada dokładnie jedna dyskretna wartość sygnału, **uważana za analogową**, z zakresu zmienności sygnału wyjściowego z przetwornika C/A (z przedziału $[FS_{min},FS_{max}]$ o długości $FS = FS_{max} - FS_{min}$ (Full Scale)).

Różnica pomiędzy otrzymywanymi wyjściowymi wartościami analogowymi n-bitowego przetwornika C/A to $\frac{FS}{2^n - 1}$.

Aby cokolwiek wnioskować o pojedynczych sygnałach lub konfrontować z sobą różne sygnały, należy je najpierw odpowiednio opisać ilościowo, tzn. opisać je odpowiednimi wartościami liczbowymi, czyli mówiąc trywialnie: odpowiednio je **zmierzyć** (w kontekście wcześniej przedstawionych treści: **wyznaczyć wartości parametrów diagnostycznych**).

Sygnał traktuje się jako zmienną losową – do ilościowego opisu można zastosować aparat i pojęcia statystyki matematycznej – miarami sygnału mogą być, zależnie od potrzeb danego przypadku analitycznego, różne parametry zmiennej losowej.

Stosowane są
- miary funkcyjne (zależne od określonego parametru, definicyjnie od czasu, ale także np. od częstotliwości, napięcia, prądu, masy itd.),
- miary punktowe (wartości miar funkcyjnych dla określonych wartości parametrów),

- miary wymiarowe (określone w konkretnych jednostkach, np. V, A, kg, s, m, m/s czy m/s^2),
- miary bezwymiarowe (w najprostszych przypadkach ilorazy miar wymiarowych),

- miary własne (dotyczą jednego (tego samego) zarejestrowanego sygnału),
- miary wzajemne (dotyczą dwu lub więcej sygnałów (zarejestrowanych w różnych punktach pomiarowych, w różnym czasie itp.)).

Miary własne sygnałów

Prowadząc eksperyment pomiarowy (np. w celu pozyskania informacji niezbędnej do diagnozowania) wykonuje się albo jeden pojedynczy pomiar (gdy stosowana metoda pomiarowa jest pewna) albo lepiej kilka powtórzeń pomiaru następujących bezpośrednio po sobie. Wyniki pomiarów mogą być wtedy traktowane jako realizacja x zmiennej losowej X. Sformułowanie „pojedynczy pomiar" dotyczy pomiaru dla jednej chwili czasu lub rejestracji przebiegu czasowego sygnału o odpowiedniej długości.

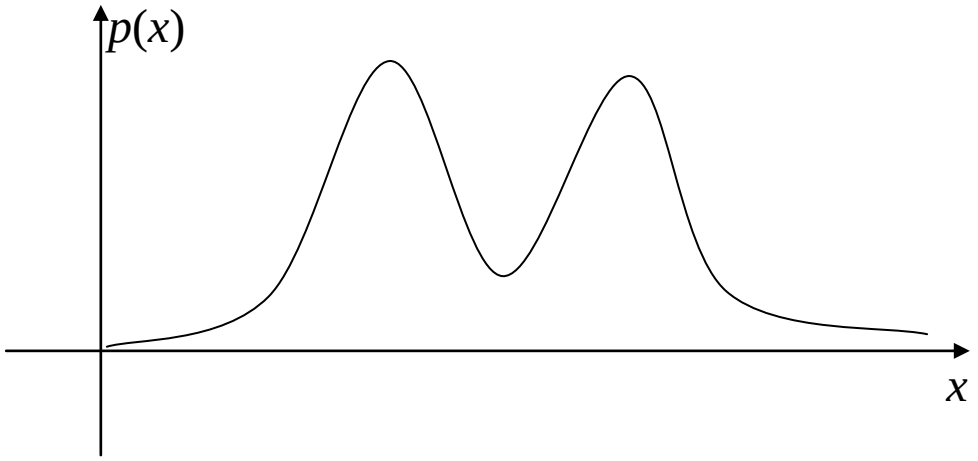
Wyróżnia się zmienne losowe:
- ciągłe – sygnał może przyjmować dowolne wartości – jest analogowy,
- dyskretne – sygnał może przyjmować tylko określone wartości – jest skwantowany.

Występowanie każdej wartości sygnału ma określone prawdopodobieństwo – zależność tego prawdopodobieństwa od wartości sygnału dla zmiennych losowych:
- ciągłych jest funkcją ciągłą i nazywa się gęstością prawdopodobieństwa $p(x)$,
- dyskretnych jest funkcją dyskretną i nazywa się funkcją prawdopodobieństwa $p(x_i)$.

Każda zmienna losowa może być opisana różnymi parametrami. Najważniejsze z nich to niektóre **parametry pozycyjne** oraz **momenty**.

Parametry pozycyjne:

– **wartość modalna** (moda, dominanta) – wartość x_0 zmiennej losowej X, dla której rozkład prawdopodobieństwa posiada maksimum lokalne;

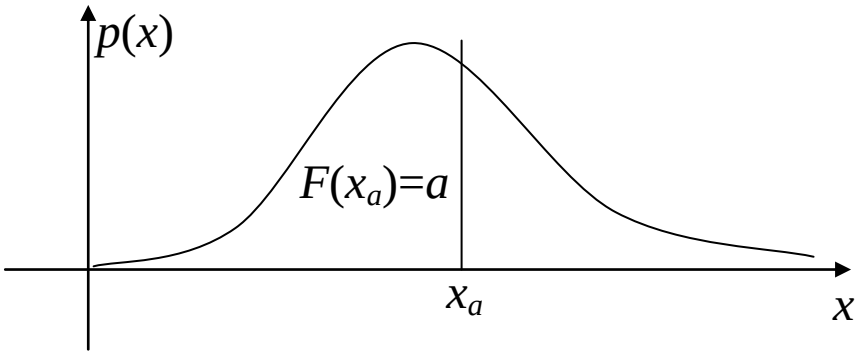


liczba wartości modalnych określa wielomodalność rozkładu:
- rozkłady **jednomodalne** mają jedno maksimum, przykładem jest rozkład normalny (Gaussa) o symetrycznym rozkładzie prawdopodobieństwa względem wartości oczekiwanej,
- rozkłady **dwumodalne** (o dwóch wartościach modalnych) opisujące np. uciąglone sygnały prostokątne,
- rozkłady **trzymodalne** itd.,
- rozkłady **antymodalne** (niemające wartości modalnych, „oczekiwanych”) opisujące sygnały o ciągle rosnących lub ciągle malejących wartościach);

– **kwantyl** *a*-tego rzędu – wartość x_a zmiennej losowej X, dla której dystrybuanta jest równa a:

$$F(x_a)=\int_{-\infty}^{x_a} p(x)\,\mathrm{d}\,x = a \quad \text{lub} \quad F(x_a)=\sum_{i=1}^{x_a} p(x_i) \geq a$$

(prawdopodobieństwo tego, że wartość zmiennej losowej nie będzie większa od x_a jest równe a):



  – dla a = ½ kwantyl nazywa się **medianą** (dla dyskretnej serii wartości wyników jest to ta z nich, która znajduje się w środku tej serii po uporządkowaniu jej rosnąco; dla serii o parzystej liczbie elementów bierze się średnią arytmetyczną dwóch takich środkowych wartości),
  – dla a = ¼ lub ¾ kwantyl nazywa się **górnym** lub **dolnym kwartylem**,
  – dla a = 0,1, 0,2, ..., 0,9 kwantyle nazywają się **decylami**;

– **prawdopodobne odchylenie zmiennej losowej od mediany** $q=\dfrac{x_{0,75}-x_{0,25}}{2}$.

Moment k-tego rzędu zmiennej losowej X względem jakiegoś punktu (liczby) c – wartość oczekiwana $\mathrm{E}(X-c)^k$:

$\int_{-\infty}^{\infty}(x-c)^k\,p(x)\,\mathrm{d}\,x$ dla zmiennych losowych ciągłych, gdzie $p(x)$ jest gęstością prawdopodobieństwa zmiennej losowej X,

$\sum_{i=1}^{\infty}(x_i-c)^k\,p(x_i)$ dla zmiennych losowych dyskretnych, gdzie $p(x_i)$ jest prawdopodobieństwem wystąpienia wartości x_i zmiennej losowej X.

Momenty zwyczajne dla $c = 0$: $m_k = \mathrm{E}X^k$.

W analizie sygnałów dla potrzeb praktyki eksperymentalno-pomiarowej znaczenie mają trzy momenty zwyczajne:

– pierwszego rzędu: $m_1 = \mathrm{E}X = \int_{-\infty}^{\infty} xp(x)\mathrm{d}\,x$ lub $\sum_{i=1}^{\infty} x_i\, p(x_i)$, czyli **wartość oczekiwana** (przeciętna, najbardziej prawdopodobna, „najczęściej występująca”) sygnału – miara poziomu (składowej stałej) sygnału,

– drugiego rzędu: $m_2 = \mathrm{E}X^2 = \int_{-\infty}^{\infty} x^2\, p(x)\mathrm{d}\,x$ lub $\sum_{i=1}^{\infty} x_i^2\, p(x_i)$, czyli **wartość średniokwadratowa** – jej pierwiastek kwadratowy jest **wartością skuteczną** sygnału – miarą energetyczności sygnału,

– „nieskończonego” rzędu: $\mathrm{E}X^k\big|_{k\to\infty} = \int_{-\infty}^{\infty} x^k\big|_{k\to\infty}\, p(x)\mathrm{d}\,x$ lub $\sum_{i=1}^{\infty} x_i^k\big|_{k\to\infty}\, p(x_i)$ – jego pierwiastek tego samego „nieskończonego” rzędu jest **wartością szczytową** sygnału – miarą rozpiętości (amplitudy) sygnału.

Mody, mediany i wartości oczekiwanej nie należy mylić:

- dla mody (dostępnej pomiarowo) rozkład prawdopodobieństwa ma maksimum,
- dla mediany (dostępnej pomiarowo) pola pod krzywą obrazującą rozkład prawdopodobieństwa na lewo (dystrybuanta) i prawo (uzupełnienie dystrybuanty do jedności) od niej są równe i wynoszą po ½,
- wartość oczekiwana jest obliczana (nie musi pokrywać się z wartością dostępną pomiarowo) jako średnia ważona z wagami będącymi prawdopodobieństwami występowania poszczególnych wartości.

Moda, mediana i wartość oczekiwana są równe dla rozkładów jednomodalnych symetrycznych, np. dla rozkładu normalnego.

Dla sygnałów przyjmujących równomiernie wartości zarówno dodatnie jak i ujemne, momenty zwyczajne bierze się w postaci $\int_{-\infty}^{\infty} |x|^k \, p(x) \mathrm{d}\, x$ lub $\sum_{i=1}^{\infty} |x|^k \, p(x_i)$, bo klasycznie wyznaczane momenty zwyczajne nieparzystych rzędów nie mają istotnego znaczenia (zwykle są bliskie zeru, niezależnie od postaci przebiegu czasowego sygnału).

Dla parzystych *k* postacie miar określone wzorami bez i z modułem są tożsame.

Odpowiednia miara amplitudy, związana z momentem zwyczajnym pierwszego rzędu nazywana jest **wartością średniego poziomu** sygnału.

Czasem wykorzystuje się także tzw. **wartość pierwiastkową** sygnału, określoną dla $k = ½$, w większym stopniu uwzględniającą małe wartości chwilowe sygnału (podczas gdy wartość oczekiwana wszystkie wartości chwilowe sygnału uwzględnia w równym stopniu, a wartość skuteczna w większym stopniu uwzględnia duże wartości chwilowe sygnału).

Momenty centralne dla $c = \mathrm{E}X$: $\mu_k = \mathrm{E}(X - \mathrm{E}X)^k$.

Dla symetrycznych rozkładów prawdopodobieństwa wszystkie momenty centralne nieparzystego rzędu mają wartość 0.

W praktyce eksperymentalno-pomiarowej znaczenie ma moment centralny drugiego rzędu: $\mu_2 = \mathrm{E}(X - \mathrm{E}X)^2 = \int_{-\infty}^{\infty} (x - m_1)^2 \, p(x) \mathrm{d}\, x = m_2 - m_1^2 \equiv \sigma^2$ lub $\sum_{i=1}^{\infty} (x_i - m_1)^2 \, p(x_i)$, czyli **wariancja sygnału**, oraz jego dodatni pierwiastek kwadratowy $\sigma = \sqrt{\mu_2}$, czyli **odchylenie standardowe** – miara rozproszenia rejestrowanych wartości sygnału wokół jego wartości oczekiwanej (poziomu).

Czasem stosuje się też moment centralny pierwszego rzędu: $\mu_1 = \mathrm{E}(X - \mathrm{E}X) = \int_{-\infty}^{\infty} (x - m_1) p(x) \mathrm{d}\, x$ lub $\sum_{i=1}^{\infty} (x_i - m_1) p(x_i)$, czyli **odchylenie przeciętne** – jako moment nieparzystego rzędu dla sygnału o symetrycznym rozkładzie prawdopodobieństwa (np. typu gaussowskiego) ma wartość 0 i nie obrazuje rozproszenia pomiarów wokół ich wartości oczekiwanej.

Momenty sygnału są jego miarami wymiarowymi → ich wymiar jest *k*-tą potęgą jednostki, w jakiej wyrażany jest sygnał → używając momentów nie pierwszego rzędu, aby móc je porównywać, zwykle używa się ich 1/*k*- tej potęgi. Np. dla momentów drugiego rzędu (wartości średniokwadratowej i wariancji) – pierwiastka kwadratowego (odpowiednio wartości skutecznej i odchylenia standardowego).

Czasami w praktyce nie stosuje się bezwzględnych wartości tych miar sygnałów, a tylko ich wartości względne, odniesione do pewnych umownych wartości, np.:

- zwykły stosunek wartości zarejestrowanej i umownej przemnożony przez 100 % (lub przez 1000 ‰), czyli wyrażanie wartości sygnału w procentach (lub w promilach) wartości umownej,
- logarytm stosunku wartości zarejestrowanej i umownej, czyli wyrażanie wartości sygnału w decybelach (dB):
  - $L\ \mathrm{dB} = 10 \log_{10} \frac{P}{P_o}$ dla wielkości „energetycznych” P (mocy, energii, ciepła),
  - $L\ \mathrm{dB} = 20 \log_{10} \frac{p}{p_o}$ dla wielkości „podstawowych” p (napięć, natężeń, ciśnień) ($P \sim p^2$).

Oprócz czystych momentów (miar wymiarowych) stosowane są też ich stosunki. Najczęściej wykorzystuje się stosunki dające miary bezwymiarowe:

– **współczynnik zmienności**: $\gamma = \frac{\sigma}{m_1} = \frac{\sqrt{\mu_2}}{m_1} = \sqrt{\frac{m_2}{m_1^2} - 1}$, często mnożony przez 100 % i pokazujący procentowy stosunek rozproszenia sygnału do jego wartości przeciętnej,

– **współczynnik asymetrii (skośność)**: $a = \frac{\mu_3}{\sigma^3} = \frac{\mu_3}{\sqrt{\mu_2^3}}$,

jeżeli $a < 0$ to rozkład prawdopodobieństwa wartości sygnału posiada asymetrię ujemną (lewostronną – więcej pomiarów sygnału jest mniejsze od wartości modalnej),

jeżeli $a = 0$ to rozkład prawdopodobieństwa wartości sygnału jest symetryczny względem wartości modalnej,

jeżeli $a > 0$ to rozkład prawdopodobieństwa wartości sygnału posiada asymetrię dodatnią (prawostronną – więcej pomiarów sygnału jest większe od wartości modalnej),

– **współczynnik spłaszczenia** (**eksces**, **kurtoza**): $e = \frac{\mu_4}{\sigma^4} - 3 = \frac{\mu_4}{\mu_2^2} - 3,$

jeżeli $e < 0$ to rozkład prawdopodobieństwa wartości sygnału jest bardziej rozmyty wokół wartości modalnej niż rozkład prawdopodobieństwa odpowiedni dla rozkładu normalnego $N(m_1,\sigma)$,

jeżeli $e = 0$ to rozkład prawdopodobieństwa wartości sygnału jest identyczny jak rozkład prawdopodobieństwa odpowiedni dla rozkładu normalnego $N(m_1,\sigma)$,

jeżeli $e > 0$ to rozkład prawdopodobieństwa wartości sygnału jest bardziej skupiony wokół wartości modalnej niż rozkład prawdopodobieństwa odpowiedni dla rozkładu normalnego $N(m_1,\sigma)$,

– **współczynnik kształtu**: $\gamma_K = \frac{\sqrt{m_2}}{m_1},$

– **współczynnik szczytu**: $\gamma_S = \frac{\sqrt[k]{m_k}\big|_{k\to\infty}}{\sqrt{m_2}},$

– **współczynnik impulsowości**: $\gamma_I = \frac{\sqrt[k]{m_k}\big|_{k\to\infty}}{m_1}.$

Dla sygnałów przyjmujących równomiernie wartości zarówno dodatnie jak i ujemne, dla których $m_1 = 0$, dopuszcza się zastąpienie m_1 przez wartość średniego poziomu sygnału.

Dla trzech ostatnich miar zachodzą zależności:

- $\gamma_I \geq \gamma_S \geq \gamma_K \geq 1$,
- $\gamma_I = \gamma_K \gamma_S$ bo $\dfrac{\sqrt[k]{m_k}\Big|_{k\to\infty}}{m_1} = \dfrac{\sqrt{m_2}}{m_1} \dfrac{\sqrt[k]{m_k}\Big|_{k\to\infty}}{\sqrt{m_2}}$.

Wszystkie wymienione miary sygnałów są przedstawione w postaci teoretycznej – zakładają dokładną znajomość jawnej postaci (w postaci wzoru analitycznego) rozkładu prawdopodobieństwa sygnału, co jest w praktyce niemożliwe.

Wartość pomiaru jest ustalana na podstawie realizacji eksperymentu – jak na podstawie skończonej liczby pomiarów estymować (wyznaczać empirycznie) przedstawione miary – parametry sygnałów?

Najlepszymi **estymatorami** (wielkościami wyliczanymi tylko na podstawie pomiarów i możliwymi do przyjęcia jako najlepsze oszacowania wartości teoretycznych) są te, które są:
- **nieobciążone** (wartość oczekiwana estymatora jest równa estymowanemu parametrowi),
- **najefektywniejsze** (błąd średniokwadratowy estymatora wobec estymowanego parametru jest najmniejszy spośród wszystkich innych estymatorów),
- **zgodne** (błąd średniokwadratowy estymatora wobec estymowanego parametru dla liczności próby rosnącej do nieskończoności maleje do zera).

Wykazuje się, że najlepszymi estymatorami przedstawionych miar sygnału są:

– dla wartości oczekiwanej (składowej stałej) – **średnia arytmetyczna** uzyskanego sygnału:

$$\bar{x} = x_{sr} = \frac{1}{N}\sum_{i=1}^{N} x_i = \frac{1}{T}\int_0^T x(t)\mathrm{d}t \equiv x_{av},$$

  – N - liczba pomiarów sygnału (tu liczba próbek pobieranych co czas próbkowania T_s),
  – T - przedział czasu t, w którym przeprowadzono obserwację sygnału dynamicznie zmieniającego się (np. natężenie prądu przemiennego, przyspieszenia drgań itd.) – czas uśredniania wyniku),
  – dla sygnałów okresowych T musi być równe całkowitej krotności okresu sygnału, a N oznacza liczbę pomiarów przypadających na tę krotność okresu ($NT_s = T$);

– dla wartości średniego poziomu wyrażenie: $x_{srp} = \frac{1}{N}\sum_{i=1}^{N}|x_i| = \frac{1}{T}\int_0^T |x(t)|\mathrm{d}t,$

– dla wartości pierwiastkowej wyrażenie: $x_p = \left(\frac{1}{N}\sum_{i=1}^{N}\sqrt{|x_i|}\right)^2 = \left(\frac{1}{T}\int_0^T \sqrt{|x(t)|}\mathrm{d}t\right)^2,$

– dla wartości skutecznej – pierwiastek kwadratowy estymatora wartości średniokwadratowej, czyli **średnia kwadratowa** uzyskanego sygnału: $x_{sk} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2} = \sqrt{\frac{1}{T}\int_0^T x^2(t)\mathrm{d}t} \equiv x_{RMS}$ (root mean square),

– dla wartości szczytowej wyrażenia:
  – górne $x_{sz}^{+} = \max_{i=\overline{1,N}}\{x_i\} = \max_{t\in(0,T)} x(t) \equiv x_{peak}^{+}$,
  – dolne $x_{sz}^{-} = \min_{i=\overline{1,N}}\{x_i\} = \min_{t\in(0,T)} x(t) \equiv x_{peak}^{-}$,
  – absolutne $x_{sz} = \max_{i=\overline{1,N}}\{|x_i|\} = \max_{t\in(0,T)}|x(t)| \equiv x_{peak}$,
  – międzyszczytowe $x_{msz} = x_{sz}^{+} - x_{sz}^{-} = x_{peak-to-peak}$,

– dla odchylenia standardowego – **średnie odchylenie kwadratowe** uzyskanego sygnału:

$$s = \sqrt{\frac{1}{N-1}\sum_{i-1}^{N}(x_i - x_{sr})^2} = \sqrt{\frac{1}{T}\int_0^T (x(t) - x_{sr})^2\,\mathrm{d}t},$$

– dla odchylenia przeciętnego wyrażenie: $d = \frac{1}{N-1}\sum_{i-1}^{N}(x_i - x_{sr}) = \frac{1}{T}\int_0^T (x(t) - x_{sr})\mathrm{d}t$,

– dla współczynnika zmienności: $\gamma = \frac{s}{x_{sr}}$ lub $\gamma = \frac{s}{x_{srp}}$,

– dla współczynnika asymetrii: $a = \frac{\frac{1}{N-1}\sum_{i-1}^{N}\left(x_i - x_{sr}\right)^3}{s^3} = \frac{\frac{1}{T}\int_0^T \left(x(t) - x_{sr}\right)^3 \mathrm{d}\,t}{s^3}$,

– dla współczynnika spłaszczenia: $e = \frac{\frac{1}{N-1}\sum_{i=1}^{N}\left(x_i - x_{sr}\right)^4}{s^4} - 3 = \frac{\frac{1}{T}\int_0^T \left(x(t) - x_{sr}\right)^4 \mathrm{d}\,t}{s^4} - 3$,

– dla współczynnika kształtu: $\gamma_K = \frac{x_{sk}}{x_{sr}}$ lub $\gamma_K = \frac{x_{sk}}{x_{srp}}$,

– dla współczynnika szczytu: $\gamma_S = \frac{x_{sz}}{x_{sk}}$,

– dla współczynnika impulsowości: $\gamma_I = \frac{x_{sz}}{x_{sr}}$ lub $\gamma_I = \frac{x_{sz}}{x_{srp}}$.

Dla sygnałów okresowych, czyli w pełni zdeterminowanych, okres lub jego całkowita wielokrotność reprezentuje cały sygnał, więc wyrażenia zawierające momenty centralne (na d, s, a i e) sygnałów próbkowanych w swoich mianownikach zamiast $N - 1$ mają N (liczy się te momenty nie dla próby, ale dla populacji sygnału jako zmiennej losowej).

Dystrybuanta wartości sygnału $P(A)$ – prawdopodobieństwo, z jakim wartość sygnału $x(t)$ lub x_i znajdzie się poniżej zadanej wartości A:

$$P(A) \equiv \mathrm{P}\left[x(t) \langle A\right] \approx \frac{1}{T}\sum_i {t_i}^P \qquad P(A) \equiv P\left[x_i \langle A\right] \approx \frac{1}{K}\sum_l {k_l}^P$$

t_i^P - czasy trwania przypadków gdy $x(t) < A$

k_i^P - liczby przypadków gdy $x_i < A$

T - czas trwania (pomiaru) badanego sygnału

K - całkowita liczba zarejestrowanych wartości badanego sygnału

Rozkład prawdopodobieństwa wartości sygnału $p(B)$ – prawdopodobieństwo, z jakim wartość sygnału $x(t)$ lub x_i znajdzie się w granicach $(B, B + \mathrm{d}B)$ w unormowaniu do długości $\mathrm{d}B$ tego przedziału:

$$p(B) = \frac{\mathrm{P}\left[B \langle x(t) \langle B + \mathrm{d}B\right]}{\mathrm{d}B} \approx \frac{1}{T \cdot \mathrm{d}B}\sum_i {t_i}^p \qquad p(B) = \frac{P\left[B \langle x_i \langle B + \mathrm{d}\,B\right]}{\mathrm{d}\,B} \approx \frac{1}{K}\sum_l {k_l}^p$$

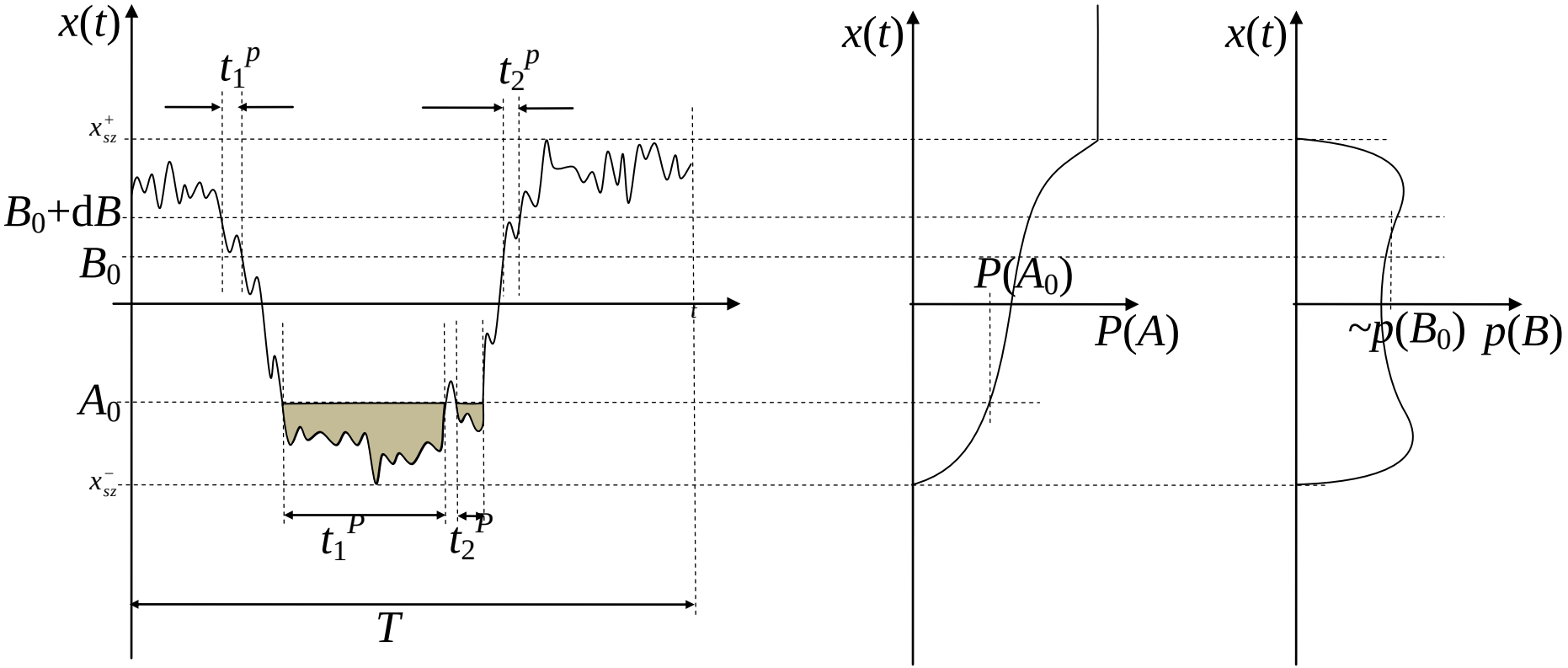
t_i - czasy trwania przypadków gdy $B < x(t) < B + \mathrm{d}B$

k_i^P - liczby przypadków gdy $B < x_i < B + \mathrm{d}B$

T - czas trwania (pomiaru) badanego sygnału

K - całkowita liczba zarejestrowanych wartości badanego sygnału

$\mathrm{d}B$ - znikomo mały przedział wartości sygnału


x(t)
t1p
t2p
x+sz
B0+dB
B0
A0
x−sz
t1P
t2P
T
t
x(t)
P(A0)
P(A)
x(t)
~p(B0)
p(B)

Częstotliwość Rice'a – częstotliwość przejścia sygnału ciągłego przez zadany poziom $x(t) = A = \text{const}$ w określonym kierunku, np. wartości większych od A – obliczana w ogólnym przypadku sygnałów przypadkowych z zależności: $f(A) = f(0)\dfrac{p(A)}{p(0)}$,

$f(0)$ - częstotliwość przejścia sygnału przez poziom $x(t) = 0$,
$p(A)$ i $p(0)$ - odpowiednie wartości gęstości prawdopodobieństwa wartości sygnału.

Dla sygnałów, dla których gęstości prawdopodobieństwa wartości opisuje się rozkładem normalnym, częstotliwość Rice'a dla poziomu $x(t) = 0$ przyjmuje wartości:

– dla wartości sygnału $f_x(0) = \dfrac{1}{2\pi}\dfrac{\dot{x}_{RMS}}{x_{RMS}}$,

– dla prędkości zmian wartości sygnału $f_{\dot{x}}(0) = \dfrac{1}{2\pi}\dfrac{\ddot{x}_{RMS}}{\dot{x}_{RMS}}$,

– dla przyspieszeń zmian wartości sygnału $f_{\ddot{x}}(0) = \dfrac{1}{2\pi}\dfrac{\dddot{x}_{RMS}}{\ddot{x}_{RMS}}$.

Na podstawie częstotliwości Rice'a oblicza się tzw. **współczynniki harmoniczności:**

przyspieszeń zmian wartości sygnału $H_{\ddot{x}} \equiv \frac{f_{\ddot{x}}}{f_{\dot{x}}} = \frac{\dddot{x}_{RMS} \cdot \dot{x}_{RMS}}{\ddot{x}_{RMS}^2}$,

prędkości zmian wartości sygnału $H_{\dot{x}} \equiv \frac{f_{\dot{x}}}{f_x} = \frac{\ddot{x}_{RMS} \cdot x_{RMS}}{\dot{x}_{RMS}^2}$.

Przyjmują one wartości z przedziału [1,∞), przy czym wartość 1 odpowiada procesowi harmonicznemu o ściśle określonym paśmie, a ∞ odpowiada białemu szumowi o nieskończonym paśmie.

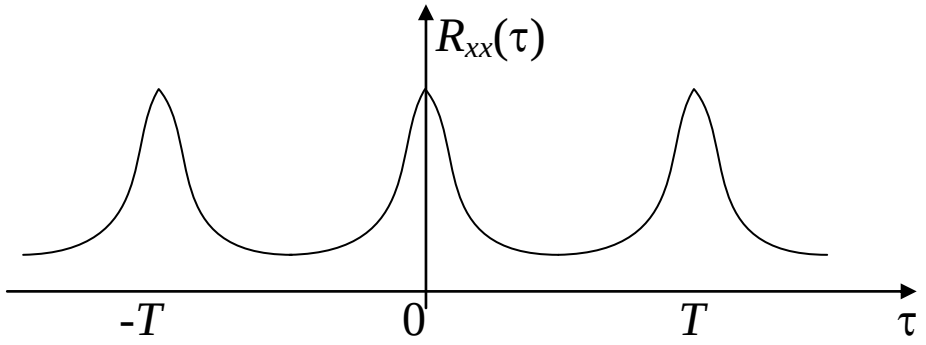
Funkcja autokorelacji sygnału – uogólniona wartość średniokwadratowa sygnału (uśredniona po czasie pomiaru suma iloczynów wartości fragmentów tego sygnału przesuniętych o pewien odcinek czasu).

Podaje współzależność między fragmentami tego sygnału opóźnionymi o odcinek czasu τ: $R_{xx}(\tau)=\frac{1}{2T}\int_{-T}^{T}x(t)x(t+\tau)\mathrm{d}t=\overline{x(t)x(t+\tau)}$, T - czas obserwacji porównywanych sygnałów (czas uśredniania dynamicznie zmiennych sygnałów) lub ewentualny okres sygnału, o ile sygnał jest okresowy (czas rejestracji sygnału przy jego powielaniu jest jego okresem, co dla sygnałów ergodycznych jest przyjmowane w sposób naturalny).

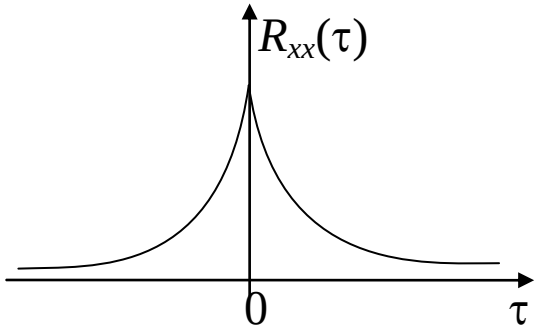
Aby uniknąć ujemnych wartości czasu funkcję autokorelacji przedstawia się w postaci $R_{xx}(\tau)=\frac{1}{T}\int_{0}^{T}x(t)x(t+\tau)\mathrm{d}t$.

Cechą funkcji autokorelacji jest własność separacji sygnału okresowego od nieokresowego.

Funkcja autokorelacji sygnału okresowego też jest funkcją okresową z maksymalnymi wartościami $R_{xx}(\tau)$ dla przesunięć sygnału o czasy będące całkowitymi wielokrotnościami okresu – czasu rejestracji (dla $\tau = kT$), a dla $\tau = 0$ funkcja autokorelacji jest po prostu wartością średniokwadratową sygnału.



Funkcja autokorelacji sygnałów nieokresowych też nie jest funkcją okresową, przy czym posiada maksimum globalne dla zerowego przesunięcia sygnału korelowanego (dla $\tau = 0$) i jest wtedy po prostu wartością średniokwadratową sygnału.



Gęstość widmowa mocy sygnału przedstawia w dziedzinie częstotliwości uśrednione po czasie pomiaru własności energetyczne sygnału i jest uśrednionym po czasie pomiaru kwadratem modułu transformaty Fouriera tego sygnału: $G_{xx}(f) \equiv \frac{1}{2T}|X(f)|^2$, gdzie $X(f) = |X(f)|e^{j\psi(f)} = \int_{-\infty}^{\infty} x(t)e^{-j2\pi ft}\mathrm{d}t$ jest **gęstością widmową wartości sygnału**.

Gęstość widmowa mocy sygnału jest funkcją rzeczywistą, a gęstość widmowa wartości sygnału jest funkcją zespoloną i jako taka ma rzeczywiste moduł i argument.

Gęstość widmowa mocy jest związana z funkcją autokorelacji zależnościami Wienera-Chinczyna: $R_{xx}(\tau) = \int_{-\infty}^{\infty} G_{xx}(f)e^{j2\pi f\tau}\mathrm{d}f$ i $G_{xx}(f) = \int_{-\infty}^{\infty} R_{xx}(\tau)e^{-j2\pi f\tau}\mathrm{d}\tau$.

Moduł gęstości widmowej wartości sygnału jest popularnie stosowanym w analizie dynamicznie zmiennych sygnałów widmem amplitudowo-częstotliwościowym, bardzo przydatnym w analizie okresowości (powtarzalności) rejestrowanego sygnału, a argument gęstości widmowej wartości sygnału jest widmem fazowo-częstotliwościowym.

Cepstrum $C(\tau)=\left|\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}(\lg G_{xx}(f))e^{-j2\pi ft}\mathrm{d}f\right|^2$ - szczególnie przydatne gdy w sygnale pojawiają się lub zanikają składowe harmoniczne.

Miary wzajemne sygnałów

Funkcja korelacji wzajemnej - współzależność między dwoma różnymi sygnałami $x(t)$ i $y(t)$ (podobieństwo w sensie statystycznym): $R_{xy}(\tau) = \frac{1}{2T}\int_{-T}^{T} x(t)y(t+\tau)\mathrm{d}t = \overline{x(t)y(t+\tau)}$,

T - czas obserwacji porównywanych sygnałów (czas uśredniania dynamicznie zmiennych sygnałów) lub ewentualny wspólny okres korelowanych sygnałów, o ile sygnały są okresowe (czas rejestracji sygnałów przy ich powielaniu jest ich okresem, co dla sygnałów ergodycznych jest przyjmowane w sposób naturalny).

Aby uniknąć ujemnych wartości czasu funkcję korelacji wzajemnej przedstawia się w postaci $R_{xy}(\tau) = \frac{1}{T}\int_{0}^{T} x(t)y(t+\tau)\mathrm{d}t$.

Dla sygnałów identycznych $x(t) = y(t)$ funkcja korelacji wzajemnej jest tożsama z funkcją autokorelacji.

Powszechnie wykorzystywaną cechą funkcji korelacji wzajemnej jest to, że wartość jej postaci unormowanej: $\rho_{xy}(\tau) = \frac{R_{xy}(\tau)}{\sqrt{R_{xx}(0)R_{yy}(0)}}$ przyjmuje wartości z przedziału <–1,1>, przy czym im bliższe są one wartościom granicznym, tym bardziej badane sygnały są do siebie podobne (dla $|\rho_{xy}| = 1$ procesy są identyczne w sensie statystycznym) i im bliższe zeru, tym badane sygnały są mniej podobne.

Splot sygnałów: $\otimes_{xy}(t) = x(t) * y(t) = \int_{-\infty}^{\infty} x(\tau) y(t-\tau) \mathrm{d}\tau$

$$\otimes_{xy}(\tau) = x(\tau) * y(\tau) = \int_{-\infty}^{\infty} x(t) y(\tau - t) \mathrm{d}t = \int_{-\infty}^{\infty} x(t) y(-t + \tau) \mathrm{d}t$$

(definicja dość podobna do funkcji korelacji wzajemnej, jednak nie identyczna – argument całkowania występujący w argumencie jednej ze splatanych funkcji jest brany jako przeciwny (funkcja jest odwracana względem chwili czasu przyjętej jako zero), całkowanie jest wykonywane po całym kontinuum czasowym i nie ma uśredniania iloczynu sygnałów w czasie).

Waga splotu w analizie sygnałów wynika z powszechnego wykorzystywania własności przekształceń całkowych, zwłaszcza Laplace'a i Fouriera, polegających na tym, że iloczyn oryginałów funkcji przekształcany jest w splot obrazów i odwrotnie, że splot oryginałów przekształcany jest w iloczyn obrazów (oryginałami są np. przebiegi czasowe sygnałów, a obrazami ich widma częstotliwościowe lub przedstawienia operatorowe).

W dobie powszechnego próbkowania sygnałów dokonywanego przez komputerowe (cyfrowe) systemy pomiarowe, przy wykonywaniu widm sygnałów niezwykle istotny jest splot z ciągami funkcji impulsowych.

Funkcja koherencji jest miarą spójności (podobieństwa źródeł) dwu sygnałów: $x(t)$ i $y(t)$: $\gamma_{xy}^2(f) \equiv \frac{\left|G_{xy}(f)\right|^2}{G_{xx}(f) \cdot G_{yy}(f)}$, gdzie G_{xy}, G_{xx} i G_{yy} są gęstościami widmowymi mocy sygnałów powiązanymi zależnościami Wienera-Chinczyna z odpowiednimi funkcjami korelacji.