

C 3246
0100

**ZASTOSOWANIA
MATEMATYKI
W BUDOWIE
MASZYN**



**OGÓLNOPOLSKIE
SEMINARIUM
NAUKOWE**

KOSZALIN - KOŁOBRZEG 19 - 20 X 1994

V OGÓLNOPOLSKIE SEMINARIUM NAUKOWE

**ZASTOSOWANIA MATEMATYKI
W BUDOWIE MASZYN**

Komitet Naukowy Seminarium

Prof. dr hab. inż. Wojciech Kacalak - Przewodniczący

Prof. dr hab. inż. Tadeusz Burakowski

Prof. dr hab. inż. Tadeusz Lechowski

Prof. dr hab. inż. Zbigniew Polański

Prof. dr hab. inż. Janusz Rybak

Prof. dr hab. inż. Jan Szadkowski

Prof. dr hab. inż. Wojciech Tarnowski

Prof. dr hab. inż. Krzysztof Wawryn

Prof. dr inż. dr h.c. Kazimierz Wieczorowski

Dr inż. Tadeusz Zaborowski

Prof. dr inż. Henryk Żebrowski

Komitet Organizacyjny Seminarium

Dr inż. Piotr Stępień - Przewodniczący

Dr inż. Tadeusz Zaborowski

Dr inż. Ryszard Lewkowicz

Dr inż. Stanisława Plichta

Mgr inż. Ryszard Ściegienka

Wydanie publikacji dofinansowane przez
Ministerstwo Edukacji Narodowej

Prace recenzowane przez członków
Komitetu Naukowego Seminarium
Materiały przygotowane do druku przez autorów

SEKCJA TECHNOLOGII MASZYN KBM
POLSKIEJ AKADEMII NAUK

KOMISJA BUDOWY MASZYN PAN ODDZIAŁ W POZNANIU
WYDZIAŁ MECHANICZNY WYŻSZEJ SZKOŁY INŻYNIERSKIEJ
W KOSZALINIE

INSTYTUT BADAŃ I EKSPERTYZ NAUKOWYCH - IBEN
GORZÓW WIELKOPOLSKI

ZASTOSOWANIA MATEMATYKI W BUDOWIE MASZYN

V OGÓLNOPOLSKIE SEMINARIUM NAUKOWE

pod redakcją naukową
prof.dr hab.inż. WOJCIECHA KACALAKA

KOSZALIN - KOŁOBRZEG, 19 - 20 października 1994

Zeszyt Naukowy Wydziału Mechanicznego nr 18

ISSN 0860-0325

Przewodniczący Uczelnianej Rady Wydawniczej

Szymon Pałkowski

Redaktor

Jerzy Ignaciuk

Projekt okładki

Tadeusz Walczak

Wydano za zgodą Rektora Wyższej Szkoły Inżynierskiej
w Koszalinie

Wydawnictwo Uczelniane Wyższej Szkoły Inżynierskiej
w Koszalinie

75-620 Koszalin, ul. Raławicka 15-17

Koszalin 1994, wyd.1, ark. wyd. 12,5, nakład 150, format B-5

Przedmowa

Seminaria naukowe pod nazwą "ZASTOSOWANIA MATEMATYKI W BUDOWIE MASZYN" mają już wieloletnią tradycję - obecnie jest piątym seminarium ogólnopolskim. Przed organizacją pierwszych seminariów pojawiały się pytania - o to czy takie seminarium jest potrzebne - bowiem, jak argumentowano, we wszystkich współczesnych pracach naukowych wykorzystuje się znane osiągnięcia matematyki. Tymczasem coraz częściej okazuje się, iż rozwój matematyki i informatyki jest tak szybki i wielokierunkowy, że wielu pracowników naukowych z trudem, a często i ze znacznym opóźnieniem zaczyna wykorzystywać najnowsze teorie matematyczne oraz rozwiązania i narzędzia informatyczne. Im lepsze zaś staje się wykorzystanie metod matematycznych, tym silniejsza jest inspiracja do rozwoju ich postaw. Już to dostatecznie przesądza o potrzebie i celu seminarium.

Dylemat, co jest ważniejsze - teoria czy eksperyment, jest problemem pozornym. Niewielu jest bowiem ortodoksyjnych empirystów z przekonania, podobnie jak też stosunkowo niewielu zwolenników nauki bez doświadczenia. Dobrym empirystą można być jedynie wówczas, gdy posiada się skłonność do pracy z wieloma alternatywnymi teoriami i umiejętności posługiwania się możliwie pełną gamą narzędzi do poznawania i opisywania nowych teorii. Warto nawet podkreślić, iż różnorodność teorii jest naturalną cechą wszelkiej wiedzy i poszukiwanie jedyne prawdziwego mega-opisu jest błędnym wyborem.

Doświadczenia krajów przodujących w rozwoju cywilizacyjnym potwierdzają, że nie ma nic bardziej praktycznego od dobrych teorii. Tymczasem zdarzają się opinie, nawet w środowiskach naukowych, typu "to tylko teoria". Zapewne jest to skutkiem często pośpiesznego

przedstawiania nadmiernie uproszczonych opisów rzeczywistości, niekiedy z zastosowaniem metod mało zgodnych z charakterem problemu.

Nauki matematyczne można uznać za najważniejszą prądziecinę rozwoju współczesnej cywilizacji. Wcześniej niektóre, wówczas uważane za abstrakcyjne, idee matematyczne długo czekały na swoje zastosowania. Obecnie modelowanie matematyczne tak się rozpowszechniło, iż tworzą się oddzielne dyscypliny zastosowań. Badania operacyjne, statystyka, matematyczne planowanie eksperymentów, informatyka, symulacja komputerowa, sztuczne sieci neuronowe, metody przetwarzania informacji niepewnej i nieścisłej to przykłady działów matematyki stosowanej i zastosowań informatyki.

Dawniej najpełniej z nowych narzędzi matematycznych korzystały nauki opisujące podstawy rzeczywistości takie jak fizyka i wywodzące się z niej dyscypliny jak np.: mechanika, elektrotechnika i elektronika, optyka, technika jądrowa. Obecnie również budowa maszyn, inżynieria produkcji i poszczególne technologie dla rozwiązywania już nawet typowych problemów wykorzystują najnowsze osiągnięcia nauk matematycznych, wręcz poszukują nowych narzędzi matematycznych i integrują się z nimi do tego stopnia, że powstają nowe dyscypliny wiedzy i techniki.

W wielu przypadkach implementacja złożonych teorii i metod matematycznych jest trudnym zadaniem. Celem seminariów z zakresu zastosowań matematyki w budowie maszyn jest tworzenie pomostu pomiędzy nowymi osiągnięciami matematyki i informatyki a ich zastosowaniami w budowie maszyn.

prof.dr hab.inż. Wojciech Kacalak

Spis treści

Przedmowa.....	5
1. Wojciech Kacalak, Anna Kacalak: Biologiczne podstawy tworzenia sztucznych sieci neuronowych.....	11
2. Wojciech Kacalak, Krzysztof Wawryn: Modyfikacje zwiększające efektywność algorytmów uczenia konkurencyjnego sztucznych sieci neuronowych w wybranych zastosowaniach technicznych.....	21
3. Szymon Grymek: Neuronowe modele obiektów dynamicznych.....	33
4. Zbigniew Polański, Marek Radwański: Zastosowanie metod analizy procesów stochastycznych w badaniach okresowości struktury materiału	39
5. Edward Miko: Stochastyczny model konstituowania chropowatości powierzchni w procesie frezowania czołowego ostrzami o zaokrąglonym narożu.....	47
6. Roman Ulbrich: Możliwości zastosowania analizy procesów stochastycznych do oceny zjawisk ciepłno- przepływowych mieszanin wielofazowych	55
7. Jerzy Kozak, Lucjan Dąbrowski, Marek Rozenek: Modelowanie matematyczne impulsowej obróbki elektro- chemicznej (PECM).....	63
8. Konrad Łubkowski, Artur Jarkiewicz: Symulacja komputerowa drażenia elektrochemicznego elektrodą kształtowaną.....	73

9. Tadeusz Zaborowski: Model matematyczny drgań powstających przy szlifowaniu zębów kół zębatych sposobem Nilesa.....	83
10. Krzysztof Tubielewicz, Andrzej Zaborski: Matematyczna analiza krzywych udziału nośnego liniowego	91
11. Bronisław Choroszy: Wzór na umowny kąt ścinania dla toczenia wibracyjnego	99
12. Stefan Bartkiewicz: Wybory złożone i oceny rozmyte.....	107
13. Wojciech Kacalak: Wprowadzenie do wnioskowania rozmytego w zastosowaniu do analizy cech procesu szlifowania	115
14. Wojciech Kacalak: Metodyka przetwarzania informacji niepewnej i tworzenia algorytmów do sterowania procesami szlifowania z wykorzystaniem teorii zbiorów rozmytych	123
15. Wojciech Kacalak: Algorytm rozmytej kompensacji nieregularnych zakłóceń w układach technologicznych	133
16. Józef Gawlik, Krzysztof Karbowski, Adam Ruszaj: Koncepcja sterowania wydajnością procesu obróbki ECM z wykorzystaniem teorii zbiorów rozmytych	145
17. Ryszard Filipowski: Identyfikacja badanego procesu liniowym równaniem regresji w oparciu o rachunek macierzowy	153
18. Wojciech Tarnowski: Poliptymalizacja ułatwia formalne ujęcie problemów projektowych w systemach CAD	165

19. Krzysztof Wawryn, Józef Drabarek:
Poszukiwanie uszkodzeń w urządzeniach mechaniczno-
elektronicznych173
20. Bogdan Wilczyński:
Application of numerical methods for shape optimization
of notches in machine or structural components181
21. Lucjan Dąbrowski, Jerzy Kozak:
Projektowanie komputerowe elektrod w obróbce
elektrochemicznej (ECM)189
22. Krzysztof Tubielewicz, Andrzej Zaborski:
Analiza zmian stanu naprężeń własnych
i innych cech eksploatacyjnych warstwy wierzchniej
poddanej procesowi zużycia201

BIOLOGICZNE PODSTAWY TWORZENIA SZTUCZNYCH SIECI NEURONOWYCH

Wiele dyscyplin techniki, nauki przyrodnicze, medycyna i nauki ekonomiczne w coraz większym stopniu korzystają z najnowszych metod komputerowego rozwiązywania złożonych problemów. Ważnym etapem rozwoju metod sztucznej inteligencji było opracowanie systemów ekspertowych. Umożliwiły one automatyzację wnioskowania o stanie obiektów na podstawie zbiorów praw nimi rządzących i pewnego zasobu danych. Wadą systemów ekspertowych jest niemożność interpretacji danych wykraczających poza zasoby bazy. Osiągnięcia współczesnej biologii, a zwłaszcza neurobiologii stworzyły podstawy sztucznych sieci neuronowych, które dzięki możliwości uczenia się mogą interpretować informacje niepełne, nieściśle i niepewne.

1. Charakterystyka problemu

Komputerowe metody rozwiązywania złożonych problemów stają się współcześnie najważniejszym narzędziem nauki i jej zastosowań. Prognozy firm decydujących o kierunkach rozwoju techniki komputerowej stwierdzają, że do roku 2000 prawie połowa zautomatyzowanych maszyn, urządzeń technologicznych, zestawów pomiarowych i diagnostycznych oraz systemów wspomagających zarządzanie wykorzystywać będzie cyfrowe i analogowe sztuczne sieci neuronowe.

Opracowanie podstaw systemów ekspertowych umożliwiło ich wykorzystanie w wielu ważnych zastosowaniach, zwłaszcza w diagnostyce. Systemy te umożliwiły automatyzację wnioskowania o stanie obiektów i procesów na podstawie zbiorów praw nimi rządzących i pewnego zasobu danych. Istnieje co prawda możliwość poszerzenia zbioru danych systemów ekspertowych, ale zasada wnioskowania pozostaje ta sama - rozpoznawanie i identyfikacja przebiega wyłącznie przy ścisłej zgodności charakterystyk rozpoznawanych z charakterystykami zapisanymi w bazie danych.

Typową cechą rozwoju cywilizacji jest coraz powszechniejsza automatyzacja oraz dążenie do coraz wyższej elastyczności systemów technicznych, diagnostycznych i pomiarowych. Powoduje to, iż coraz bardziej prawdopodobne stają się takie sytuacje, w których dla nowych warunków i nowych metod konieczne jest automatyczne dokonywanie doboru parametrów obiektów i procesów bez dysponowania wyczerpującym analitycznym rozwiązaniem problemu [1, 2, 3, 4, 5, 6, 7]. Systemy ekspertowe w takich sytuacjach tracą swoją przydatność.

* Wyższa Szkoła Inżynierska w Koszalinie, ul. Raławicka 15/17, 75-620 Koszalin

** Pomorska Akademia Medyczna, Szczecin, ul Rybacka 1, 70-204 Szczecin

Najnowszym narzędziem opracowanym dla rozwiązywania takich problemów są sztuczne sieci neuronowe, zwane również sieciami neuropodobnymi lub sieciami neuralnymi.

W badaniach doświadczalnych i zastosowaniach praktycznych efektywność konwencjonalnych metod analizy problemów staje się drastycznie niewystarczająca, gdy ilość możliwych rozwiązań wielowymiarowych problemów zaczyna się wykładniczo lub potęgowo powiększać do tak dużych liczb, że staje się, to problemem nawet dla bardzo wydajnych komputerów. Ponadto trudno jest akceptować typowe algorytmy poszukiwań rozwiązań optymalnych lub zbliżonych do optymalnych, w których wskutek małej efektywności, liczba sprawdzanych rozwiązań jest wielka. Tych negatywnych cech nie posiadają sieci neuronowe dzięki swej umiejętności "uczenia się".

W przypadku wdrażania sieci do konkretnego zastosowania zamiast wprowadzania reguł odpowiednich dla danej dziedziny, które wtedy jeszcze nie są znane lub nie mogą być określone, uczymy sieć na przykładach jednostkowych. Faza uczenia nie musi zresztą być fazą wyodrębnioną. W niektórych zastosowaniach nauczanie następuje w trakcie poszukiwań rozwiązania najlepszego. Większość badanych problemów ma w swej naturze charakter probabilistyczny, a okres zbierania danych jest ograniczony. To powoduje, że ich opis następuje na podstawie rozmytych zbiorów informacji niepełnych i losowych.

2. Biologiczne wzorce tworzenia sztucznych sieci neuronowych

Tkanka nerwowa, a zwłaszcza OUN (ośrodkowy układ nerwowy) jest najdoskonalszym przykładem samoorganizującego się układu. W tworzeniu oprogramowania zwanego siecią neuronową wykorzystuje się wiele cech neuronów biologicznych, a zwłaszcza takie cechy jak:

- ♦ progowość reakcji,
- ♦ występowanie połączeń pobudzających i hamujących (synaps), o określonej lub zmiennej sile połączenia na synapsie,
- ♦ rozprzestrzenianie się pobudzeń neuronów.

Podstawowym przejawem życia neuronu jest zdolność do generowania i przewodzenia potencjałów elektrycznych [8, 9]. W stanie spoczynku występuje ujemny bioelektryczny potencjał spoczynkowy, będący wynikiem przestrzennej separacji jonów przez błonę komórkową, co wymaga wydatkowania pewnej energii.

Na poziomie jednostki sieci - neuronu (biologicznego) działa prawo "wszystko albo nic", opisujące skutki działania bodźca depolaryzującego błonę neuronalną. Bodźce o zbyt małej wartości wywołują depolaryzacje wygasające przed powstaniem potencjału progowego. Warunkiem powstania lokalnego potencjału czynnościowego (iglicowego) jest więc odpowiednio wysoki poziom depolaryzacji skumulowanej w czasie i w przestrzeni. Dalsze pobudzenie nie zmienia już amplitudy powstałego potencjału, której wartość jest cechą

charakterystyczną dla danego neuronu. Istotną cechą potencjału czynnościowego jest jego zdolność do rozprzestrzeniania się wzdłuż błony neuronalnej.

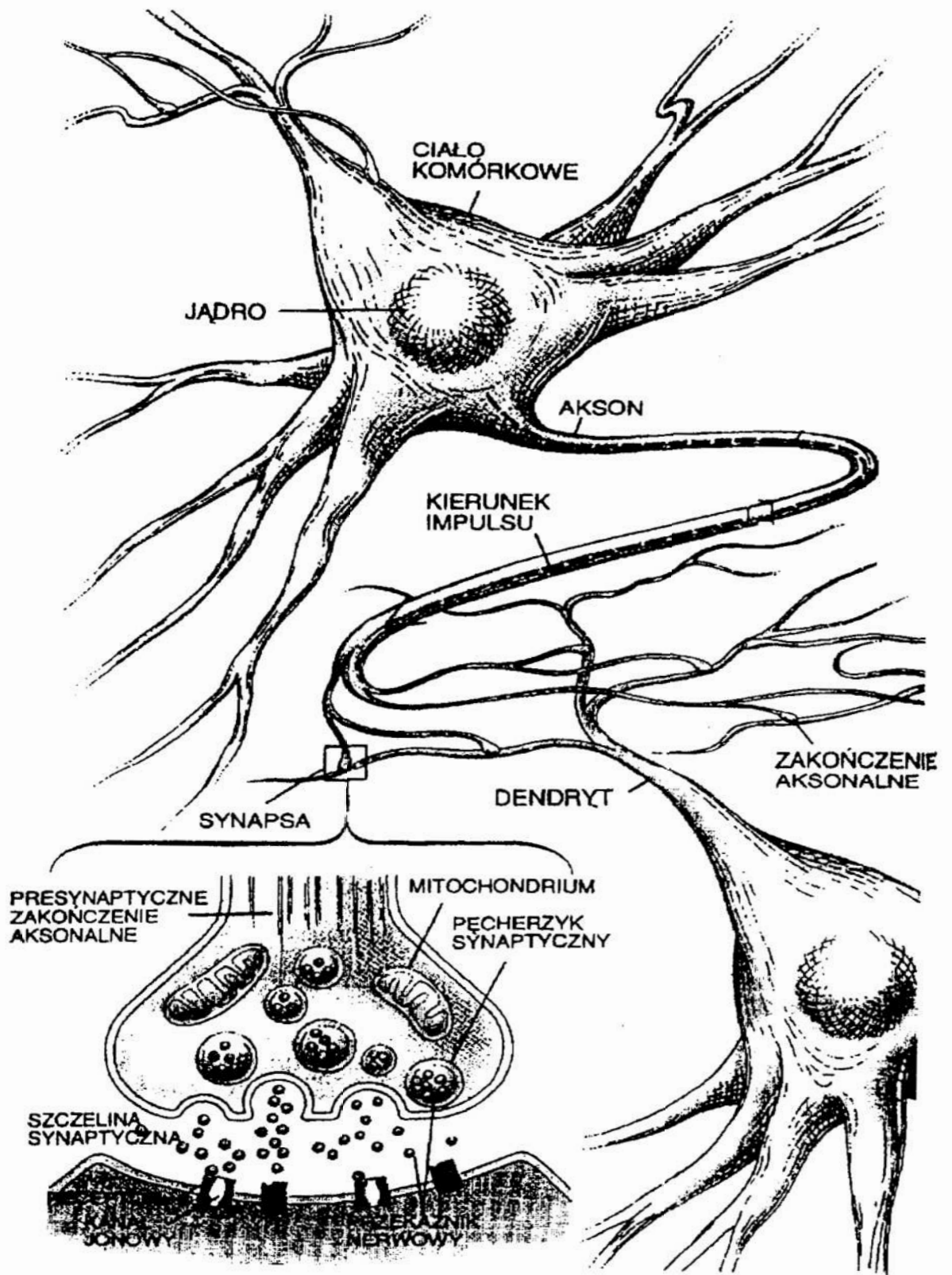
Zjawiska te są skutkiem szybkich zmian w przestrzennym rozkładzie jonów głównie sodu i potasu w neuronie i jego otoczeniu. Zwiększenie przepuszczalności błony neuronu dla jonów sodu, zwane aktywacją sodową wywołuje szybkie narastanie wartości potencjału iglicowego (0,5 ms). Wygasanie potencjału jest wynikiem jednoczesnej inaktywacji sodowej i aktywacji potasowej - prądu odkomórkowego.

Przekazywanie informacji wewnątrz i między komórkami odbywa się poprzez kanały jonowe - dzięki wbudowanym w błony komórkowe białkom tworzącym mikropory. Technika stabilizacji potencjału skrawka błony komórkowej [10, 11] umożliwiła badanie pojedynczych kanałów jonowych. Wykazano, że działanie kanałów ma charakter dyskretny (impulsowo-kwantowy). Kanały sodowe bramkowane przez potencjał są odpowiedzialne za szybką fazę narastania potencjału czynnościowego w neuronach. Zmiana potencjału błony jest skutkiem procesów chemicznych, można zatem stwierdzić, że działają one jak niezwykle szybkie, chemicznie bramkowane przełączniki prądu. Kanały jonowe są tylko częścią złożonego systemu oddziaływań, w którym pewną rolę wypełniają również wtórne przekaźniki, czyli cząsteczki zdolne do przenoszenia sygnałów z powierzchni komórki do jej wnętrza.

Neurony w różnych obszarach układów nerwowych mają różnorodne kształty, lecz ich budowa i fizjologia są podobne. Typowym elementem wejścia są, zazwyczaj liczne, dendryty neuronu, charakteryzujące się wysokim progiem pobudliwości i powolnym przewodzeniem impulsu. Elementem wyjścia neuronu jest akson, który jest główną drogą komunikowania z innymi komórkami nerwowymi, łącząc się najczęściej z ich dendrytami. Łącze to zwane synapsą nerwową nie jest połączeniem strukturalnym lecz czynnościowym. Szczelina występująca między zakończeniem aksonu i dendrytu (synapsa aksono-dendrytyczna) jest obszarem przekaźnictwa chemicznego. W skład synapsy wchodzi: zakończenie aksonu zwane elementem presynaptycznym, szczelina synaptyczna o szerokości ok. 20nm, oraz element postsynaptyczny, tworzony przez błony zakończeń dendrytów (rys. 1).

Szczelina synaptyczna jest drogą dyfuzji neurotransmitera (przekaźnika chemicznego), uwalnianego kwantowo z pakietów zakończeń presynaptycznych. Uwalnianie to następuje w zależności od zmian potencjału elektrycznego tego obszaru (sprzężenie elektrochemiczne). Przejmowanie sygnałów chemicznych przez receptory elementu postsynaptycznego prowadzi do wytwarzania, zależnie od charakteru transmitera, pobudzających lub hamujących potencjałów postsynaptycznych.

Systemy biologiczne zorganizowane są hierarchicznie. Neurony bardziej odległe od receptora (odległość może być wyrażona liczbą synaps) odpowiadają za bardziej przetworzone informacje, które będąc skutkiem jednego bodźca dochodzą do nich wieloma odrębnymi drogami.



Rys. 1. Schemat elementów neuronów i ich połączeń

3. Ważniejsze różnice między biologicznymi i sztucznymi sieciami neuronowymi

Funkcjonalne różnice między biologicznymi a komputerowymi sieciami neuronowymi, mimo starań twórców algorytmów opisujących sztuczne sieci neuronowe, nadal są jednak głębokie [8, 10, 11]. Do najważniejszych można zaliczyć:

- ♦ wielką liczbę elementów podstawowych (neuronów) - mózg człowieka zawiera ponad 10^{10} neuronów, z których każdy połączony jest synapsami z dziesiątkami tysięcy innych neuronów,
- ♦ neurony komputerowe mogą symulować działanie progowe a także pewny jeszcze zbiór prostych funkcji, natomiast neurony biologiczne realizują funkcje o dużym stopniu złożoności,
- ♦ organizacja układów biologicznych, ich struktura wewnętrzna i dynamika są znacznie bardziej złożone, przy tym budowa, rozmiary oraz właściwości neuronów biologicznych są bardzo zróżnicowane,
- ♦ mechanizmy uczenia się w sieciach neuronowych są proste, natomiast w organizmach żywych bardzo skomplikowane - od najprostszego jakim jest habituacja polegająca na wyzbywaniu się pewnych reakcji, przez klasyczne odruchy warunkowe, naukę metodą prób i błędów, a następnie wpajanie aż do uczenia się przez rozumowanie,
- ♦ mechanizmy gromadzenia informacji i ich wykorzystywania są w układach biologicznych znacznie bardziej zróżnicowane niż w systemach komputerowych, inne jest bowiem działanie pamięci krótkotrwałej i długotrwałej, a także zróżnicowany jest proces przekształcania i konsolidowania pamięci w różnych fazach uczenia się,
- ♦ w organizmach żywych podczas uczenia się możliwe są zmiany siły oddziaływań synaptycznych między neuronami, co przesądza o niepowtarzalnych cechach osobowości.

Sprzężenia zwrotne w sieci neuronowej są niezbędne dla realizacji poboru danych. U człowieka odpowiada to wyobrażeniom oraz jest podstawą zjawiska świadomości.

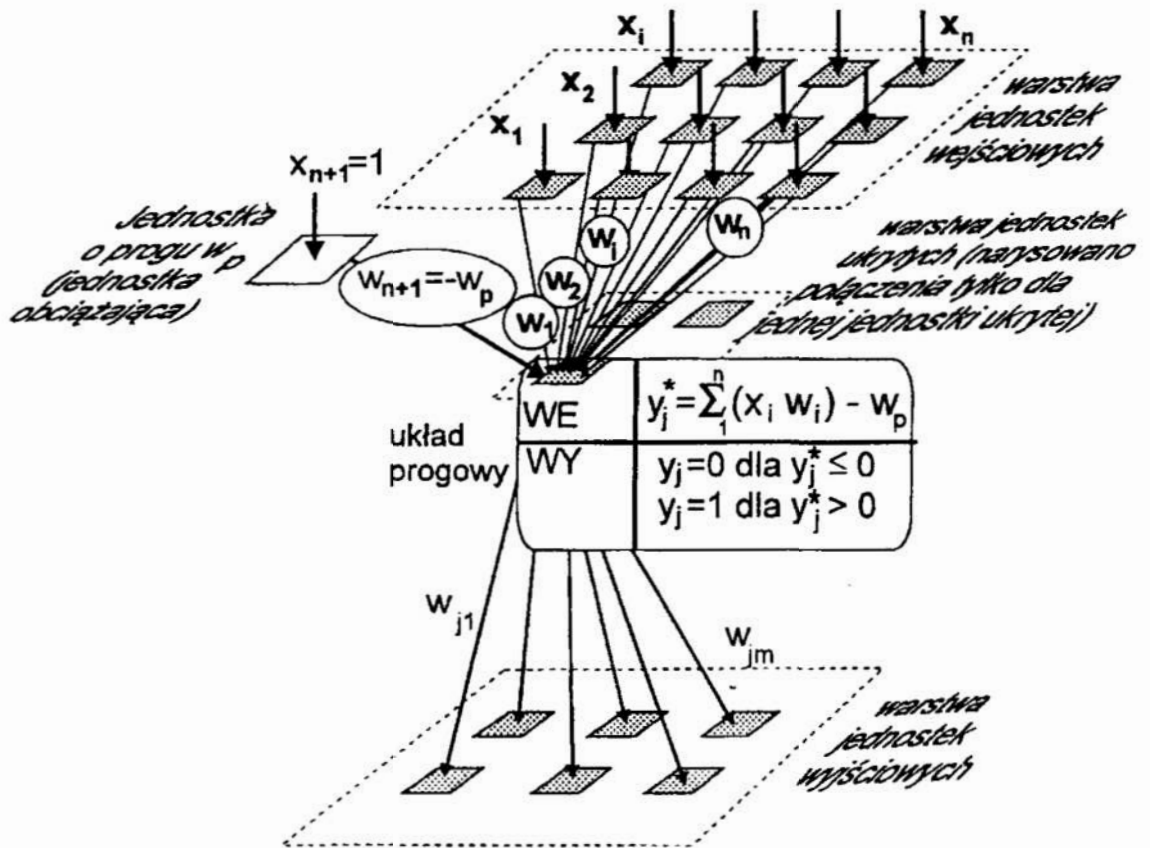
4. Analogie między neuronem a prostymi modelami technicznymi

Wszystkie komórki organizmów żywych oprócz tego, że zawierają zapis genetyczny swojej struktury, mają jeszcze jedną cechę wspólną. Zmiana zachowania (stanu komórki) może nastąpić tylko wtedy, gdy jej potencjał elektryczny przekroczy pewną wartość progową, zwaną progiem pobudzenia. Ta zmiana stanu jest wywoływana wpływami zewnętrznymi i to na ogół z komórek sąsiednich.

Reakcję na wpływy zewnętrzne, w sztucznych sieciach neuronowych można wyrazić wzorem:

$$y = | \left[\sum_{i=1}^n (x_i w_i) + w_{n+1} \right], \text{ gdzie } w_{n+1} = -w_p$$

oraz przedstawić w formie schematu (rys. 2).



Rys. 2. Schemat reakcji na wpływy zewnętrzne jednostki w modelu sieci neuronowej

Na rysunku 2 x_1, x_2, x_n są wpływami zewnętrznymi (np. impulsami z innych neuronów), a w_1, w_2, w_n parametrami (wagami) charakteryzującymi wrażliwość neuronu.

Suma iloczynów $(x_i w_i)$ jest sumą wpływów zewnętrznych i dopóki jest mniejsza od progu pobudzenia, to sygnał wyjściowy jest zerowy, co oznacza brak reakcji na wpływy zewnętrzne. Dobierając wagi i progi można modelować dowolną funkcję logiczną. Ważną cechą jest to, że nie jest konieczne dobieranie wag i progów przez użytkownika sieci. Działanie sieci może rozpocząć się z dowolnymi wagami i progami. Następnie przez sukcesywne przekazywanie wzorców wejścia i wyjścia (lub ocenę jakości rozwiązania) można modyfikować wagi tak, aby uzyskiwać wyniki zbliżone do wzorców lub rozwiązania coraz lepsze. W ten sposób sieć uczy się rozpoznawać i rozwiązywać w przyszłości nawet nieznanne sytuacje - potrafi zatem generalizować na podstawie procesu uczenia się.

Każda jednostka o dodatnim progu w_p i pewnej liczbie połączeń n może być zastąpiona przez jednostkę o progu zerowym z $n+1$ połączeniami, z których to dodatkowe połączenie ma wagę $w_{n+1} = -w_p$ i przychodzi od innej jednostki, która ma na wyjściu zawsze 1. Zabieg taki polegający na zamianie progu w wagę nazywa się obciążaniem sieci.

Typowa, najprostsza sieć neuronowa składa się z trzech warstw, przy czym każdy neuron danej warstwy połączony jest ze wszystkimi neuronami warstw sąsiednich. Neurony w danej warstwie nie są połączone ze sobą. Pierwsza warstwa zawiera jednostki wejściowe, odpowiedzialne za przyjęcie dostarczanej informacji. Podobnie działają biologiczne jednostki neuronowe ze zmodyfikowanymi zakończeniami w formie receptorów. Warstwa druga zawiera jednostki "ukryte", spełniające funkcję podobną do neuronów kojarzeniowych i pośredniczących OUN. Trzecia warstwa zawiera jednostki wyjściowe. Jest to oczywiście znaczącym uproszczeniem w stosunku do rozbudowanych połączeń komórek nerwowych w organizmach żywych, lecz wystarczającym dla wielu praktycznych zastosowań.

Wzorcem trudnym do naśladowania są łąca synaptyczne w biologicznych sieciach neuronów. W sztucznych sieciach neuronowych każdy z neuronów danej warstwy jest połączony z licznymi jednostkami warstw sąsiednich. Umożliwia to kumulowanie i rozprzestrzenianie sygnałów. Jest to podobieństwem procesów konwergencji i dywergencji w sieciach biologicznych. Dywergencja następuje wskutek występowania rozgałęzień aksonalnych. Dzięki temu następuje zwielokrotnienie torów równoczesnego przepływu informacji. Powstają też obwody zamknięte, w których krążenie impulsów powoduje ich wzmacnianie. Ten proces w OUN człowieka jest odpowiedzialny za konsolidację pamięci.

W organizmach żywych już na poziomie komórkowym występują odrębne systemy uczenia się - system uczenia świadomego i system uczenia utajonego. Odpowiadające temu zmiany w sile połączeń synaptycznych (zachodzące podczas uczenia się) mogą następować w efekcie jednoczesnego pobudzenia neuronu pre- i postsynaptycznego lub, w drugim przypadku, bez pobudzenia neuronu postsynaptycznego, ale za to z pobudzeniem neuronu wstawkowego (modulującego).

Współczynniki wag, decydujące o modyfikacji przekazywanych danych w neuronach sieci sztucznych są uproszczonym odpowiednikiem procesów pobudzenia i hamowania w połączeniach synaptycznych. Stan neuronu jest wypadkową oddziaływań pobudzających i hamujących.

Największą trudnością w naśladowaniu biologicznych sieci neuronowych jest duża różnorodność w sposobie przekazywania informacji między neuronami. Istotną cechą jest to, że w przypadku szerokich szczelin synaptycznych (rzędu 20nm), następuje zmiana nośnika informacji z potencjału elektrycznego na chemiczny.

Potencjały postsynaptyczne mające charakter depolaryzacyjny mogą prowadzić do depolaryzacji krytycznej, czego skutkiem jest wyzwolenie potencjału

czynnościowego, są nazywane postsynaptycznymi potencjałami pobudzającymi. Potencjały postsynaptyczne mogą mieć też charakter hiperpolaryzacyjny. W takim przypadku przez zwiększenie spoczynkowej różnicy potencjałów (progu) następuje działanie hamujące pobudzenie neuronu.

Do innych istotnych różnic między neuronami biologicznymi i ich odpowiednikami komputerowymi można zaliczyć zróżnicowanie pobudliwości elektrycznej i chemicznej neuronów biologicznych w obszarze samego neuronu a także złożoność mechanizmów czasowego i przestrzennego kumulowania pobudzeń, zwłaszcza wówczas, gdy neuron ma wejście zarówno pobudzające, jak i hamujące. Interesującą właściwością, zasługującą na wykorzystanie w komputerowych sieciach neuronowych, jest występowanie różnych typów hamowania presynaptycznego.

5. Algorytmy uczenia sieci neuronowych

Ogólnie algorytmy uczenia sieci neuronowych można podzielić według kryterium nadzorowania na procedury uczenia nadzorowanego i bez nadzorowania [12]. W procedurach uczenia nadzorowanego dla każdego wzorca wejściowego potrzebny jest wynik w postaci wzorca wyjściowego. W procedurach uczenia nienadzorowanego wagi sieci dobiera na podstawie badania obiektywnej jakości wyniku, a nie na podstawie analizy wzorców wejściowych i wzorców wynikowych.

Najprostszym algorytmem uczenia nadzorowanego jest algorytm wstecznej propagacji [12], [13]. Trenowanie sieci według tego algorytmu polega, dla przyjętego zbioru danych wejściowych, na porównywaniu odpowiadających im wielkości wyjściowych z wielkościami wzorcowymi dla określonych danych wejściowych. Następnie zmienia się wagi połączeń neuronów tak, aby zmniejszyć błąd. Zmiana wagi następuje proporcjonalnie do stosunku zmiany błędu do zmiany wagi.

Algorytm wstecznej propagacji składa się z następujących kroków:

- ▶ obliczenia pochodnej błędu względem aktywności jednostki wyjściowej,
- ▶ obliczenia pochodnej błędu względem sygnału wejściowego do jednostki wyjściowej,
- ▶ obliczenia pochodnej błędu względem wagi połączenia do jednostki wyjściowej,
- ▶ obliczenia pochodnej błędu względem aktywności jednostki poprzedzającej,
- ▶ propagacji powyższych kroków wstecz poprzez sieć.

Algorytm wstecznej propagacji jest przydatny, gdy dysponuje się licznym zbiorem wzorców. Okazuje się być skutecznym w sieciach wielowarstwowych, zwłaszcza dla nieliniowych związków między wejściem i wyjściem [12, 13]. Algorytm ten został zastosowany w sieciach do rozpoznawania obrazów [1, 2, 4, 5] np. znaków pisma, obiektów w ruchu, predykcji cech procesów ekonomicznych [6], maksymalizacji złożonych procesów chemicznych, identyfikacji

preparatów biologicznych, optymalizacji i filtrowaniu sygnałów sterujących, sterowaniu pracą robotów w warunkach zmiennej konfiguracji otoczenia [3, 4], sterowaniu pracą elastycznych systemów pomiarowych, od których wymaga się odporności na niekonwencjonalne zakłócenia, diagnostyki złożonych obiektów technicznych np. silników spalinowych, przekładni, a także diagnostyki różnorodnych procesów technologicznych.

Procedury uczenia sieci bez nadzorowania (bez wzorców) polegają na poprawianiu wyniku na podstawie analizy jakości reprezentacji. Jedną z procedur uczenia bez nadzorowania zwana metodą czynników głównych jest oparta na założeniu, że w przypadku występowania korelacji aktywności jednostek wejściowych, nie jest potrzebne osobne opisywanie każdej z nich. Wystarczy określenie czynników, charakteryzujących wiele jednostek. Zmniejsza się wtedy liczba jednostek w sieci, co znacznie obniża czas obliczeń. W procesie uczenia takiej sieci zadaje się zadanie wytworzenia na wyjściu wzorca zgodnego z wzorcem wejścia. Nie są zatem potrzebne wzorce wyjścia, co oznacza uczenie bez nadzorowania.

Inne algorytmy zwane algorytmami uczenia konkurencyjnego polegają, na tym, że traktuje się jednostki w warstwie ukrytej jako współzawodniczące o to, która z nich lepiej będzie reprezentować wzorzec wejściowy. Wtedy dla różnych wzorców wejściowych poszczególne jednostki ukryte "specjalizują się" w reprezentowaniu innego wzorca.

Ważną modyfikacją tego algorytmu dokonana przez Teuvo Kohonena [14, 15, 16, 17], polega na tworzeniu w jednostkach sąsiadujących podobnej reprezentacji wzorców wejściowych jak w jednostce wygrywającej. Dzięki temu algorytm staje się bardziej zgodny z naturalnym mechanizmem tworzenia topograficznych map w korze wzrokowej, co w praktycznych zastosowaniach oznacza zmniejszenie wrażliwości na nieprawidłowe działanie pojedynczej jednostki ukrytej.

Najnowszymi algorytmami nad którymi trwają jeszcze prace badawcze są algorytmy hierarchicznego kodowania populacyjnego, w których informacja jest przekazywana nie przez pojedyncze jednostki lecz przez pewną populację neuronów. Możliwe jest wtedy badanie obiektów wielowymiarowych, przy czym zakłócenia w pewnej małej grupie jednostek nie mają dużego wpływu na wynik globalny.

6. Literatura

1. *Koppen M., Nickolay B., Schwarze G.*: Anwendung Nueronalen Netze für die Texturklassifikation in gemusterten Materialien. *Vision&Voice Magazine* 6, 3/ 1992.
2. *Nickolay B., Vollmerhaus D.*: Automatisierte visuelle Inspektion auf der Basis einer hierarchischen Bildauflösung. *ZWF*, 1/1992, s.3437.

3. *Ritter H., Martinez T., Schulten K.*: Neuronale Netze. AddisonWesley GmbH, (II wydanie), 1992.
4. *Schöneburg E.*: Industrielle Anwendung Neuronale Netze. Addison Wesley (Deutschland) GmbH, 1993.
5. *Schöneburg E., Hansen N., Gawelczyk A.*: Neuronale Netzwerke. Markt und Technik. 1992.
6. *Schöneburg E., Niess J., Sautter U.*: Aktienkurs-Prognose mit Neuronale Netzwerken. Chip Plus 7/1991.
7. *Kacalak W.*: Zastosowania sieci neuronowych w budowie i eksploatacji maszyn. Mechanik 2/1994, s.45-49.
8. *Traczyk W.Z., Trzebski A.*: Fizjologia człowieka z elementami fizjologii stosowanej i klinicznej. Państwowy Zakład Wydawnictw Lekarskich, Warszawa 1989.
9. *Villee C.A.*: Biologia. Państwowe Wydawnictwo Rolnicze i Leśne. Warszawa 1990.
10. *Pener R., Neher E.*: The patch-clamp technique in the study of secretion. Trends in Neurosciences, vol.12, nr 4, pp.159-163, IV/1989.
11. *Unwin N.*: The structure of ion channels in membranes of excitable cells. Neuron, vol.3, nr 6, pp.600-612, IX/1989.
12. *Tadeusiewicz R.*: Sieci neuronowe. Akademicka Oficyna Wydawnicza. Warszawa 1993.
13. *Hinton G.,E.*: Connectionist learning procedures. Artificial Intelligence, vol.40. Nr.13, pp.185-234, 1989.
14. *Kohonen T.*: Self-Organized Formation of Topologically Correct Feature Maps", *Bio. Cybern.*, 43, 1982, pp. 59-69.
15. *Kohonen T.*: Selforganization and associative memory. Berlin, Springer Verlag 1984.
16. *Kohonen T.*: The Self-Organizing Map, Proceedings of the IEEE, special issue on Neural Networks, vol.78, Nr. 9, 1990, pp.1464-1480.
17. *Kohonen T.*: Unsupervised learning algorithms. Proceedings of International Conference "Neural Networks-Biological Computers or Electronic Brains", Lyon, 1990.

Biological principles of neural networks

Summary

Biological theory and properties of neuron and nervous system. A possible organization of animal memory and learning. Differences between biological and computer neural networks. Analogies between a neuron and simple technical models. Learning and generalization in neural networks. Neural networks models of learning and adaptation. Features of neuron networks and examples of their applications.

MODYFIKACJE ZWIĘKSZAJĄCE EFEKTYWNOŚĆ ALGORYTMÓW UCZENIA KONKURENCYJNEGO SZTUCZNYCH SIECI NEURONOWYCH W WYBRANYCH ZASTOSOWANIACH TECHNICZNYCH

Sieci neuronowe znajdują coraz więcej zastosowań w budowie i eksploatacji maszyn. Problem komiwojażera (TSP - travelling salesperson problem) jest problemem optymalizacji kombinatorycznej. Problem ten pojawia się w licznych zagadnieniach technicznych, zwłaszcza w optymalizacji procedur kontrolnych i optymalizacji procesów technologicznych. Nie jest znane jego rozwiązanie analityczne. Zastosowanie sztucznych sieci neuronowych pozwala uzyskiwać rozwiązania zbliżone do optymalnych. Opracowane przez autorów modyfikacje algorytmu uczenia konkurencyjnego sieci zwiększają efektywność poszukiwań.

1. Charakterystyka problemu

Problem komiwojażera [1], [2], [3], [4], [5] pojawia się w coraz większej liczbie zastosowań technicznych. Wynika to między innymi z dążenia do dużej elastyczności systemów wytwórczych i coraz pełniejszej automatyzacji produkcji wyrobów wytwarzanych w małych seriach. Ważnym zastosowaniem sieci neuronowych w technologii maszyn jest np. zagadnienie minimalizacji drogi przemieszczeń zespołów układu technologicznego lub pomiarowego w warunkach, gdy liczba położeń jest duża i położenia te tworzą mało regularny układ. Problem ten polega zatem na minimalizacji sumarycznej drogi przemieszczeń z powrotem do punktu początkowego.

Innym przykładem może być wybór kolejności produkcji n wyrobów w sytuacji, gdy czas i koszty dostosowania systemu produkcyjnego do wytwarzania kolejnych wyrobów zależą od typu wyrobu produkowanego ostatnio. W tym przypadku, w graficznej interpretacji problemu, odległości między punktami mają wymiar inny niż geometryczny. Podobne problemy występują w kontroli jakości, gdy o klasyfikacji decyduje wiele cech o różnych kosztach ich wyznaczania i różnym prawdopodobieństwie dyskwalifikacji wyrobu. Również wiele innych problemów szeregowania operacji i organizacji procesów produkcyjnych w swej istocie sprowadza się do problemu komiwojażera.

Najlepsze rozwiązanie tego problemu można byłoby otrzymać poprzez analizę całej przestrzeni rozwiązań. W zadaniu o n położeniach istnieje $n!$ permutacji. W tym jednak jest 2^n różnych przypadków o tej samej długości drogi sumarycznej, różniących się tylko punktem i kierunkiem startu.

* Wyższa Szkoła Inżynierska w Koszalinie, ul. Raławicka 15/17, 75-620 Koszalin

Tak więc mamy $n!/2n$, czyli $(n-1)!/2$ różnych długości drogi sumarycznej. W tabelicy 1 zestawiono liczby możliwych rozwiązań dla różnych liczb n .

Tablica 1. Liczby różnych rozwiązań dla określonych n

n	$(n-1)!/2$	n	$(n-1)!/2$
5	1.2e+001	10	1.8144e+005
15	4.35891456e+010	20	6.082255020442e+016
25	3.102242008666e+023	30	4.42088099687e+030
35	1.476163995198e+038	40	1.01989410406e+046
45	1.329135787394e+054	50	3.041409320171e+062

Dla $n=50$, zastosowanie komputerów pozwalających uzyskać ponad $10e+007$ rozwiązań na sekundę, pozwoliłoby uzyskać wynik nie wcześniej niż po upływie ok. $10e+048$ lat. Taki sposób postępowania nie jest zatem możliwy nawet dla $n>20$ (czas oczekiwania ok. 198 lat), podczas gdy wiele zadań dotyczy właśnie $n>20$.

W niniejszej pracy autorzy przedstawiają sposoby modyfikacji uczenia z rywalizacją (uczenia konkurencyjnego) Kohonena, zapewniające wyższą efektywność procesu uczenia sztucznej sieci neuronowej.

2. Algorytm uczenia konkurencyjnego

W regule uczenia Kohonena [6], [7], [8], [9] neurony w sieci są uporządkowane i uczeniu podlega nie tylko neuron wygrywający w konkurencji z innymi, ale także neurony, które z nim sąsiadują [10]. Dzięki temu algorytm naśladuje naturalny mechanizm tworzenia topograficznych map w korze wzrokowej [6], co w praktycznych zastosowaniach oznacza zmniejszenie wrażliwości na nieprawidłowe działanie pojedynczej jednostki ukrytej.

Zastosowanie sieci neuronowych z algorytmem uczenia konkurencyjnego Kohonena do rozwiązania problemu komiwojażera polega na tym [11], [12], że w układzie sprowadzonym do jednej płaszczyzny neurony reprezentują współrzędne i tworzą zamknięty łańcuch. Każdy neuron jest połączony z dwoma sąsiednimi. Łańcuch ten tworzy drogę, sumarycznych przemieszczeń. Początkowo droga ta nie przechodzi przez punkty stanowiące zbiór obiektów.

W typowym postępowaniu, na początku procesu, po losowej inicjalizacji współrzędnych neuronów, suma odległości neuronów (sumaryczna droga) jest długa i odległa od punktów, będących reprezentacją obiektów. Dla poprawy rozwiązania stosuje się najczęściej następujący algorytm:

- ▶ #1 wybiera się losowo obiekt (punkt) $Loc(NumCentrAdapt)(LocX, LocY)$, który w danym kroku staje się centrum adaptacji,
- ▶ #2 poszukuje się neuronu $Neur(NearCentrAdapt)(NeX, NeY)$ najbliższego wybranemu obiektowi $Loc(NumCentrAdapt)(LocX, LocY)$,
- ▶ #3 znaleziony neuron $Neur(NearCentrAdapt)(NeX, NeY)$ wraz z jego bliskimi i dalszymi sąsiadami z obu stron $Neur(NearCentrAdapt+i)(NeX, NeY)$, $Neur(NearCentrAdapt-i)(NeX, NeY)$ jest przemieszczany w kierunku chwilowego centrum adaptacji, przy czym wielkość względnego przesunięcia określonego neuronu $dx(i)$, $dy(i)$ jest:
 - ♦ tym większa, im mniejsza jest odległość neuronu od centrum adaptacji ($|LocX-NeX|$, $|LocY-NeY|$),
 - ♦ tym mniejsza, im dalszym (i) sąsiadem neuronu wygrywającego jest dany neuron.

Po tym powtarza się postępowanie od #1 zmieniając w kolejnych cyklach współczynniki we wzorach opisujących przemieszczenia neuronów.

W tym algorytmie zmiana położenia neuronów może być opisana następująco:

```

void SetXYNeur( int NearNeur, int NumCentrAdapt,
                float LeConst1, float LeConst2)
{
    // wskazanie na pierwszą w tablicy strukturę współrzędnych neuronów
    struct NEUR far *ne = pne;
    // wskazanie na strukturę obiektów (miejsc) ze współrzędnymi centrum adaptacji
    struct LOC far *loC = plo + NumCentrAdapt;
    int NewNrNeur; // nowy numer neuronu licząc od NearNeur
    int HalfNr = NumNeur / 2; // połowa liczby neuronów
    float h; // współczynnik siły przyciągania
    double PowAdapt; // siła przyciągania malejąca dla dalszych sąsiadów
    float dx, dy; // odległości neuronu od centrum adaptacji

    for(int i=0; i<NumNeur; ne++, ++i) // dla wszystkich neuronów
    {
        // numer liczony od neuronu najbliższego centrum adaptacji
        if((NewNumNeur = NearNeur - i) < 0)
            NewNumNeur = -NewNumNeur;
        if(NewNumNeur > HalfNr)
            NewNumNeur = NumNeur - NewNumNeur;
        h = (NewNumNeur / (float)HalfNr) * LeConst2;
        PowAdapt = exp((double) - h*h) * LeConst1;
        dx = loC->x - ne->x; // odległość od centrum adaptacji
        dy = loC->y - ne->y;
        ne->x += dx * PowAdapt; // nowe współrzędne neuronów
        ne->y += dy * PowAdapt;
    }
}

```


Współczynniki *LeConst1* i *LeConst2* przyjmowane są najczęściej w postaci wyrażeń liniowych:

$$\begin{aligned} \text{LeConst1(NumStep)} &= \\ &\text{Learn1Beg} + (\text{double})\text{NumStep} * (\text{Learn1End} - \text{Learn1Beg}) / \text{MaxStep} \\ \\ \text{LeConst2(NumStep)} &= \\ &\text{Learn2Beg} + (\text{double})\text{NumStep} * (\text{Learn2End} - \text{Learn2Beg}) / \text{MaxStep} \end{aligned}$$

w których *NumStep* oznacza numer bieżącego kroku, *MaxStep* założoną maksymalną liczbę kroków, a *Learn1Beg*, *Learn1End*, *Learn2Beg* i *Learn2End* są współczynnikami dobieranymi dla określonych realizacji.

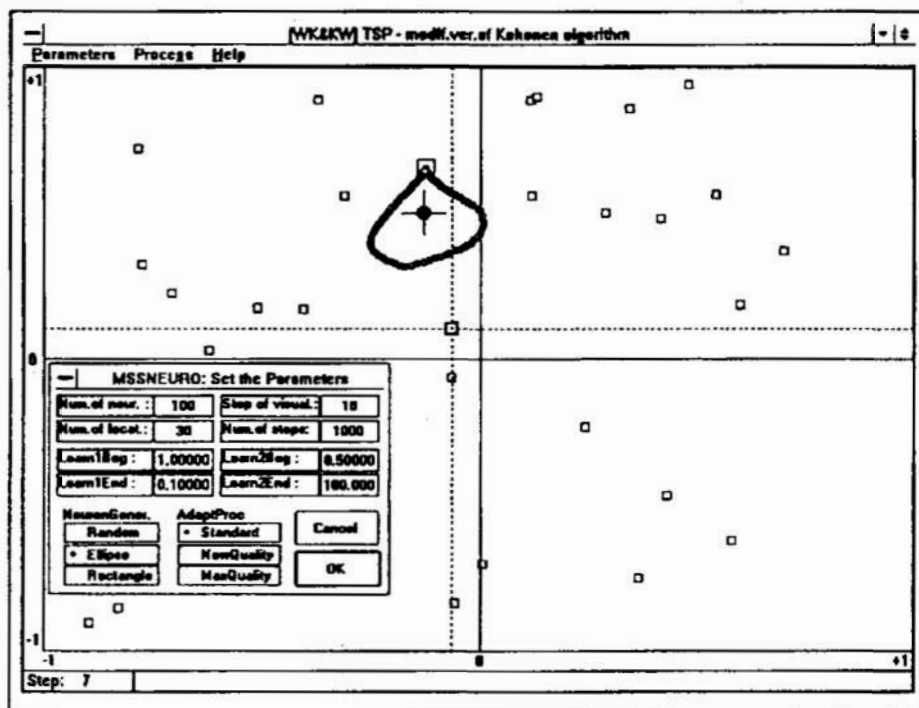
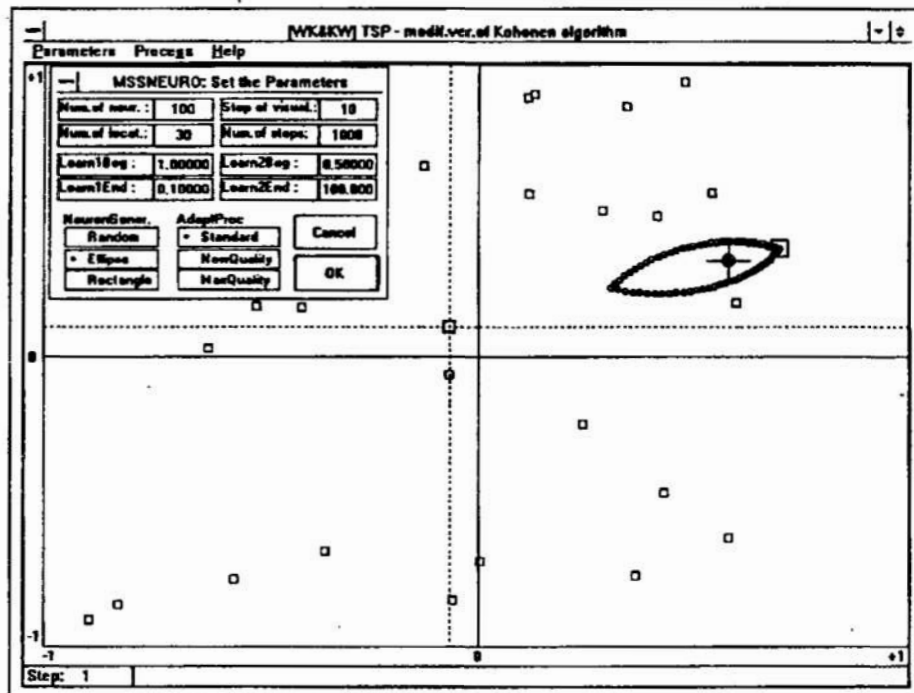
Testy przeprowadzone przez autorów wykazały liczne wady opisanego powyżej sposobu adaptacji. Współczynnik *Learn1(Beg, End)* może być interpretowany jako moc adaptacji. Jego wartości powinny zmieniać się od początkowej 1,0 do wartości mniejszej od 1,0 np. do 0,1 na końcu procesu. Dobór wartości tego współczynnika nie sprawia trudności.

Współczynnik *Learn2(Beg, End)* ma wielki wpływ na efektywność procesu. Zbyt mała wartość *Learn2Beg* oznacza na początku procesu dużą siłę przyciągania również neuronów odległych od centrum adaptacji. Wtedy cały łańcuch neuronów niepotrzebnie przemieszcza się w pobliże centrum adaptacji.

Przykładowe wyniki dla *Learn1Beg=1,0* i *Learn1End=0,1* oraz *Learn2Beg=0,5* i *Learn2End=100*, przedstawiono na rysunku 1. Na tym i dalszych rysunkach małe kwadraty oznaczają położenie obiektów (punktów) a okręgi położenie neuronów. Pogrubionym kwadratem zaznaczono środek ciężkości obiektów (punktów-miejsc), okręgami współśrodkowymi środek ciężkości łańcucha neuronów a większym kwadratem wyróżniono centrum adaptacji.

Zbyt duża wartość *Learn2Beg* oznacza od początku adaptacji małą siłę przyciągania neuronów odległych od centrum adaptacji. Wtedy proces staje się nieefektywny - łańcuch neuronów nie zdąża do miejsc adaptacji. Neurony położone blisko centrum adaptacji są w danym kroku silnie przemieszczane, lecz pozostałe neurony przemieszczają się za mało. Dla danych *Learn1Beg=1,0* i *Learn1End=0,1* oraz *Learn2Beg=25,0* i *Learn2End=100,0* otrzymano wyniki przedstawione na rysunku 2. W takiej, niekorzystnej sytuacji, lepsze wyniki uzyskuje się, gdy początkowe położenia neuronów są skupione wewnątrz obszaru, w którym rozmieszczone są obiekty (punkty).

Zbyt duża wartość *Learn2End* osłabia siłę przyciągania neuronów odległych od centrum adaptacji. Wtedy przemieszcza się wyraźnie tylko neuron najbliższy centrum. Duża wartość tego współczynnika jest korzystna w przypadku, gdy początkowe położenia neuronów są skupione wewnątrz obszaru. Zbyt mała wartość *Learn2End* powoduje, że zróżnicowanie siły przyciągania jest za małe i cały łańcuch neuronów ciągle wędruje za kolejnym centrum adaptacji.

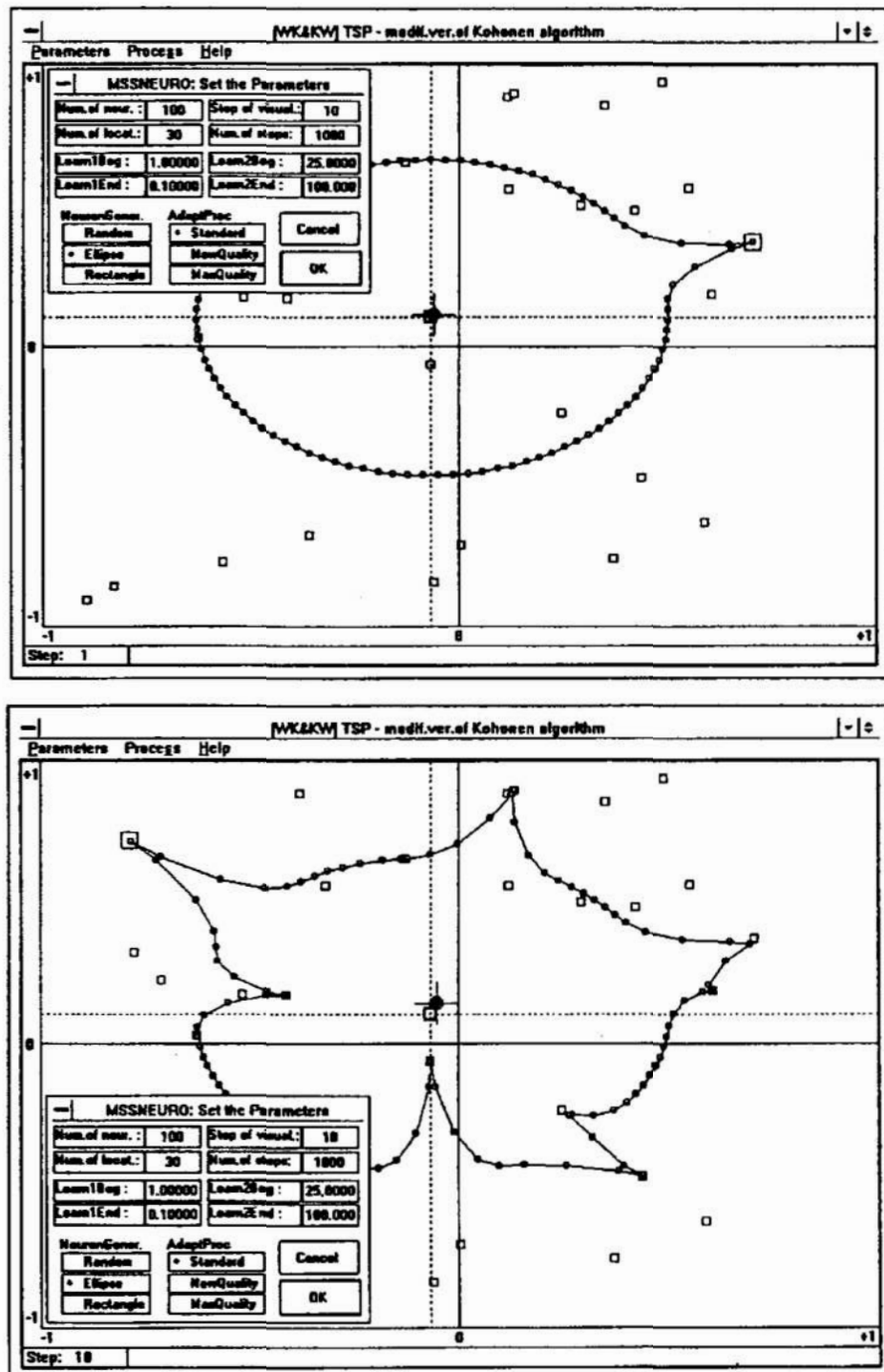


Rys.1. Przykład wpływu zbyt małej wartości $Learn2Beg=0,5$ na proces adaptacji w kroku nr 1 i 7

3. Modyfikacje zwiększające szybkość uczenia sieci oraz stabilność procesu adaptacji

Pierwszym celem opracowanych modyfikacji algorytmu adaptacji było uniezależnienie procesu od poprawności doboru przez realizatora współczynników uczenia, oraz dodatkowo, zwiększenie szybkości uczenia poprzez uzależnienie

procesu adaptacji od globalnej, a zarazem prostej oceny szybkości uczenia. Drugim celem było opracowanie innej strategii procesu uczenia, eliminującej kroki mało efektywne i zwiększającej przez to szybkość adaptacji.



Rys.2. Ilustracja wpływu zbyt dużej wartości $Learn2Beg=25,0$ na proces adaptacji w krokach nr 1 i 10

Dla uwolnienia użytkownika od doboru współczynników uczenia $Learn1Beg$, $Learn1End$, $Learn2Beg$, $Learn2End$ opracowano zależności, w których nie występują te współczynniki.

Zależności te, dające poprawne wartości w szerokim zakresie warunków charakterystycznych dla problemu komiwojażera, są następujące:

```
LeConst1 = 1.0 / log10(NumStep + 10.0),  
LeConst2 = sqrt((double)NumStep * sqrt((double)NumLoc / NumNeur)),
```

przy czym współczynnik *LeConst2* zależy od numeru kroku *NumStep* oraz stosunku liczby obiektów *NumLoc* do liczby neuronów *NumNeur*.

Dodatkowe przyspieszenia procesu uczenia się sieci uzyskano poprzez exponentialny wzrost wartości *LeConst2* w zależności od wartości względnego przesunięcia środka ciężkości łańcucha neuronów (*RelativeOffCentrNeur*) oraz stosunku średniego promienia obiektów (punktów) (*MRadLoc*) do średniego promienia łańcucha neuronów (*MRadMeur*) w poprzednim kroku:

```
LeConst2 *= exp(RelativeOffCentrNeur + (MRadLoc / MRadNeur)).
```

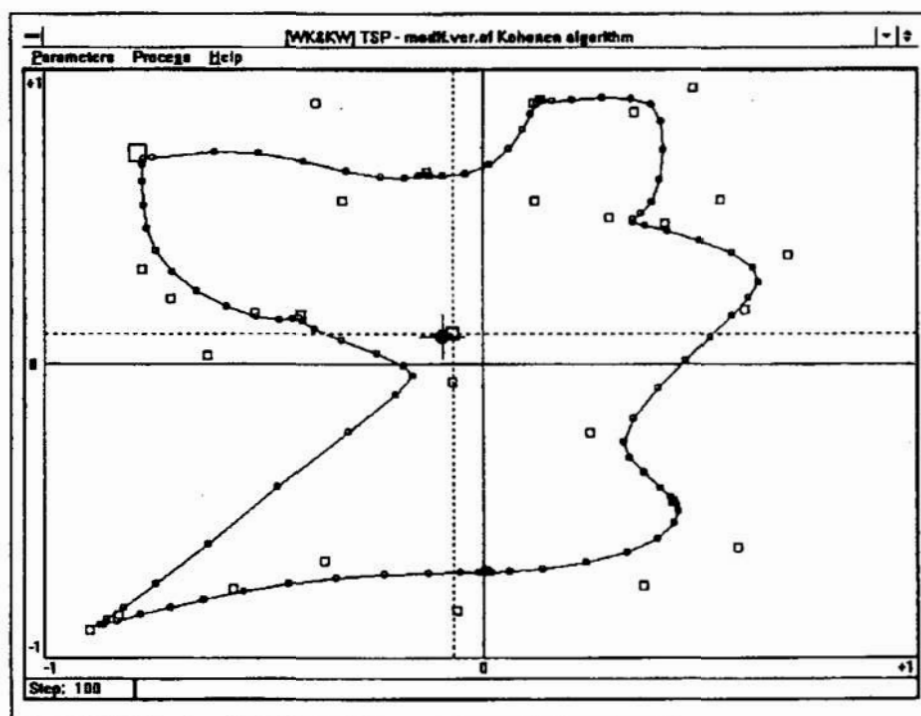
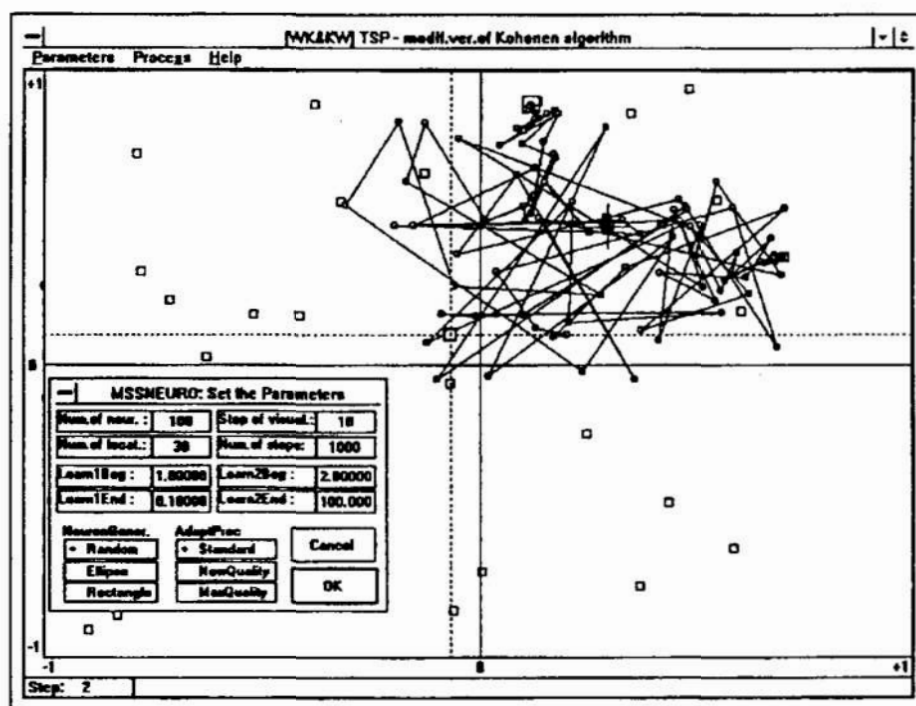
Skutek takiego postępowania jest bardzo korzystny. Współczynnik *LeConst2* zwiększa się, gdy z powodu zbyt dużej siły przyciągania cała sieć neuronów nadmiernie przemieszcza się za kolejnym centrum, lub gdy sieć jest zbyt mało rozwinięta.

Innym korzystnym rozwiązaniem, zwłaszcza na początku procesu uczenia jest zastąpienie losowego początkowego rozmieszczenia neuronów, rozmieszczeniem regularnym, z łańcuchem w formie eliptycznej, środkiem ciężkości w środku ciężkości obiektów i promieniu łańcucha równym średniemu promieniowi obiektów. Zapewnia to wyższą efektywność procesu adaptacji.

Rozwiązanie powyższego problemu zostało zawarte w programie MSSNEUR opracowanym przez autorów w systemie Borland C++ 4.0 dla środowiska Windows. Przykładowe porównanie procesu adaptacji dla algorytmu bez modyfikacji oraz algorytmu opracowanego przez autorów przedstawiono na rysunkach 3 i 4.

Na rysunku 3 pokazano dla kroku nr 2 i kroku 100 położenie łańcucha neuronów, będące wynikiem zastosowania typowego algorytmu uczenia konkurencyjnego Kohonena. Początkowe położenie neuronów (krok nr 0) było wynikiem losowej generacji współrzędnych, przy zastosowaniu generatora liczb pseudolosowych o rozkładzie jednostajnym.

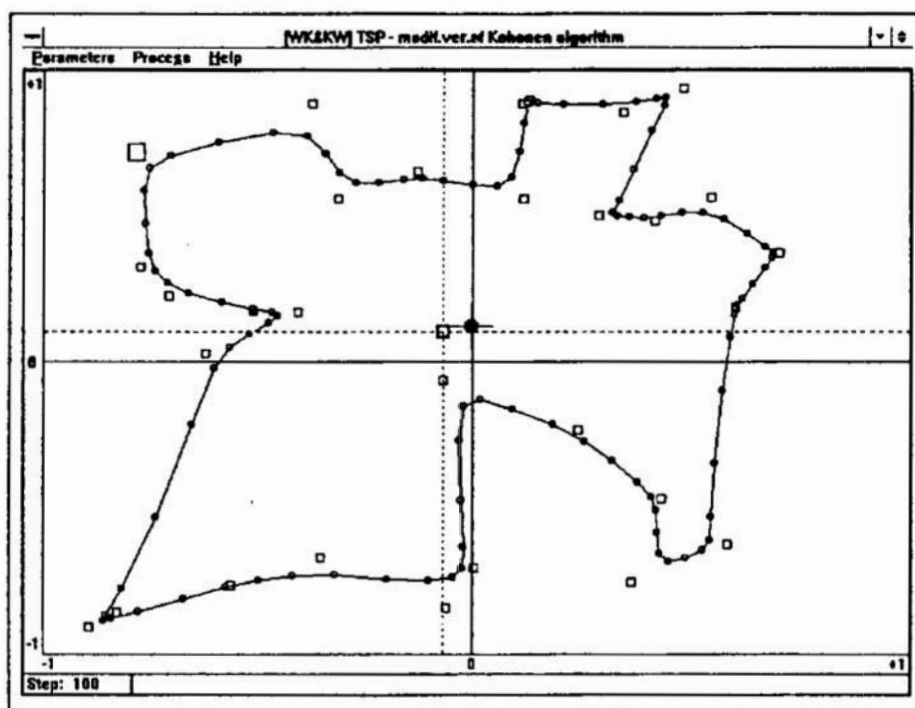
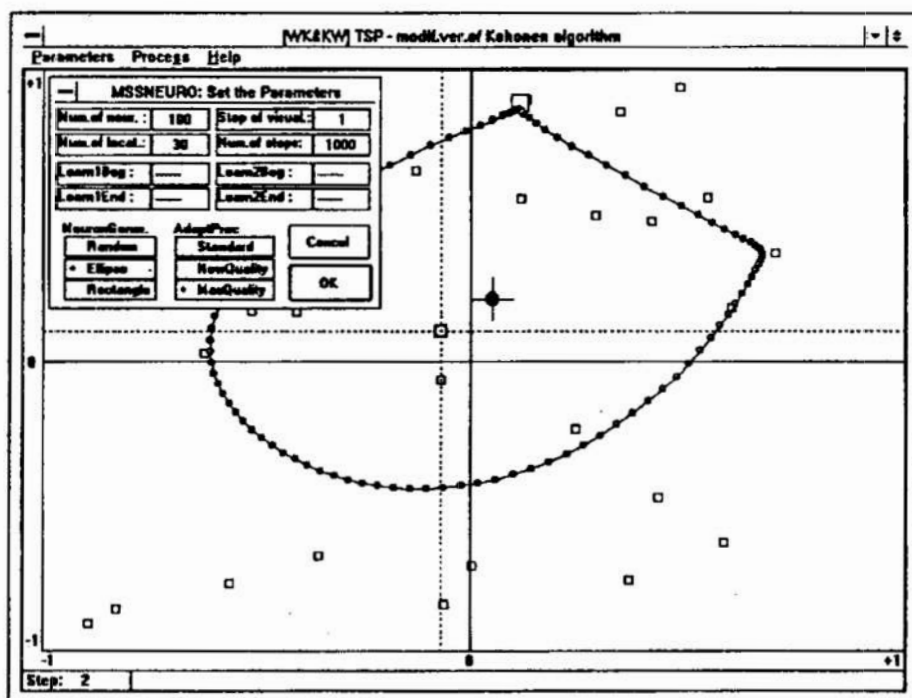
Na rysunku 4 pokazano, dla porównania przykładowe wyniki procesu z zastosowaniem opracowanych modyfikacji, zwiększających efektywność uczenia konkurencyjnego w rozpatrywanych problemach. Widoczny jest, już od pierwszych kroków, znaczny wzrost jakości rozwiązania.



Rys.3. Wyniki procesu adaptacji (uczenia konkurencyjnego sieci) dla typowego algorytmu Kohonena w krokach nr 2 i 100

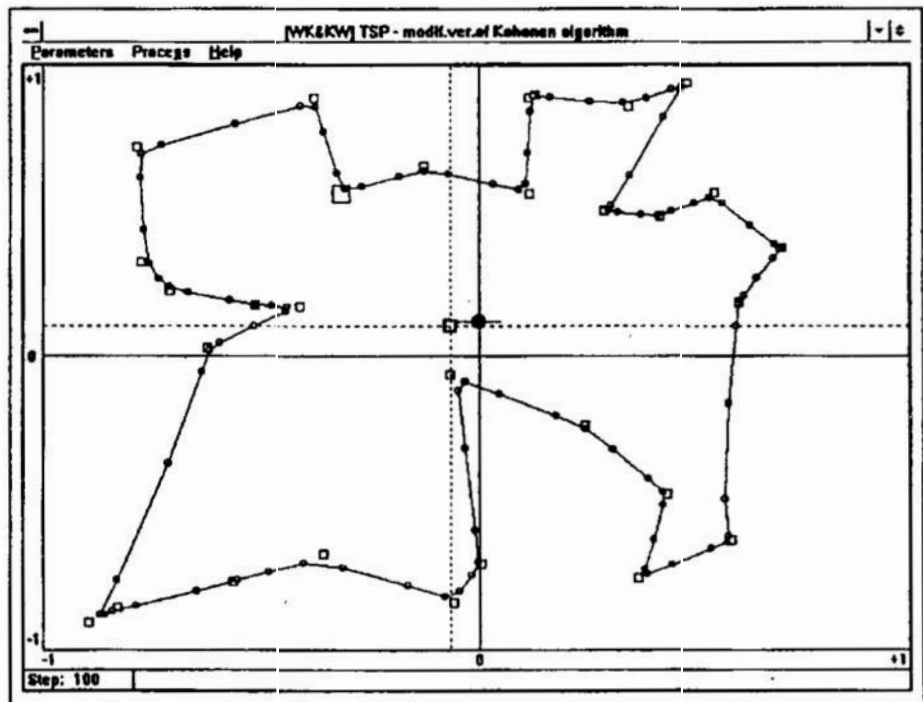
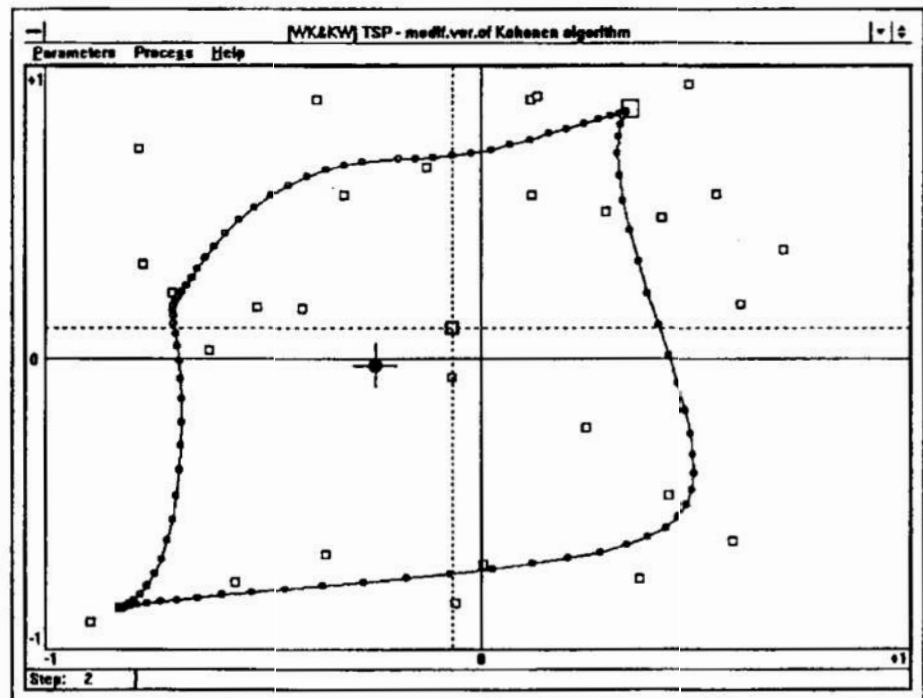
Jeszcze większą efektywność i stabilność procesu uczenia można uzyskać przez odstąpienie od zasady jednego centrum adaptacji w danym kroku i wprowadzenie dwóch lub więcej centrów. W przypadku dwóch centrów adaptacji w danym kroku losowo (lub systematycznie) wybiera się pierwsze centrum adaptacji i określa drugie, korzystnie najbardziej odległe od pierwszego. Zamiast

przemieszczać neurony do jednego centrum adaptacji, oblicza się przemieszczenie wynikowe od przyciągania obu centrów (rys.5). Dzięki temu nie następują zbędne przemieszczenia całej sieci neuronów.



Rys.4. Wyniki procesu adaptacji dla nowego, zmodyfikowanego algorytmu uczenia konkurencyjnego sieci w krokach nr 2 i 100

Podobnie można posłużyć się czterema centrami, lub jeszcze większą ich liczbą. Następne prace autorów dotyczyć będą maksymalizacji szybkości uczenia sieci w takich układach oraz wykorzystywaniu sieci neuronowych do nadzorowania pracy innych sieci.



Rys.5. Wyniki procesu adaptacji w krokach nr 2 i 100 dla zmodyfikowanego algorytmu uczenia konkurencyjnego sieci z dwoma centrami adaptacji

4. Podsumowanie

Problem komiwojażera jest charakterystycznym problemem wielu ważnych zastosowań technicznych. Potrzeba jego rozwiązania pojawia się coraz częściej, gdyż upowszechniają się elastyczne systemy wytwórcze i następuje coraz pełniejsza automatyzacja produkcji jednostkowej i małoseryjnej. Minimalizacja sumarycznej drogi przemieszczeń zespołów układu technologicznego lub pomiarowego czy też optymalizacja kolejności produkcji wielu wyrobów należą do problemów, w których rozwiązania powinny być uzyskiwane w krótkim czasie.

Znane i stosowane dotąd metody poszukiwania rozwiązań zbliżonych do optymalnego nie zapewniają wystarczającej efektywności. Zastosowanie sieci neuronowych z algorytmem uczenia konkurencyjnego Kohonena pozwala szybko uzyskiwać zadowalające rozwiązania, pod warunkiem, iż poprawnie zostaną dobrane współczynniki uczenia sieci. Jest to istotną wadą typowego algorytmu, szczególnie uciążliwą dla użytkowników o mniejszym doświadczeniu.

Opracowany przez autorów algorytm uczenia sieci nie zawiera stałych współczynników dobieranych przez użytkownika lecz zmienne współczynniki uczenia modyfikowane samoczynnie na podstawie analizy procesu adaptacji. Zapewnia to wyższą efektywność uczenia już od pierwszych kroków. Jeszcze korzystniejsze wyniki można uzyskać stosując opracowane algorytmy uczenia z wieloma centrami adaptacji w każdym kroku.

5. Literatura

1. *Sysło M.M., Deo N., Kowalik J.S.*: Algorytmy optymalizacji dyskretnej. Wydawnictwo Naukowe PWN, Warszawa 1993.
2. *Garey M.R., Johnson D.S.*: Computer and Intractability- a Guide to the Theory of NP-completeness, W.H. Freeman, San Francisco 1979.
3. *Lawler E.L., Lenstra J.K., Rinnooy Kan A.H.G., Shmoys.*: The Travelling Salesman Problem, Wiley, New York 1985.
4. *Durbin R., Willshaw D.*: An Analogue Approach to the Travelling Salesman Problem Using Elastic Net Method, *Nature*, 326, 1987, pp. 689-691.
5. *Angeniol B., de la Croix Vaubois G. and le Texier J.Y.*: Self-Organizing Feature Maps and the Travelling Salesman Problem, *Neural Networks*, 1, 1988, pp. 289-293.
6. *Kohonen T.*: Self-Organized Formation of Topologically Correct Feature Maps", *Bio. Cybern.*, 43, 1982, pp. 59-69.
7. *Kohonen T.*: Selforganization and associative memory. Berlin, Springer Verlag 1984.

8. *Kohonen T.*: The Self-Organizing Map, Proceedings of the IEEE, special issue on Neural Networks, vol.78, Nr. 9, 1990, pp.1464-1480.
9. *Kohonen T.*: Unsupervised learning algorithms. Proceedings of International Conference "Neural Networks-Biological Computers or Electronic Brains", Lyon, 1990.
10. *Tadeusiewicz R.*: Sieci neuronowe. Akademicka Oficyna Wydawnicza. Warszawa 1993.
11. *Kacalak W.*: Zastosowania sieci neuronowych w budowie i eksploatacji maszyn. Mechanik 2/1994, s.45-49.
12. *Kacalak W., Wawryn K.*: Algorytm i program MSSNEUR do minimalizacji drogi przemieszczeń zespołów wykonawczych w układach technologicznych, 1994 (praca nie publikowana).

**MODIFICATIONS INCREASING EFFECTIVENESS
OF THE COMPETITIVE SELF LEARNING
NEURAL NETWORK ALGORITHMS
IN CHOSEN TECHNICAL APPLICATIONS**

Summary

Recently, neural networks applications have been receiving substantial consideration for the solution of the machine construction and exploitation problems. The Travelling Salesman Problem (TSP) is a combinatorial problem and makes possible several interesting technical applications, especially in control systems and optimisation of technological processes. Unfortunately, a global efficient analytically tractable solution of the TSP does not exist. Alternative approaches relying upon artificial neural network methodologies enable to obtain almost optimal solutions. Proposed in the paper, modifications in competitive self learning neural network TSP algorithm increase search effectiveness significantly.

Neuronowe modele obiektów dynamicznych

Artykuł przedstawia zastosowanie Sztucznych Sieci Neuronowych w procesie modelowania nieliniowych obiektów dynamicznych. Dwa przykłady prezentują zastosowanie do tego celu częściowo rekurencyjnej zmodyfikowanej sieci Jordana oraz klasycznej statycznej sieci jednokierunkowej. W zakończeniu przedstawiono niektóre cechy obu typów sieci w odniesieniu do modelowania obiektów dynamicznych.

1. Wprowadzenie

Literatura prezentuje trzy główne nurty zastosowań modeli obiektów w nauce i w technice [1]:

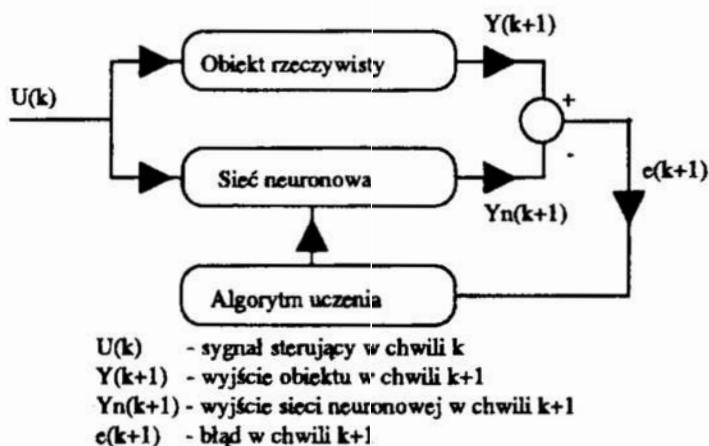
- badania naukowe,
- projektowanie urządzeń technicznych,
- sterowanie.

W badaniach naukowych modele wykorzystuje się do zdobycia nowych informacji o interesujących nas obiektach.

W projektowaniu urządzeń technicznych modele umożliwiają wariantowanie konstrukcji oraz ich wielokryterialną optymalizację.

W automatyce dają one możliwość doboru sterowań dla obiektu (szczególnie w sterowaniu adaptacyjnym i optymalnym), określenia zachowania obiektu na granicach stabilności i inne. Ostatnio wykorzystywane są w procesie uczenia regulatorów neuronowych w przypadku procesów nieliniowych i strukturalnie niestabilnych lub o nieznanym modelu matematycznym.

Modelowanie obiektów dynamicznych, szczególnie nieliniowych, jest trudne, pracochłonne i nie zawsze daje zadowalające wyniki. W ostatnich latach pojawiła się nowa metoda modelowania takich obiektów polegająca na wykorzystaniu Sztucznych Sieci Neuronowych (ANN-Artificial Neural Network) [2,3,4].



Rys. 1 Struktura układu modelującego

* Politechnika Gdańska Wydział Mechaniczny
Zakład Konstrukcji i Eksploatacji Maszyn
80-952 Gdańsk Narutowicza 11/12
e-mail sgrzymek@sunrice.pg.gda.pl

Do modelowania wykorzystywane są zarówno klasyczne jednokierunkowe sieci neuronowe (feedforward networks) jak i sieci rekurencyjne (recurrent networks).

Klasyczne sieci neuronowe jednokierunkowe są układami dynamicznymi zerowego rzędu. Symulację dynamiki obiektu wyższego rzędu osiąga się w nich przez wprowadzenie na wejście sieci odpowiedniej liczby poprzednich stanów obiektu i sygnałów sterujących działających w przeszłości na obiekt.

Rekurencyjne sieci neuronowe są układami dynamicznymi. Charakteryzują się posiadaniem połączeń (sprzężeń) zwrotnych pomiędzy warstwami sieci. Umożliwia im to zapamiętanie stanu obiektu i historii oddziaływań na obiekt sygnałów sterujących.

Modelowanie z wykorzystaniem sztucznych sieci neuronowych (uczenie sieci) prowadzi się w czasie rzeczywistym w układzie jak na rys. 1.

W dalszej części artykułu przedstawiono zastosowanie częściowo rekurencyjnej zmodyfikowanej sieci Jordana do zamodelowania nieliniowego obiektu reaktora chemicznego oraz zastosowanie klasycznej sieci jednokierunkowej do zamodelowania zbiornika ze swobodnym odpływem, w celu porównania możliwości i parametrów uczenia oraz dokładności modelu.

2. Neuronowy model reaktora chemicznego

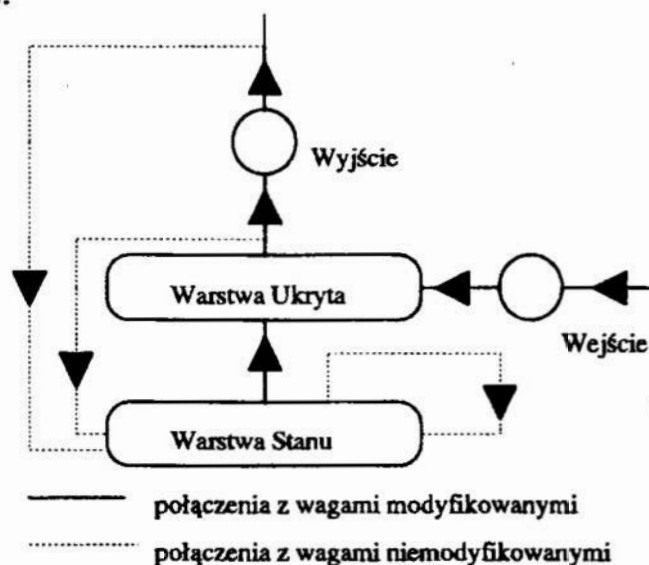
Niech przykładowy badany reaktor chemiczny będzie reprezentowany przez następujące dyskretne równanie 'wejścia - wyjścia':

$$y(k+1) = \frac{y(k)}{1.5 + y^2(k)} - 0.3y(k-1) + 0.5\sin(u(k)) \quad (1)$$

gdzie k jest numerem dyskretnego kroku czasu.

Do zamodelowania tego obiektu zastosowano sztuczną częściowo rekurencyjną zmodyfikowaną sieć Jordana (rysunek 2). Sieć ta jest układem typu SISO (Single Input-Single Output), o jednym wejściu i jednym wyjściu. Strukturę wewnętrzną sieci tworzą dwie warstwy o jednakowej liczbie neuronów: warstwa ukryta (hidden layer) i warstwa stanu (state layer, context layer). W sieci występują dwa rodzaje połączeń:

- z wagami modyfikowanymi w procesie uczenia,
- sprzężenia zwrotne z wagami niemodyfikowanymi (stąd nazwa 'częściowo rekurencyjna').



Rys. 2 Zmodyfikowana sieć Jordana

Działanie zmodyfikowanej sieci Jordana dla neuronów z liniowymi funkcjami przejścia można opisać następującym układem równań:

$$X_S(k) = c_O \cdot N \cdot y(k) + c_S \cdot M \cdot X_S(k-1) + c_H \cdot M \cdot X_H(k-1) \quad (2)$$

$$X_H(k+1) = W_H \cdot X_S + W_I \cdot u(k) \quad (3)$$

$$y(k+1) = W_O \cdot X_H(k+1) \quad (4)$$

gdzie:

- k - kolejny krok,
- u - sygnał wejściowy,
- X_S - wektor wyjścia z warstwy stanu,
- X_H - wektor wyjścia z warstwy ukrytej,
- y - sygnał wyjściowy,
- W_I - wektor wag pomiędzy wejściem a warstwą ukrytą,
- W_H - macierz wag pomiędzy warstwami: stanu i ukrytą,
- W_O - wektor wag pomiędzy warstwą ukrytą a wyjściem,
- c_O - współczynnik sprzężenia zwrotnego,
- c_H - współczynnik sprzężenia pomiędzy warstwami: ukrytą i stanu.
- c_S - współczynnik autoasocjacji warstwy stanu,
- M - macierz kwadratowa jedynek,
- N - wektor jedynek.

Analizując powyższe równania widać ich podobieństwo do opisu obiektu dynamicznego w przestrzeni stanów.

W procesie modelowania obiektu konieczne jest podjęcie decyzji co do struktury sieci i parametrów uczenia. Przeprowadzone eksperymenty wskazują, że dobre rezultaty uzyskuje się stosując następujące zasady:

- liczba neuronów w warstwie stanu i ukrytej od 6 do 15, jednak nie mniej niż rząd dynamiczny obiektu;
- funkcja przejścia neuronów warstwy ukrytej liniowa dla obiektów liniowych i nieliniowa (sigmoid, tangens hiperboliczny) dla nieliniowych;
- współczynnik autoasocjacji $c_S = (0.75.. 0.9)$;
- współczynnik $c_H = 1$;
- współczynnik $c_O = 1$;
- współczynnik uczenia (learning rate) od 0.001 do 0.1;
- współczynnik stałego kierunku modyfikacji (momentum term) od 0.01 do 0.1;
- sygnał sterujący uczący - sygnał losowy o równomiernym rozkładzie prawdopodobieństwa.

Jako algorytmu uczenia używa się Wstecznej Propagacji Błędów (BP-Back error Propagation) [5] zwanej też Uogólnioną Regułą Delta (Generalized Delta Rule).

Dla modelowanego reaktora chemicznego opisanego równaniem (1) przyjęto następujące parametry: liczba neuronów - 10, funkcja przejścia - tangens hiperboliczny, $c_S = 0.8$, learning rate - 0.005, momentum term - 0.01.

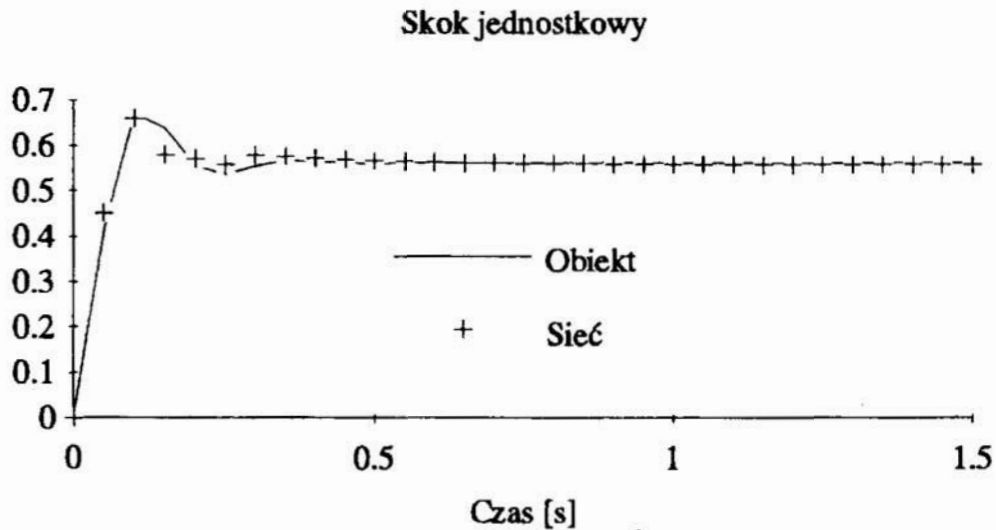
Zaprojektowano sieć neuronową, następnie utworzono jej symulator komputerowy. Ten symulator został napisany w języku Borland Pascal firmy Borland International Inc.

Proces modelowania (uczenia sieci) przeprowadzono wykorzystując sygnał losowy o wartości z zakresu od -1 do 1 i zakończono go po 700.000 kroków. Jako kryterium

oceny modelu zastosowano średni znormalizowany błąd kwadratowy wyrażony wzorem:

$$E = \frac{\sum_0^{399} (y(k) - y_n(k))^2}{\sum_0^{399} (y(k))^2} \quad (5)$$

Na rys. 3 zaprezentowano odpowiedź modelu neuronowego reaktora chemicznego na tle odpowiedzi obiektu obliczonej ze wzoru (1) dla wymuszenia w postaci skoku jednostkowego. W tym przypadku wartość błędu liczonego według wzoru (5) wynosi 0.00014.

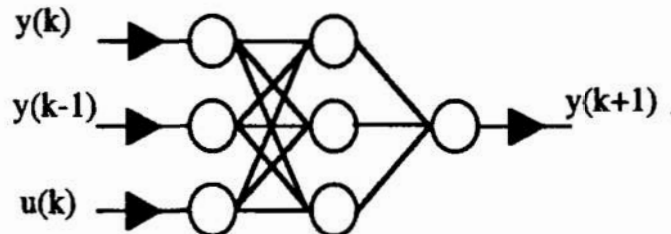


3. Neuronowy model zbiornika ze swobodnym odpływem

Proces napełniania zbiornika ze swobodnym odpływem jest reprezentowany przez następujące dyskretne równanie 'wejścia - wyjścia':

$$y(k+1) = y(k) - 0.1\sqrt{y(k)} + 0.2u(k) \quad (6)$$

Do zamodelowania tego obiektu zastosowano sztuczną klasyczną jednokierunkową sieć neuronową (rys. 4). Na wejścia sieci wprowadzono informacje o poprzednich stanach obiektu - wartość poziomu cieczy w zbiorniku w chwili obecnej i poprzedniej - oraz wartość sygnału sterującego (stopnia otwarcia zaworu na dopływie) w chwili obecnej. Spodziewaną odpowiedzią sieci jest wartość poziomu cieczy w zbiorniku w chwili następnej.



Rys. 4 Sieć jednokierunkowa

Działanie sieci z jedną warstwą ukrytą i neuronami z liniowymi funkcjami opisują następujące równania:

$$X = W_H \cdot I \quad (7)$$

$$y = W_O \cdot X \quad (8)$$

gdzie: X - wektor sygnałów wyjściowych warstwy ukrytej,
 y - sygnał wyjściowy sieci (w tym przypadku skalar),
 W_H - wektor wag pomiędzy wejściem a warstwą ukrytą,
 W_O - wektor wag pomiędzy warstwą ukrytą a wyjściem,

Przeprowadzone eksperymenty wskazują, że podczas projektowania sieci należy wziąć pod uwagę następujące wytyczne:

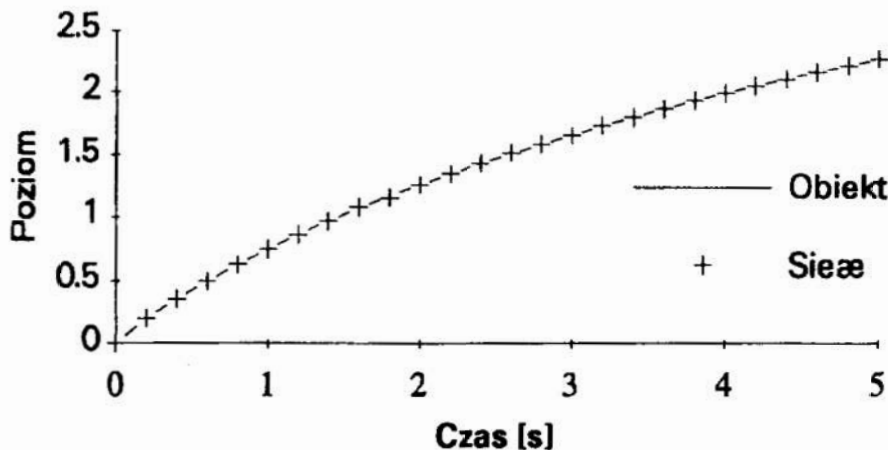
- liczba wejść nie mniejsza niż rząd obiektu;
- liczba warstw ukrytych - 1;
- liczba neuronów w warstwie ukrytej nie mniejsza niż liczba wejść;
- współczynnik uczenia (learning rate) od 0.005 do 0.5;
- współczynnik stałego kierunku modyfikacji (momentum term) od 0.01 do 0.5;
- sygnał sterujący uczący - sygnał losowy o równomiernym rozkładzie prawdopodobieństwa.

Jako algorytmu uczenia używa się Wstecznej Propagacji Błędów.

Dla neuronowego modelu obiektu zbiornika ze swobodnym odpływem opisanego równaniem (6) przyjęto następujące parametry: liczba wejść - 3, liczba neuronów w warstwie ukrytej - 3, learning rate - 0.2, momentum term - 0.1.

Symulator zaprojektowanej sieci neuronowej został napisany w języku Borland Pascal. Dla tego obiektu proces modelowania prowadzono wykorzystując losowy sygnał sterujący z zakresu od 0 (zawór dopływowy zamknięty) do 1 (zawór w pełni otwarty) i zakończono po 100.000 kroków.

Na rys. 5 zaprezentowano odpowiedź modelu neuronowego zbiornika ze swobodnym odpływem na tle odpowiedzi obiektu wyliczonej z równania (6) dla wymuszenia w postaci skoku jednostkowego (pełne otwarcie zaworu). W tym przypadku wartość błędu liczonego według wzoru (5) wynosi 0.00008.



4. Podsumowanie

Porównując sztuczne sieci neuronowe: rekurencyjną i jednokierunkową w odniesieniu do modelowania obiektów dynamicznych można poczynić następujące spostrzeżenia:

1. sieć jednokierunkowa w sposób statyczny odwzorowuje równanie 'wejścia - wyjścia' modelowanego obiektu;
2. sieć rekurencyjna odwzorowuje dynamiczny opis obiektu w przestrzeni stanów;
3. podczas modelowania całkowicie nieznanego obiektu o wiele łatwiej określić poprawną strukturę dla sieci rekurencyjnej niż dla sieci jednokierunkowej;
4. w sieci rekurencyjnej istnieją problemy z ustawieniem stanu początkowego obiektu (z wyjątkiem stanu zerowego); problem ten nie występuje przy sieci jednokierunkowej;
5. sieć rekurencyjna pamięta całą historię dynamiki obiektu; model obiektu w postaci sieci jednokierunkowej uwzględnia tylko pewien określony okres przeszłości, zależny od ilości wejść do sieci;
6. modele z siecią rekurencyjną wykazują lepszą dokładność odwzorowania niż modele z siecią jednokierunkową; prawdopodobną przyczyną jest własność wymieniona w punkcie 5;
7. uczenie sieci rekurencyjnej trwa dłużej (dla tego samego obiektu), gdyż z reguły stosuje się mniejszy współczynnik uczenia niż dla sieci jednokierunkowej

Dotychczas uzyskane rezultaty wskazują, że użycie sieci neuronowych do modelowania obiektów dynamicznych daje bardziej zadawalające rezultaty niż stosowanie metod aproksymacyjnych.

Wydaje mi się, że modele neuronowe oparte na sieciach jednokierunkowych, jako uzyskiwane przy mniejszym nakładzie czasowym i z zadawalającym z inżynierskiego punktu widzenia dokładnością, znajdą szerokie zastosowanie w technice. Modele oparte na sieciach rekurencyjnych, jako czasochłonne, ale o wysokiej dokładności znajdą zastosowanie w badaniach naukowych oraz w technice dla bardzo trudnych obiektów.

Bibliografia

- [1] Giergiel J., Uhl T., "Identyfikacja układów mechanicznych", PWN Warszawa 1990.
- [2] Thibault J., Grandjean B.P.A., "Process Control Using Feedforward Neural Networks", Journal of Systems Engineering vol.2, London 1992;
- [3] Yabuta T., Yamada T., "Learning Control Using Neural Networks", Journal of Systems Engineering vol.2, London 1992;
- [4] Pham D.T., Oh S.J., "A Recurrent Backpropagation Neural Network for Dynamic System Identification", Journal of Systems Engineering vol.2, London 1992;
- [5] Caudill M., "Neural Network Primer, Part III, Backpropagation Network", AI EXPERT, June 1988;

Neural models of dynamic plants

Modelling of nonlinear dynamic plants by means of Artificial Neural Network is presented. Two examples show how partially-recurrent modified Jordan network and classic feedforward network may be used for that purpose. Finally some features of both networks in the aspect of using them in dynamic plants modelling are discussed.

ZASTOSOWANIE METOD ANALIZY PROCESÓW STOCHASTYCZNYCH W BADANIACH OKRESOWOŚCI STRUKTURY MATERIAŁU

W referacie przedstawiono możliwości zastosowania metod analizy procesów stochastycznych w badaniach okresowości struktury materiału, przy wykorzystaniu komputerowego programu do planowania i analizy doświadczeń SPECTRA 3.1.

Podkreślono rolę analizy statystycznej struktury materiału jako istotnego elementu inżynierii materiałowej.

1. Wstęp

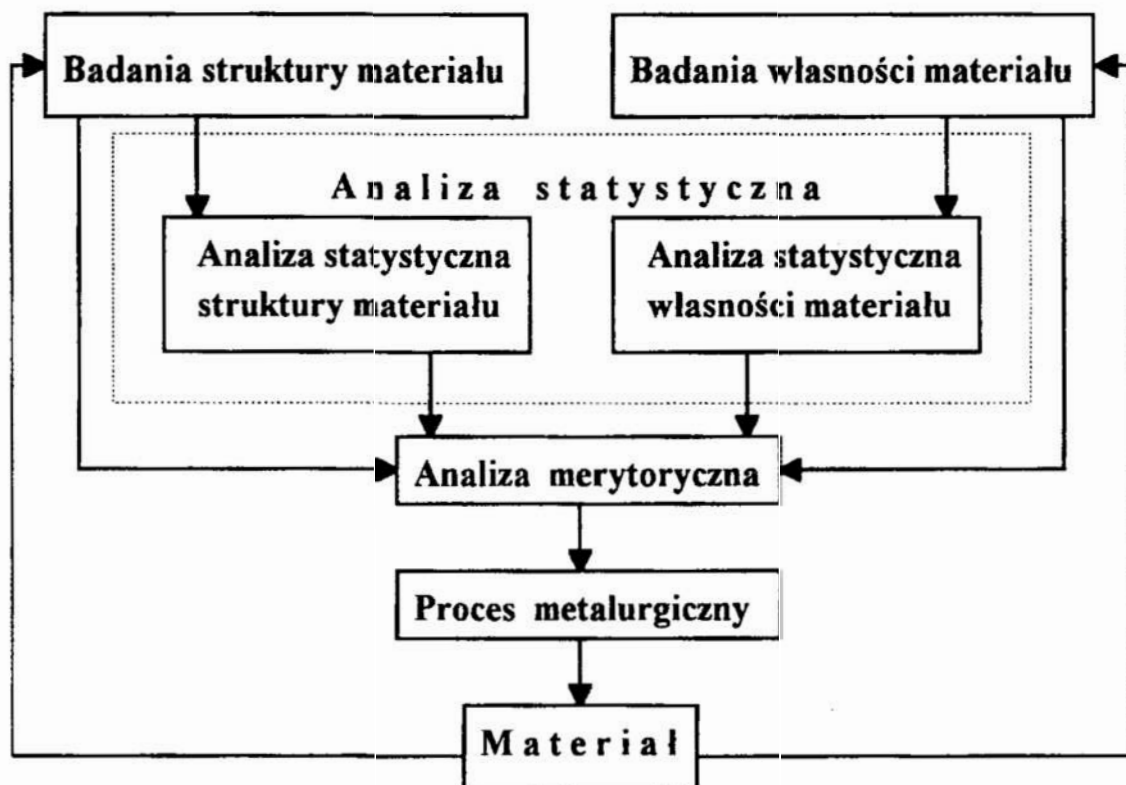
Rolą inżynierii materiałowej jest tworzenie nowych lub udoskonalonych materiałów o lepszych własnościach, przy uwzględnieniu innych kryteriów, przede wszystkim ekonomicznych oraz możliwości wytwórczych. Podstawowymi problemami tej dyscypliny są: logiczne skojarzenie przyczynowo-skutkowe wyników prowadzonych niezależnie od siebie badań własności materiału oraz jego struktury, a także sterowanie procesem metalurgicznym i technologiami pokrewnymi w celu uzyskania materiału o pożądanych parametrach.

Wyniki badań wytrzymałościowych odpowiadają na pytanie na ile materiał spełnia określone wymagania dotyczące jego własności. Natomiast badania struktury - mówią o budowie wewnętrznej materiału, rozmieszczeniu i relacjami pomiędzy jej elementami składowymi. Wspólna analiza merytoryczna tych wyników daje podstawy do opracowania nowych parametrów technologicznych procesu metalurgicznego, dającego w efekcie nowy materiał, będący następnie przedmiotem kolejnych badań. Przedstawiony na rys. 1 schemat obejmuje powyższe działania oraz zachodzące pomiędzy nimi korelacje.

Analiza statystyczna struktury materiału ma na celu określenie tych parametrów geometrycznych, które mają największy wpływ na wybrane jego własności. Są to przede wszystkim: wielkość ziarn poszczególnych składników struktury, ich kształt, rozmieszczenie i ukierunkowanie. Z tych parametrów stosunkowo najmniej rozpowszechnione w metalografii jest pojęcie okresowości rozkładu elementów struktury, która ma istotny wpływ na różne własności materiału: elektromagnetyczne, optyczne, ale przede wszystkim na wytrzymałość zmęczeniową. Jej rolę podkreśla fakt, że 90% uszkodzeń urządzeń mechanicznych spowodowanych jest przekroczeniem wytrzymałości zmęczeniowej, a tylko 10% - statystycznej. Przyjmuje się, że do $10\ 000\ \text{min}^{-1}$ [166,7 Hz] częstotliwość drgań nie ma wpływu na wytrzymałość

* Politechnika Krakowska. Instytut Materiałoznawstwa i Technologii Metali.
31-864 Kraków, Al. Jana Pawła II nr 37.

zmęczeniową. Jednak powyżej tej wartości istotny wpływ częstotliwości wskazuje na powstawanie w materiale rezonansu, a więc na istotny wpływ parametrów geometrycznych, w tym okresowości struktury.



Rys. 1. Schemat badań inżynierii materiałowej

Fig. 1. Flowchart of materials engineering

2. Identyfikacja struktury oraz analiza jej okresowości

Identyfikację struktury pod kątem analizy jej okresowości można przeprowadzić, przy wykorzystaniu aktualnych możliwości aparaturowych dwoma sposobami. Za pomocą skanowania przy kroku dyskretyzacji h , obrazu struktury w celu uzyskania ciągu binarnego (0 - brak elementu w danym punkcie obszaru, 1 - występowanie elementu): $u_i = 0, 0, 0, 1, 0, 1, 1, 0 \dots (i = 1, 2, \dots, N)$. Ciąg ten można następnie unormować tworząc wyrazy o zerowej średniej arytmetycznej: $y_i = u_i - \bar{u}$.

Innym sposobem uzyskania analogicznego ciągu jest przetworzenie zbioru współrzędnych środków ciężkości wybranych składników struktury. Odpowiednikiem kroku dyskretyzacji będzie w tym przypadku dokładność określenia współrzędnych.

Uzyskany w ten sposób materiał badawczy stanowi podstawę do obliczenia estymowanej funkcji autokorelacji (FAK):

$$R(\tau) = R(rh) = \frac{1}{n-r} \sum_{i=1}^{N-r} y_i y_{i+r}; r = 0, 1 \dots m \quad (1)$$

oraz

$$-1 \leq \rho(\tau) = r(\tau) = \frac{R(\tau)}{R(0)} \leq 1 \quad (2)$$

Unormowana funkcja autokorelacji $r(\tau)$ stanowi już podstawę do stwierdzenia okresowości. Kolejnym etapem analizy statystycznej jest wyznaczenie funkcji gęstości widmowej mocy (GWM) metodą konwencjonalną Blackmana - Turkeya z zastosowaniem funkcji wagi Hanna [1]. Funkcja GWM, otrzymana bezpośrednio z FAK, wskazuje już - w oczywisty sposób - częstotliwość odpowiadającą rozkładowi elementów struktury materiału.

3. Badania własne

Badania przeprowadzono przy użyciu programu do komputerowego wspomagania planowania i analizy doświadczeń SPECTRA 3.1, przeznaczonego do analizy widmowej sygnałów losowych. Program ten został opracowany i wykonany w Instytucie Materiałoznawstwa i Technologii Metali Politechniki Krakowskiej. Analizie poddano dwa ciągi binarne. jeden symulujący wynik identyfikacji struktury okresowej, oznaczony 01 i drugi, oznaczony 02, symulujący obraz struktury nieokresowej.

Parametry statystyczne tych ciągów były następujące:

- dla ciągu 01:

PARAMETRY STATYSTYCZNE SYGNAŁU

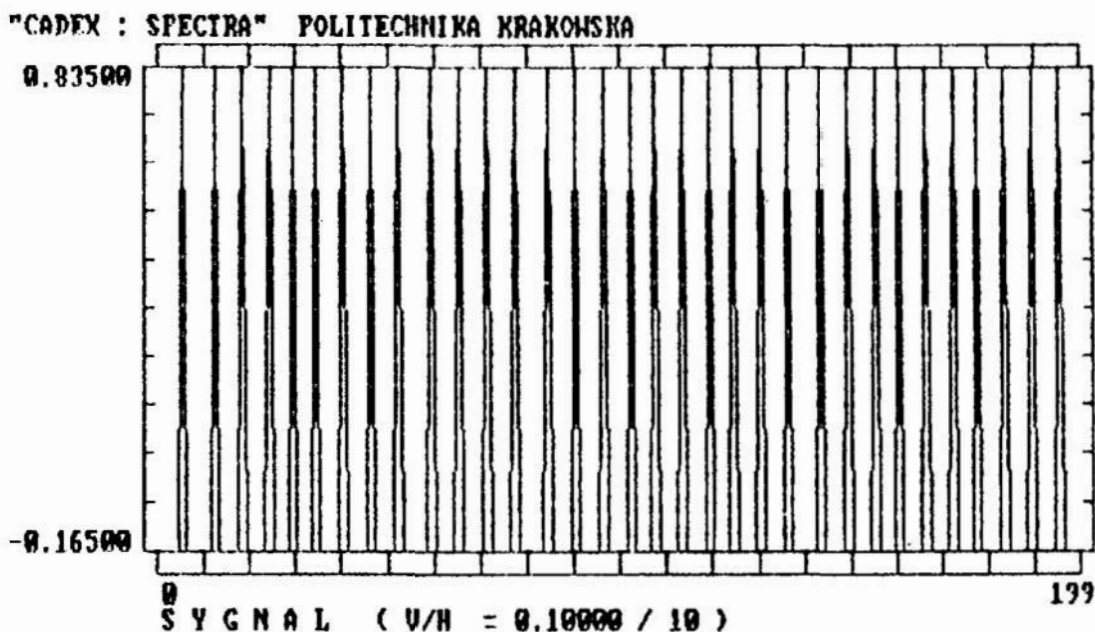
Srednia arytmetyczna sygnału (niecentrowanego) us	-0.000000
Srednia arytmetyczna sygnału (centrowanego) ys	0.0
Wariancja sygnału jako próbki	0.137775
Wariancja populacji	0.138467
Odchylenia standardowe sygnału s 0.371181 S	0.372112
Odchylenie standardowe wartosci sredniej Ss	0.026312
Wartosci ekstremalne . . . ymin -0.165000 ymax	0.835000
Rozstep w sygnale R	1.000000
Odchylenie przecietne sygnału	0.275550
Wspolczynnik skosnosci (asymetria) g1 1.805051 M3	0.092309
Wspolczynnik spłaszczenia (eksces) g2 1.258211 M4	0.080829
Sygnal centrowany	
N = 200 h = 1.000000	

- oraz dla ciągu 02:

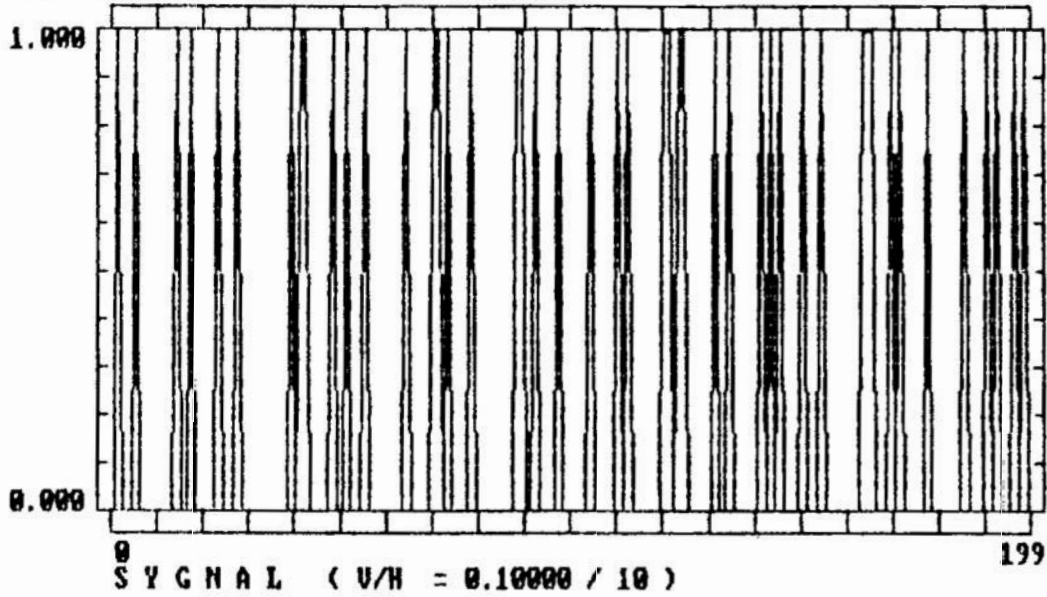
PARAMETRY STATYSTYCZNE SYGNAŁU

Srednia arytmetyczna sygnału (niecentrowanego) u_s			0.230000
Srednia arytmetyczna sygnału (centrowanego) y_s			0.0
Wariancja sygnału jako próbki			0.177100
Wariancja populacji			0.177990
Odchylenia standardowe sygnału s	0.420833	S	0.421889
Odchylenie standardowe wartości średniej S_s			0.029832
Wartości ekstremalne . . . y_{min}	-0.230000	y_{max}	0.770000
Rozstęp w sygnale		R	1.000000
Odchylenie przeciętne sygnału			0.354200
Współczynnik skosności (asymetria) g_1	1.283171	M3	0.095634
Współczynnik spłaszczenia (eksczes) g_2	-0.353473	M4	0.083007
Sygnal centrowany			
N =	200	h =	1.000000

Program SPECTRA 3.1, przeznaczony do analizy przebiegów stochastycznych daje możliwość wizualizacji badanego sygnału w postaci wykresu $u_i = f(i)$. Dla ciągów binarnych nie mają one istotnej wartości, niemniej w prezentowanych badaniach przekrój wykresu funkcją $u = 1$ daje wprost obraz rozmieszczenia badanych elementów struktury. Dlatego też na rys. 2 i 3 przedstawiono wykresy badanych sygnałów.

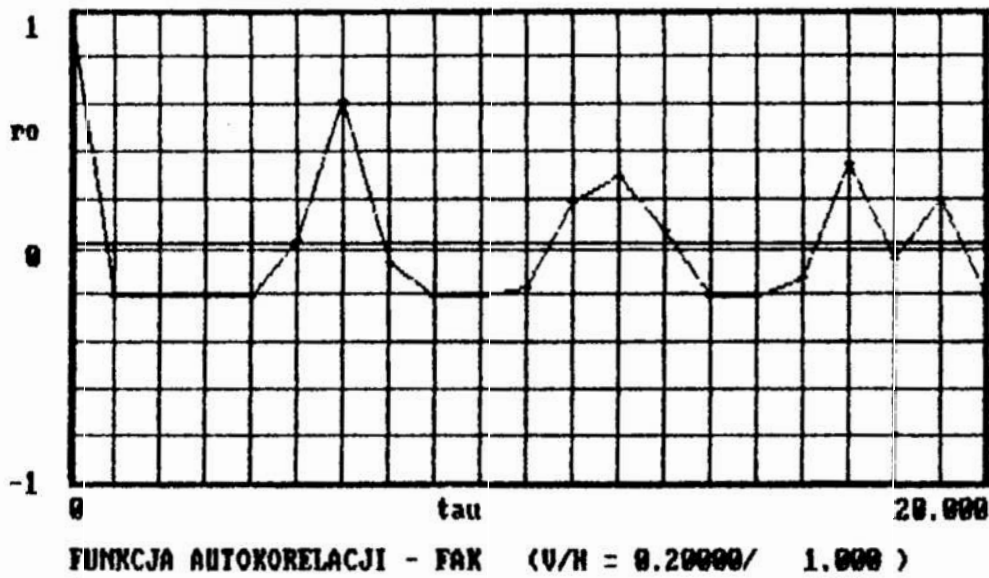


Rys. 2. Sygnal struktury okresowej
Fig. 2. Signal of periodic structure

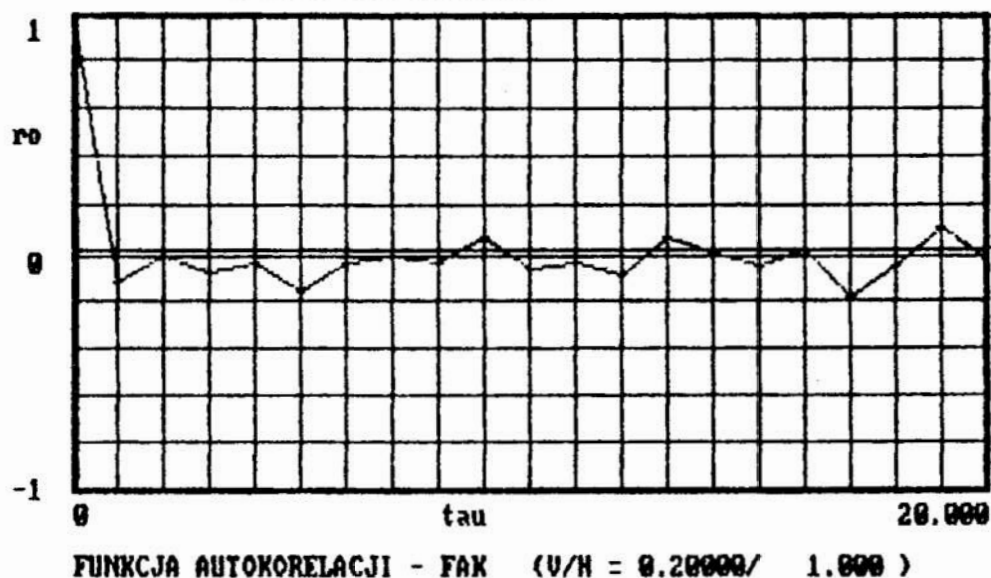


Rys. 3. Sygnal struktury nieokresowej
 Fig. 3. Signal of aperiodic structure

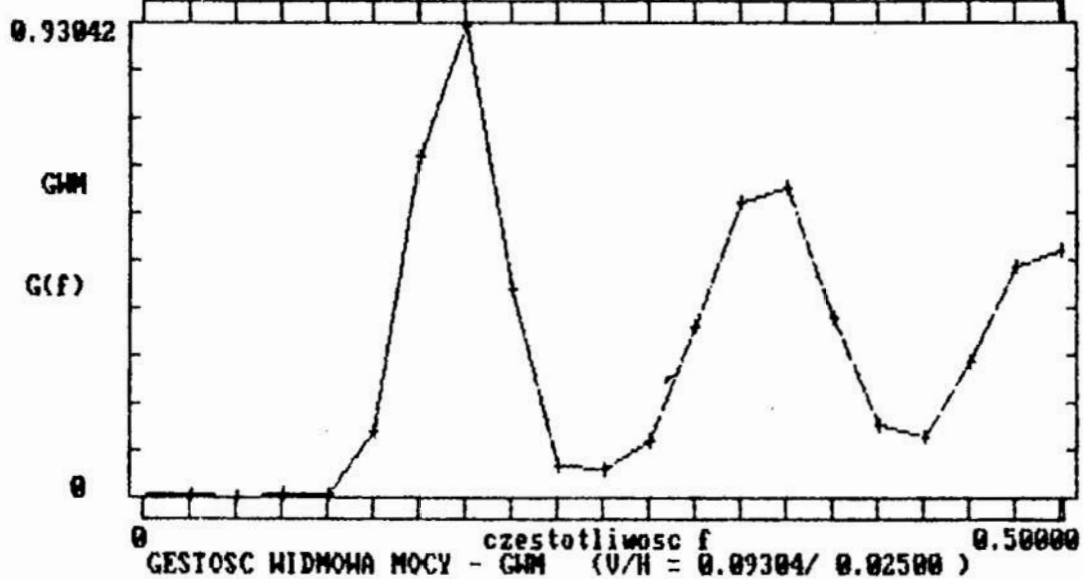
Dla każdego z tych sygnałów wyznaczono zgodnie z wzorami (1) i (2) estymatory funkcji autokorelacji, umożliwiające sporządzenie wykresów FAK:



Rys. 4. Wykres funkcji autokorelacji dla struktury okresowej
 Fig. 4. Autocorrelation function for a periodic structure



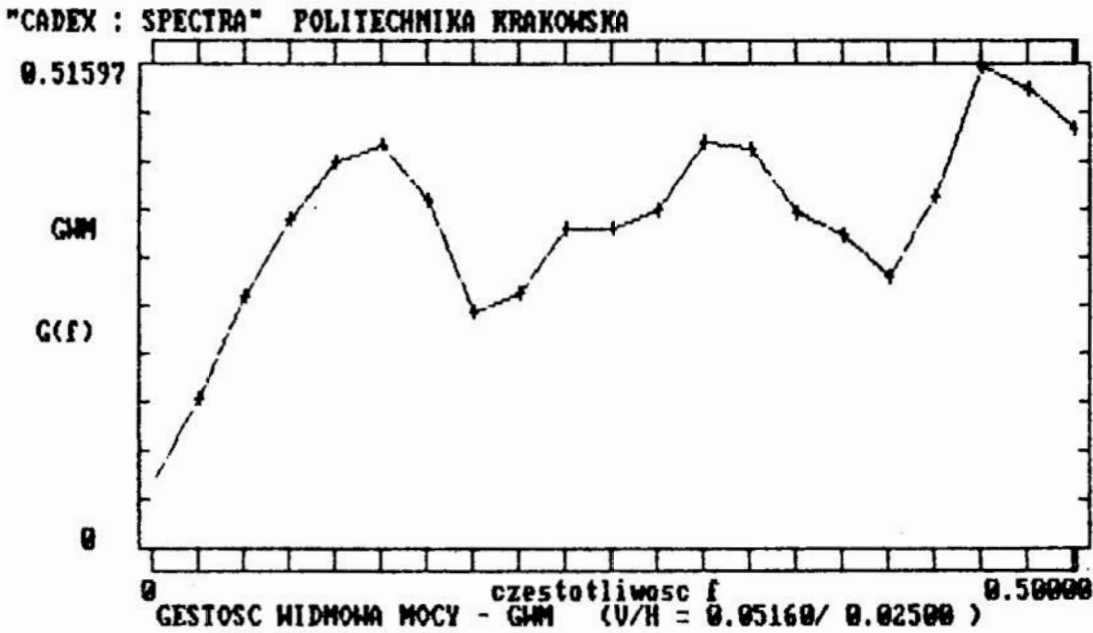
Rys. 5. Wykres funkcji autokorelacji dla struktury nieokresowej
 Fig. 5. Autocorrelation function for an aperiodic structure



Rys. 6. Wykres gęstości widmowej mocy dla struktury okresowej
 Fig. 6. Spectral density of power for a periodic structure

Analiza wykresów przedstawionych na rys. 4 i 5 wskazuje na całkowicie odmienny typ analizowanych ciągów. Wykres funkcji autokorelacji dla struktur okresowej jest typowym przebiegiem tej funkcji dla sygnału sinusoidalnego z szumem losowym. Jego asymetria względem osi odciętych wynika z badanego ciągu - większe

ilości "0" opisujących tło niż "1" odpowiadających za inne elementy struktury. Wykres z rys. 5 to przebieg FAK dla szumu losowego. Spostrzeżenia te potwierdzają analogiczne wykresy funkcji gęstości widmowej mocy przedstawione na rys. 6 i 7.



Rys. 7. Wykres gęstości widmowej mocy dla struktury nieokresowej
 Fig. 7. Spectral density of power for an aperiodic structure

4. Wnioski

Przedstawione wyniki badań okresowości struktury materiału za pomocą metod analizy stochastycznej, przy wykorzystaniu komputerowego programu do planowania i analizy doświadczeń SPECTRA 3.1. upoważniają do przedstawienia następujących wniosków:

1. Na przykładzie analizy dwóch ciągów binarnych symulujących opis struktur okresowej i nieokresowej materiału wykazano możliwość zastosowania metod analizy procesów stochastycznych w badaniach okresowości struktury materiału.

2. Zastosowany program komputerowego wspomaganie planowania i analizy doświadczeń SPEKTRA 3.1. opracowany w Instytucie Materiałoznawstwa i Technologii Metali Politechniki Krakowskiej, daje wygodne i szybkie narzędzie do analizy procesów stochastycznych, w tym również do badania okresowości struktury materiału.

3. Przedstawione badania modelowe nie rozwiązują szeregu problemów związanych z analizą i przetwarzaniem obrazu, a więc częścią badań dostarczającą materiał liczbowy do analizy statystycznej. Zagadnienia stereologiczne nie były jednak celem niniejszego artykułu.

5. Literatura

1. Górecka R., Polański Z.: Metrologia warstwy wierzchniej. WNT Warszawa 1983.

The use of stochastic process analysing methods to investigate the periodicity of material structure

Summary

In the paper the possibilities of application of stochastic processes analysing methods to investigate the periodicity of material structure is presented. SPECTRA 3.1 computer programme has been used throughout the work to design and analyse the experiments.

The role of statistical analysis of material structure as an important element of materials engineering has been emphasized.

STOCHASTYCZNY MODEL KONSTYTUOWANIA CHROPOWATOŚCI POWIERZCHNI W PROCESIE FREZOWANIA CZOŁOWEGO OSTRZAMI O ZAOKRĄGLONYM NAROZU

Przedstawiono modelowanie konstytuowania mikronierówności powierzchni w procesie frezowania czołowego ostrzami o zaokrąglonym narożu. Model ujmuje wpływ stereometryczno-kinematycznego odwzorowania ostrzy, przemieszczeń względnych w układzie narzędzie - przedmiot obrabiany oraz bicia czołowego ostrzy na chropowatość powierzchni frezowanej czołowo. Przemieszczenia względne oraz bicie czołowe ostrzy stanowią ograniczenie dla chropowatości powierzchni obrobionej.

1. Wstęp

Frezowanie czołowe z powodu swoich zalet takich jak duża wydajność oraz duża dokładność wymiarowo - kształtowa jest coraz powszechniej stosowane także jako obróbka wykańczająca. Na chropowatość powierzchni frezowanej czołowo wpływają takie czynniki jak stereometryczno - kinematyczne odwzorowanie ostrza, przemieszczenia względne narzędzia i przedmiotu obrabianego, bicie czołowe ostrzy freza, oraz inne czynniki fizyczne. Proces konstytuowania chropowatości powierzchni frezowanej czołowo mimo podejmowanych prób [1] jest mało zbadany, a także niedostatecznie opisany. W pracy podjęto próbę opisu tego procesu. Przedstawiono stochastyczny model konstytuowania mikronierówności powierzchni uwzględniający oprócz odwzorowania ostrzy, przemieszczenia względne narzędzia i przedmiotu obrabianego oraz czołowe bicie ostrzy o zaokrąglonym narożu.

2. Model uwzględniający stereometryczno - kinematyczne odwzorowanie ostrzy

Jeśli proces frezowania czołowego odbywa się frezem lub głowicą frezową z ostrzami o zaokrąglonym kształcie o promieniu r_ϵ bez bicia i drgań, to położenie zarysu tego ostrza na płaszczyźnie Oxy w chwili początkowej może być zapisane następująco

$$y_0(x) = \frac{1}{2r_\epsilon} x^2 \quad (1)$$

gdzie: r_ϵ - promień naroża ostrza.

Zależność (1) jest równaniem paraboli ściśle stycznej zastępującej rzeczywisty zarys ostrza. Przy obróbce z małymi posuwami, przyjęcie zarysu ostrza jako paraboli ściśle stycznej w miejsce zarysu kołowego daje błąd pomijalnie mały [2].

Profil poprzeczny utworzonej powierzchni po przejściu i - tego ostrza ($i = 1, 2, \dots, z$) podczas l - tego ($l = 1, 2, \dots, n$) obrotu freza przemieszczającego się wzdłuż osi x ruchem posuwowym (rys. 1) dany jest równaniem

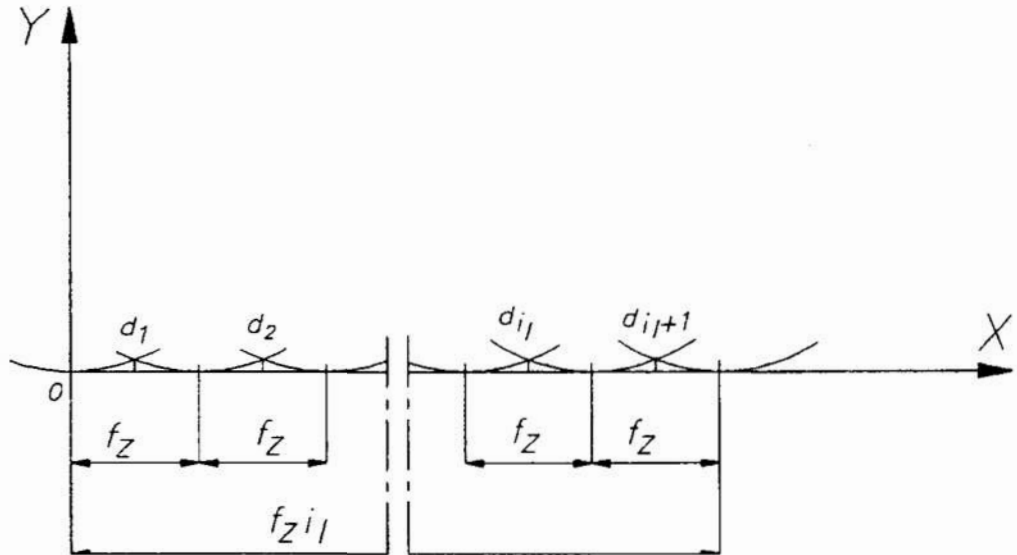
* Samodzielny Zakład Technologii Maszyn Politechniki Świętokrzyskiej w Kielcach 25-314 Kielce Al. Tysiąclecia Państwa Polskiego 3.

$$y_{i_1}(x) = \frac{1}{2r_c}(x - f_z i_1)^2 \quad \text{dla } d_{i_1} < x \leq d_{i_1+1} \quad (2)$$

gdzie : f_z - posuw na ostrze freza ,

i_1 - i - te ostrze podczas l - tego obrotu ($i_1 = z(l-1) + i - 1$)

jest to równanie krzywej okresowej o okresie f_z dla $f_z(i_1 - \frac{1}{2}) < x \leq f_z(i_1 + \frac{1}{2})$.



Rys. 1. Profil poprzeczny nierówności powierzchni frezowanej czołowo z posuwem f_z ostrzami o zaokrąglonym narożu.

Fig. 1. Transverse cross-section of surface inequality formed by face-milling with a feed rate f_z using blades with round cutting edges.

Wysokość nierówności powierzchni tego profilu wynosi

$$R_{z1} = y_{i_1}(d_{i_1}) = \frac{f_z^2}{8r_c} \quad (3)$$

Średnie arytmetyczne odchylenie profilu (danego funkcją $y_{i_1}(x)$) od położenia linii średniej \bar{y} wyznaczonym na odcinku elementarnym l_e oblicza się z zależności

$$R_a = \frac{1}{l_e} \int_0^{l_e} |y(x) - \bar{y}| dx \quad (4)$$

Wartość średnia profilu wynosi

$$\bar{y} = \frac{1}{l_e} \int_0^{l_e} y(x) dx = \frac{1}{f_z} \int_{d_{i_1}}^{d_{i_1+1}} y_{i_1}(x) dx \quad (5)$$

Podstawiając zależność (2) do wzoru (5) otrzymamy

$$\bar{y} = \frac{1}{24} \frac{f_z^2}{r_c} \quad (6)$$

Wstawiając do równania (4) otrzymamy

$$R_a = \frac{1}{f_z} \int_{d_{i1}}^{d_{i+1}} |y_{i1}(x) - \bar{y}| dx = \frac{1}{18\sqrt{3}} \frac{f_z^2}{f_e} \quad (7)$$

3. Model uwzględniający odwzorowanie ostrzy, przemieszczenia względne narzędzia i przedmiotu obrabianego oraz bicie czołowe ostrzy freza

W praktyce procesowi frezowania czołowego towarzyszą drgania względne freza oraz bicie czołowe jego ostrzy co będzie odwzorowane na powierzchni obrobionej. Podczas konstytuowania mikronierówności powierzchni frezowanej czołowo podobnie jak przy toczeniu [2,3] i frezowaniu walcowym [4] istotny wpływ posiadają drgania i bicie ostrzy w kierunku prostopadłym do powierzchni obrabianej t.j. w kierunku osi Oy .

Profil poprzeczny powierzchni utworzonej po przejściu i -tego ostrza podczas l -tego obrotu freza przemieszczającego się wzdłuż osi x ruchem posuwowym wykonującym dodatkowo drgania oraz czołowe bicie ostrzy można zapisać równaniem

$$y_{i1}(x) = \frac{1}{2r_a}(x - f_z i l)^2 + \rho_i + \xi_{i1} \quad (8)$$

gdzie: ξ_{i1} - chwilowe przemieszczenie względne i -tego ostrza podczas l -tego obrotu freza,

ρ_i - chwilowe bicie czołowe i -tego ostrza freza względem przedmiotu obrabianego.

Kolejne zarysy (parabole) $y_{i-1,1}, y_{i1}$ oraz $y_{i1}, y_{i+1,1}$ przecinają się w punktach d_{i1} i $d_{i+1,1}$ (patrz rys. 2)

$$d_{i1} = f_z(i l - \frac{1}{2}) + \frac{f_z}{f_e}(\rho_i - \rho_{i-1} + \xi_{i1} - \xi_{i-1,1}) \quad (9)$$

$$d_{i+1,1} = f_z(i l + \frac{1}{2}) + \frac{f_z}{f_e}(\rho_{i+1} - \rho_i + \xi_{i+1,1} - \xi_{i1}) \quad (10)$$

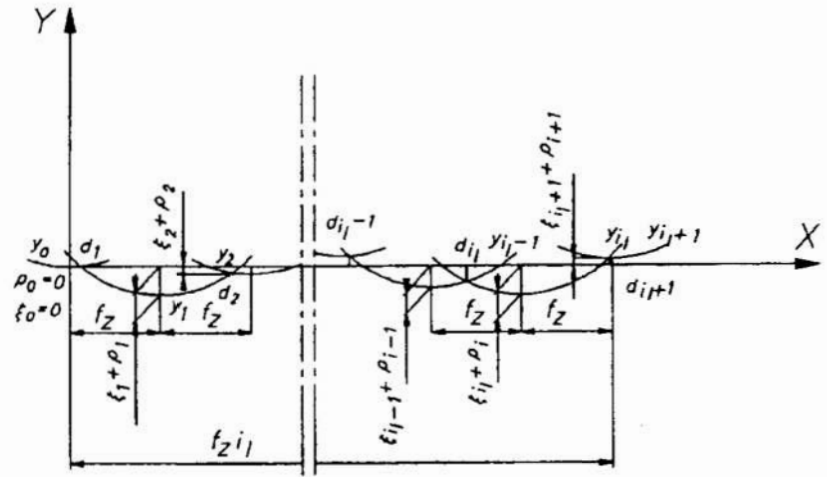
Zarys powierzchni frezowanej czołowo opisany jest równaniem (8) dla $d_{i1} < x \leq d_{i+1,1}$

Przyjmujemy, że ξ_{i1} są ciągiem niezależnych zmiennych losowych o jednakowym rozkładzie prawdopodobieństwa z wartością średnią równą zero i wariancją $D^2(\xi)$. Gdy ξ_{i1} są niezależnymi zmiennymi losowymi, to zarys powierzchni $y(x)$ jest procesem stochastycznym.

Bicie ostrzy freza ρ_i można opisać zależnością

$$\rho_i = -\frac{e}{2} \cos(i-1) \frac{2\pi}{z} \quad (11)$$

gdzie: z - ilość ostrzy freza lub głowicy frezowej,
 e - amplituda bicia czołowego ostrzy freza.



Rys. 2. Profil poprzeczny mikronierówności powierzchni frezowanej czołowo z posuwem f_z ostrzami zaokrąglonymi o promieniu r_e powstający przy uwzględnieniu przemieszczeń względnych ξ_{i1} narzędzia i przedmiotu obrabianego oraz bicia czołowego ostrzy freza ρ_i .

Fig. 2. Transverse cross-section of on finished surface formed by face-milling with a feed rate f_z using round cutting edged blades with a radius r_e formed considering relative displacements ξ_{i1} of the tool and the workpiece, as well as the face run-out of the milling cutter ρ_i .

Do wyznaczenia parametru R_a modelowanej powierzchni wymagającej określenia funkcjonału na zbiorze realizacji tego procesu stochastycznego można wykorzystać nierówność Schwartza [2], z której wynika, że

$$R_a^2 \leq [\overline{y^2} - \bar{y}^2] \quad (12)$$

Nierówność (12) można zastąpić równaniem

$$R_a = \Theta \sqrt{\overline{y^2} - \bar{y}^2} \quad (13)$$

gdzie: \bar{y} - wartość średnia modelowanego profilu,

Θ - współczynnik kształtu modelowanego profilu.

Wartości \bar{y} i $\overline{y^2}$ wyznaczmy z zależności

$$\bar{y} = \lim_{l_e \rightarrow \infty} \frac{1}{l_e} \int_0^{l_e} y(x) dx \quad \text{ i } \quad \overline{y^2} = \lim_{l_e \rightarrow \infty} \frac{1}{l_e} \int_0^{l_e} [y(x)]^2 dx$$

Podstawiając do tych zależności równanie (8) otrzymamy

$$\bar{y} = \lim_{n \rightarrow \infty} \frac{1}{nf_z} \sum_{i_1=0}^n \int_{d_{i_1}}^{d_{i_1+1}} \left[\frac{1}{2r_e} (x - f_z i_1) + \rho_i + \xi_{i_1} \right] dx \quad (14)$$

$$\overline{y^2} = \lim_{n \rightarrow \infty} \frac{1}{nf_z^2} \sum_{i_1=0}^n \int_{d_{i_1}}^{d_{i_1+1}} \left[\frac{1}{2r_e} (x - f_z i_1) + \rho_i + \xi_{i_1} \right]^2 dx \quad (15)$$

Podstawiając do równań (14) i (15) zależności (9) i (10) oraz dokonując obliczeń i pomijając momenty wyższego rzędu niż drugi zmiennych ξ_{ii} , jak i potęgi tych momentów oraz sumy zmiennych ρ_i w potęgach parzystych wyższych od drugiej jako małe wyższego rzędu otrzymamy

$$\bar{y} = \frac{f_z^2}{24r_e} - \frac{r_a}{f_z^2} [D^2(\xi) + D^2(\rho)(1 - \cos \frac{2\pi}{z})] \quad (16)$$

$$\overline{y^2} = \frac{f_z^4}{320r_e^2} + \frac{5}{6} [D^2(\xi) + D^2(\rho)(1 + \frac{1}{5} \cos \frac{2\pi}{z})] \quad (17)$$

następnie wyznaczamy różnicę

$$\overline{y^2} - \bar{y}^2 = \frac{f_z^4}{720r_e^2} + \frac{5}{6} [D^2(\xi) + D^2(\rho)(1 + \frac{1}{5} \cos \frac{2\pi}{z})] \quad (18)$$

Wstawiając do zależności (18) $D(\xi) = 0$ i $D(\rho) = 0$ oraz podstawiając do (13) i porównując z (7) otrzymamy

$$R_a = \Theta \sqrt{\frac{f_z^4}{720r_e^2}} = \frac{1}{18\sqrt{3}} \frac{f_z^2}{r_e} \quad \text{stad} \quad \Theta = \frac{2\sqrt{5}}{3\sqrt{3}}$$

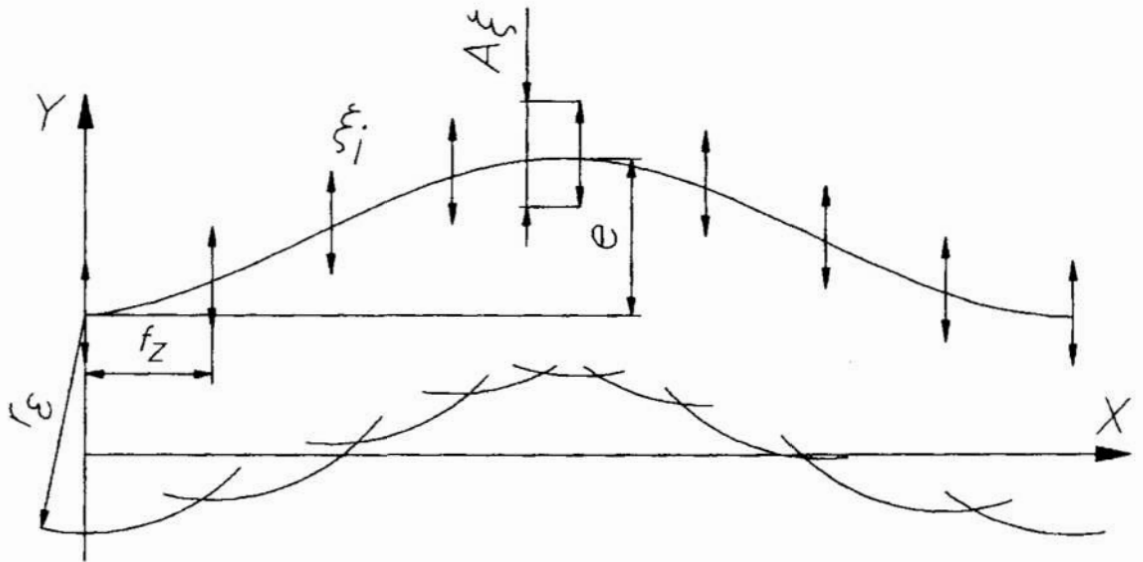
Podstawiając za $\Theta = \frac{2\sqrt{5}}{3\sqrt{3}}$ i zależność (18) do równania (13) otrzymamy

$$R_a = \sqrt{\frac{f_z^4}{972r_e^2} + \frac{50}{81} [D^2(\xi) + D^2(\rho)(1 + \frac{1}{5} \cos \frac{2\pi}{z})]} \quad (19)$$

Wysokość nierówności powierzchni można wyznaczyć korzystając z zależności (3) oraz uwzględniając przemieszczenia względne freza i bicie czołowe ostrzy (rys. 3)

$$R_{zt} = \frac{f_z^2}{8r_e} + e + A_\xi \quad (20)$$

gdzie : A_ξ - amplituda przemieszczeń względnych freza i przedmiotu obrabianego .



Rys. 3. Zarys powierzchni powstającej w procesie frezowania czołowego, frezem o zaokrąglonym ostrzu o promieniu r_e uwzględniający bicie czołowe ostrzy o amplitudzie e .

Fig. 3. Profile of a surface formed in the face-milling process using a cutter with round cutting edged blades with a radius r_e , taking into account the run-out with an amplitude e .

Dla przypadku gdy $\frac{f_z}{e} > 40$ bicie czołowe ostrzy freza będzie powodować falistość. Wtedy we wzorze (19) należy podstawić $D^2(\rho) = 0$, a we wzorze (20) $e = 0$ i otrzymamy (najczęściej występujący przypadek w praktyce)

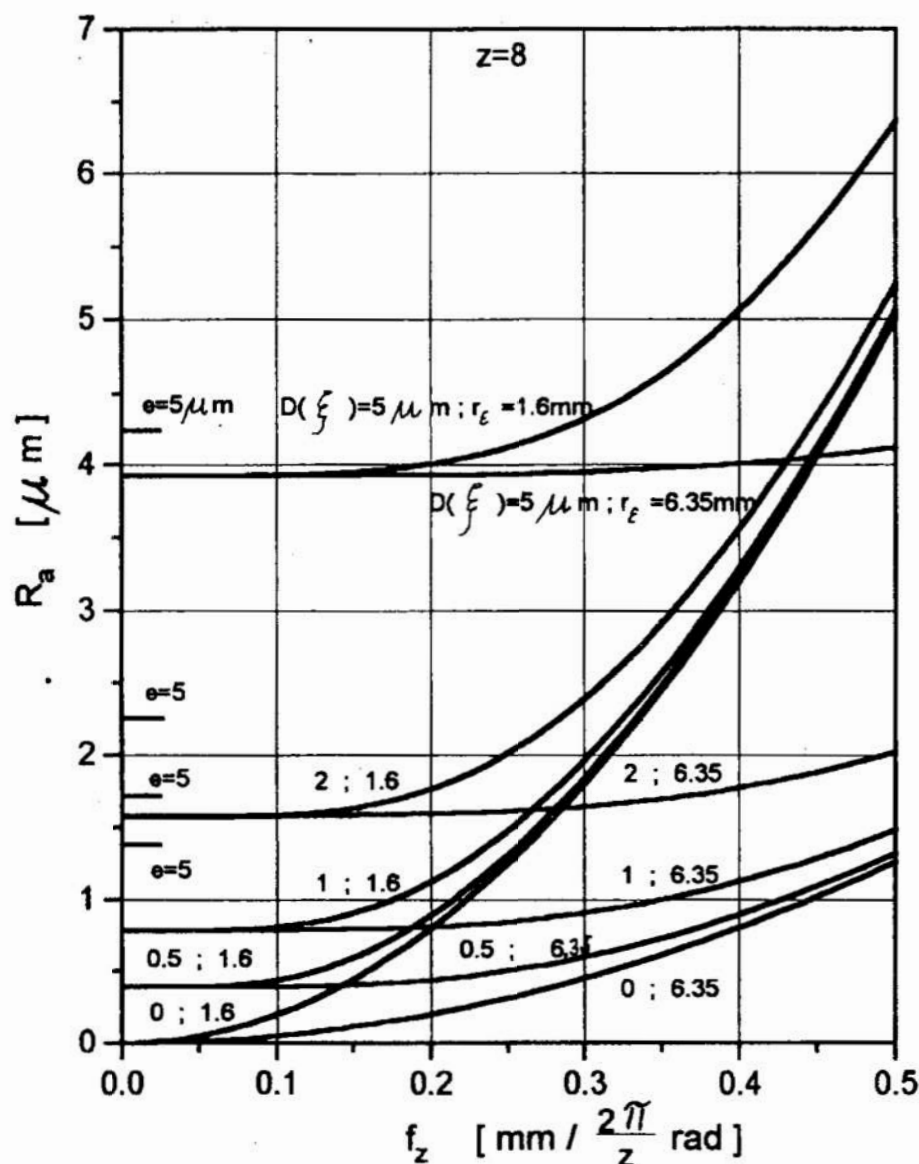
$$R_a = \sqrt{\frac{f_z^4}{972r_e^2} + \frac{50}{81}D^2(\xi)} \quad (21)$$

$$R_{z1} = \frac{f_z^2}{8r_e} + A_{\xi} \quad (22)$$

Dla przypadku jak na rys.3 we wzorze (19) można podstawić $D^2(\rho) = \frac{e^2}{8}$ stąd otrzymujemy

$$R_a = \sqrt{\frac{f_z^4}{972r_e^2} + \frac{50}{81}[D^2(\xi) + \frac{e^2}{8}(1 + \frac{1}{5}\cos\frac{2\pi}{z})]} \quad (23)$$

Zależność parametru R_a od posuwu na ostrze f_z oraz promienia naroża r_e , odchylenia standardowego przemieszczeń względnych $D(\xi)$, a także amplitudy bicia czołowego ostrzy e oraz liczby ostrzy freza z przedstawiono na rys. 4.



Rys. 4. Przykład nomogramu do prognozowania wartości parametru R_a chropowatości powierzchni frezowanej czołowo ostrzami o zaokrąglonym narożu.

Fig. 4. Example of a nomogram forecasting the value of the parameter R_a surface roughness formed by means of face-milling denoting using blades with rounded corners.

Analizując krzywe na rys. 4 widzimy, że wpływ posuwu na wartość parametru R_a jest znacznie mniejszy dla ostrza o promieniu $r_e = 6.35$ mm jak dla ostrza o $r_e = 1.6$ mm, maleje on także w miarę wzrostu odchylenia standardowego drgań $D(\xi)$. W miarę zmniejszania się posuwu wartość parametru R_a maleje ale tylko do pewnej wartości granicznej zależnej od wartości przemieszczeń względnych w układzie narzędzie przedmiot obrabiany, która wynosi $\sqrt{\frac{50}{81}} D(\xi)$. Dla bardzo małych posuwów $f_z < \frac{40e}{z}$ wartość tego parametru nawet wzrasta, ponieważ bicie czołowe ostrzy powoduje powstawanie mikronierówności, które dla tych posuwów stanowią chropowatość.

4. Wnioski

- Na chropowatość powierzchni konstytuowanej w procesie frezowania czołowego oprócz stereometryczno - kinematycznego odwzorowania ostrzy w istotny sposób wpływają przemieszczenia względne w układzie narzędzie - przedmiot obrabiany oraz bicie czołowe ostrzy .

- Wpływ posuwu na wartość parametru R_a jest coraz mniejszy w miarę wzrostu promienia naroża ostrza freza r_ϵ oraz w miarę wzrostu wartości przemieszczeń względnych $D(\xi)$.

- Przemieszczenia względne narzędzia i przedmiotu obrabianego oraz bicie czołowe ostrzy freza stanowią ograniczenie dla chropowatości powierzchni konstytuowanej w procesie frezowania czołowego .

- W celu minimalizacji chropowatości powierzchni obrobionej należy dążyć do zmniejszenia wartości przemieszczeń względnych $D(\xi)$ oraz używać ostrzy o stosunkowo dużym promieniu zaokrąglenia r_ϵ , co umożliwi stosowanie dużych posuwów f_z i spowoduje wzrost wydajności obróbki .

5. Literatura

- [1] Olszak W., Cieloszyk J., Wpływ czynników geometrycznych i kinematycznych na kształtowanie powierzchni w procesie frezowania czołowego, II Międzynarodowa Konferencja Naukowo-Techniczna " Wpływ technologii na stan warstwy wierzchniej -WW'93". Gorzów Wlkp - Lubniewice 1993 .
- [2] Rusek P. , Dynamika procesu konstytuowania chropowatości powierzchni sposobami obróbki wiórowej , Zeszyty Naukowe AGH , Elektryfikacja i Mechanizacja Górnictwa i Hutnictwa nr 122 , Kraków 1979 .
- [3] Miko E. , Rusek P. , Skrzypiński A. , Modelowanie chropowatości powierzchni konstytuowanej w procesie toczenia , Archiwum Technologii Budowy Maszyn z. 8 , WPP , Poznań 1990 .
- [4] Stępiński L. , Wpływ przemieszczeń względnych w układzie : narzędzie - przedmiot obrabiany na chropowatość powierzchni frezowanej walcowo , Praca doktorska , AGH Kraków 1982 .

THE STOCHASTIC MODEL CONSTITUTING THE FORMATION OF SURFACE ROUGHNESS IN THE FACE-MILLING PROCESS BY MEANS OF BLADES WITH ROUND CUTTING EDGES

Summary

The model constituting the formation of on finished surfaces in face-milling processes using blades with round cutting edges is represented .The model formulates the influence on irregularity of face-milled surfaces by stereo-kinematic transformation of blade , relative displacements in the tool-workpiece system and the face run-out . Relative displacement as well as face run-out of a blade limit the inequality of the machined surface.

MOŻLIWOŚCI ZASTOSOWANIA ANALIZY PROCESÓW STOCHASTYCZNYCH DO OCENY ZJAWISK CIEPLNO-PRZEPLYWOWYCH MIESZANIN WIELOFAZOWYCH

W pracy przedstawiono wyniki dotychczasowych badań eksperymentalnych przepływu mieszanin wielofazowych (gaz-ciecz, gaz-ciało stałe) w których wykorzystano do analizy parametry i funkcje stochastyczne. Na podstawie analizy fluktuacji wybranych parametrów procesowych opracowano metodę oceny struktury płynącej mieszaniny wielofazowej.

1. Wprowadzenie

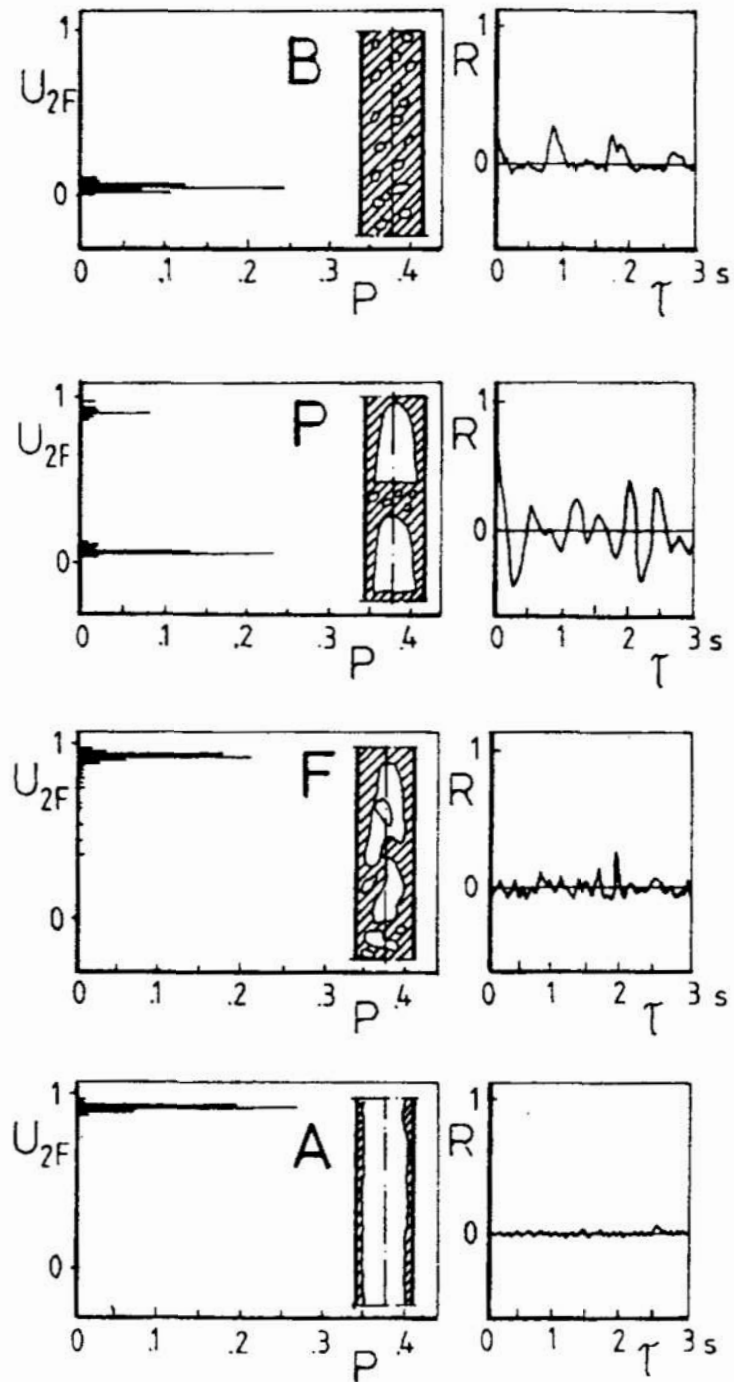
Badania przepływu wielofazowego, który występuje powszechnie w przemyśle chemicznym, energetycznym i pokrewnych, wymagają, z uwagi na swoją naturę (równoczesny przepływ wielu faz np. gaz-ciecz, gaz-ciało stałe, czy też gaz-ciecz-ciało stałe) szczególnego podejścia. Z uwagi na bardzo silne oddziaływanie wzajemne faz, znaczne oscylacje wartości pola prędkości i koncentracji faz, do analizy wyników badań należy wykorzystać teorię procesów losowych.

W pracy przedstawiono wyniki badań eksperymentalnych przepływu wielofazowego dla różnych składników mieszaniny w różnych układach geometrycznych. Poniżej omówiono wyniki badań zwracając uwagę na możliwości wykorzystania parametrów i funkcji stochastycznych do wyznaczania różnych parametrów charakteryzujących przepływ wielofazowy.

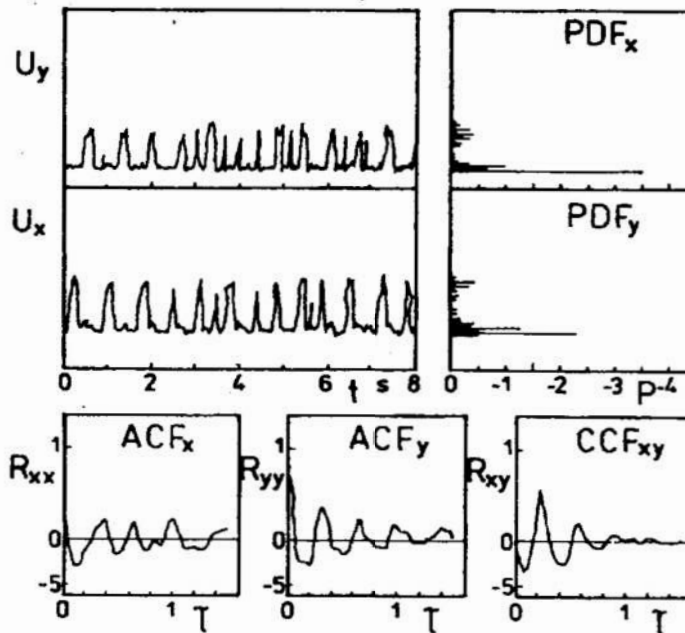
2. Przepływ dwufazowy gaz-ciecz w kanale

Na rys. 1 pokazano rozkłady funkcji gęstości prawdopodobieństwa PDF oraz funkcji autokorelacji ACF dla struktur przepływu mieszaniny gaz-ciecz w rurze pionowej. Jako cechę znamioną wykorzystano analizę fluktuacji udziału objętościowego gazu, mierząc zmiany oporności elektrycznej strugi dwufazowej w przekroju osiowym kanału. Wyraźne różnice w kształcie funkcji prawdopodobieństwa dla poszczególnych struktur skłoniły do opracowania metody identyfikacji struktur przepływu [1] z pominięciem subiektywnej metody bezpośredniej obserwacji wizualnej. Na podstawie analizy funkcji autokorelacji można wnioskować o charakterze procesu: dla przepływu pianowego jest to proces wyraźnie losowy, zaś dla korkowego wyraźnie periodyczny.

Na rys. 2 przedstawiono wyniki badań z wykorzystaniem podwójnej sondy rezystancyjnej. Na podstawie funkcji autokorelacji można wyznaczać częstość występowania struktur korkowych, zaś na podstawie funkcji interkorelacji sygnałów wyraźne ekstremum na funkcji interkorelacji świadczy o przesunięciu sygnałów w czasie; można zatem wyznaczyć prędkość przemieszczania się struktur korkowych w kanale.



Rys. 1. Funkcje stochastyczne PDF i ACF dla struktur przepływu dwufazowego gaz-ciecz w rurze pionowej : pęcherzykowej B, korkowej P, pianowej F i pierścieniowej A
 Fig. 1. Stochastic functions: PDF and ACF for particular flow patterns of two-phase gas-liquid flow in vertical pipe: bubble B, plug P, froth F and annular A



Rys. 2. Funkcje stochastyczne PDF, ACF i CCF dla struktury korkowej
 Fig. 2. Stochastic functions: PDF, ACF and CCF for plug flow

3. Przepływ dwufazowy gaz-ciecz w przestrzeni międzyrurowej

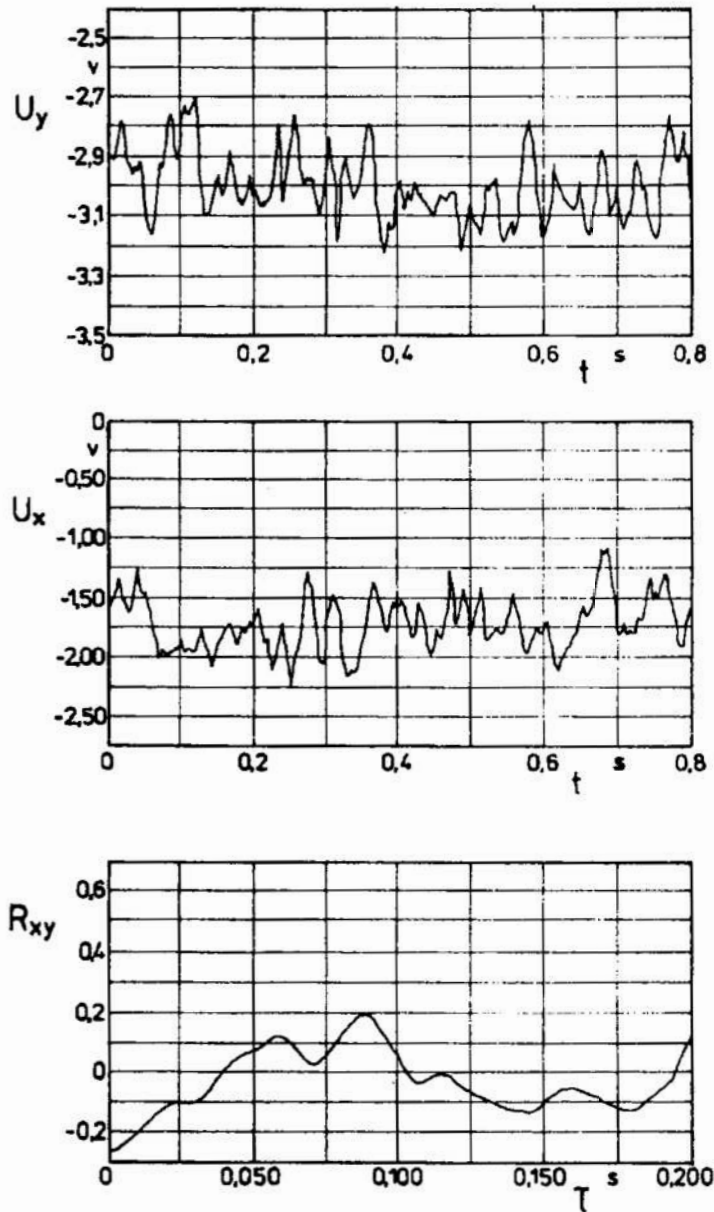
Badania przepływu w przestrzeni międzyrurowej stanowią z punktu widzenia geometrii tej przestrzeni bardzo poważny problem. Przy obliczaniu oporów przepływu mieszanin wielofazowych jednym z głównych parametrów do określenia jest udział objętościowy faz. Dla przepływu mieszaniny gaz-ciecz wyraźna różnica gęstości obu faz sprawia, że udział objętościowy gazu nie jest, z powodu tzw. poślizgu międzyfazowego, równy wartości wynikającej z udziału strumieni płynących faz. Dzięki umieszczeniu na powierzchni rur sond pomiarowych (nadajniki i odbiorniki światła dla sond optycznych oraz elektrody dla sond przewodnościowych) otrzymano nieważne metody pomiarowe [2]. Do pomiaru udziału objętościowego gazu wykorzystano korelacyjną metodę sygnałów. Po dokonaniu rejestracji napięć dla sond przewodnościowych umieszczonych w sąsiednich rzędach rur (na rys. 3 napięcia U_x i U_y), wyznaczano funkcję interkorelacji sygnałów oraz czas przesunięcia sygnałów zbieranych w sąsiednich rzędach rur dla punktu na krzywej, w którym pojawia się wartość maksymalna współczynnika interkorelacji. Średnia rzeczywista prędkość pęcherzyków wynosi

$$w_B = \frac{t}{\tau_{\max\{R_{xy}\}}} \quad (1)$$

zaś udział objętościowy gazu

$$\alpha_G = \frac{w_{Go}}{w_B} \quad (2)$$

gdzie w_{Go} jest prędkością pozorną gazu liczoną na minimalny przekrój przepływu pomiędzy rurami.

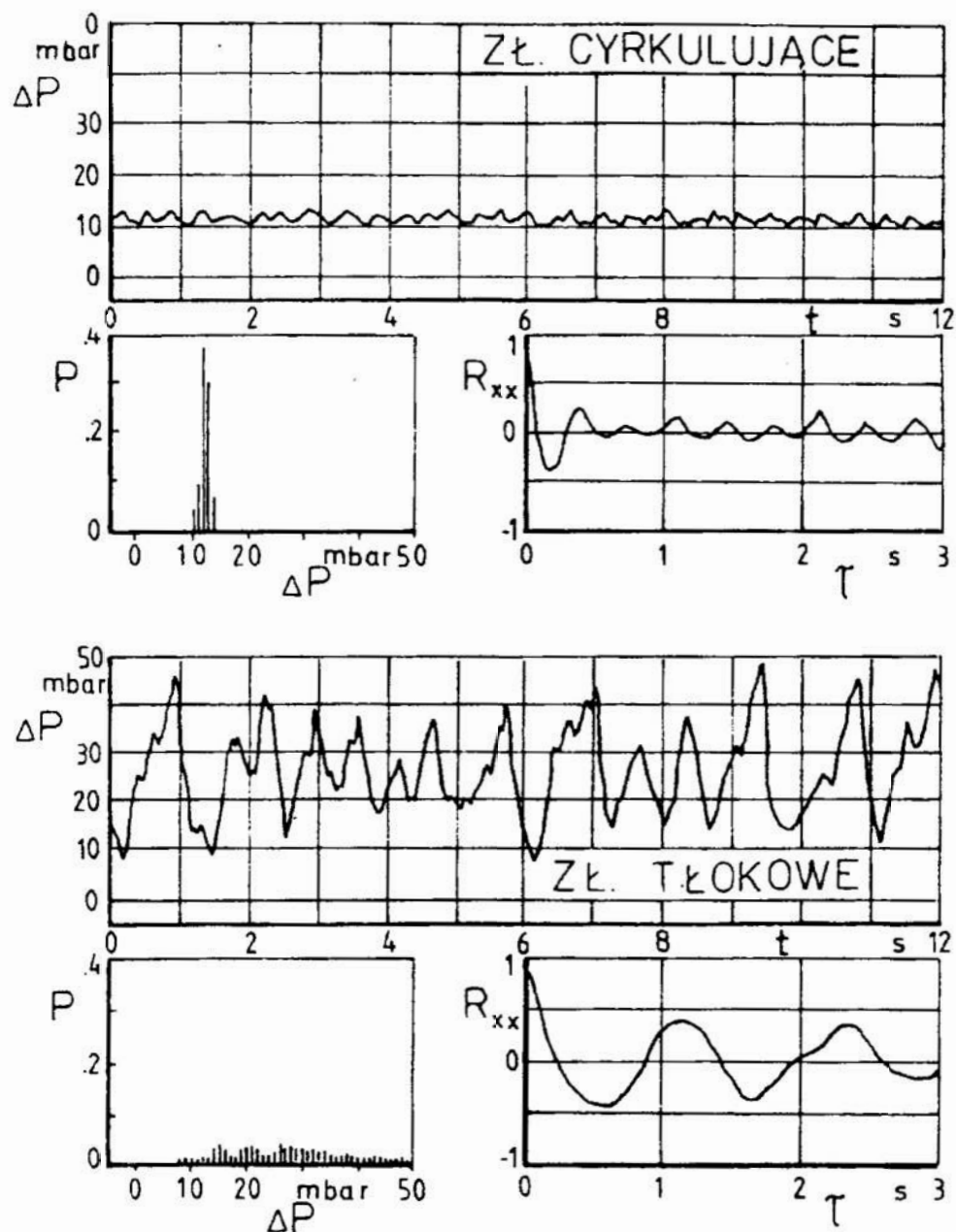


Rys.3. Przebiegi napięcia oraz funkcja CCF dla badań przepływu pęcherzykowego w przestrzeni międzyrurowej
 Fig. 3. Tracers of voltages and CCF function from conductivity probes for bubble flow across tube bundle

4. Przepływ dwufazowy gaz-ciało stałe w aparacie fluidalnym

W zależności od ilości doprowadzonego strumienia gazu a także gęstości i wielkości cząstek ciała stałego, mogą się pojawić w aparacie fluidalnym różne stany złoża [3]. Powszechnie przyjęta jest klasyfikacja stanów złoża w oparciu o obserwację wizualną. W celu zobiektywizowania oceny a także możliwości automatyzacji tej czynności zaproponowano określenie stanu złoża na podstawie analizy fluktuacji ciśnienia na ścianie aparatu [4]. Na rys. 4 przedstawiono przykładowo wyniki badań dla stanów złoża: z cyrkulującą warstwą ciała stałego oraz stan bliski zawałowi - złoże tłokowe. Dla złoża cyrkulującego fluktuacje ciśnienia mają charakter losowy, zaś dla

złoża łokowego - wyraźnie periodyczny. Na podstawie cykliczności funkcji autokorelacji można wyznaczyć częstotliwość własną pojawiania się korków gazu. Analiza funkcji gęstości widmowej mocy pozwala na ocenę zjawiska segregacji złoża polidispersyjnego będącego w stanie fluidalnym.

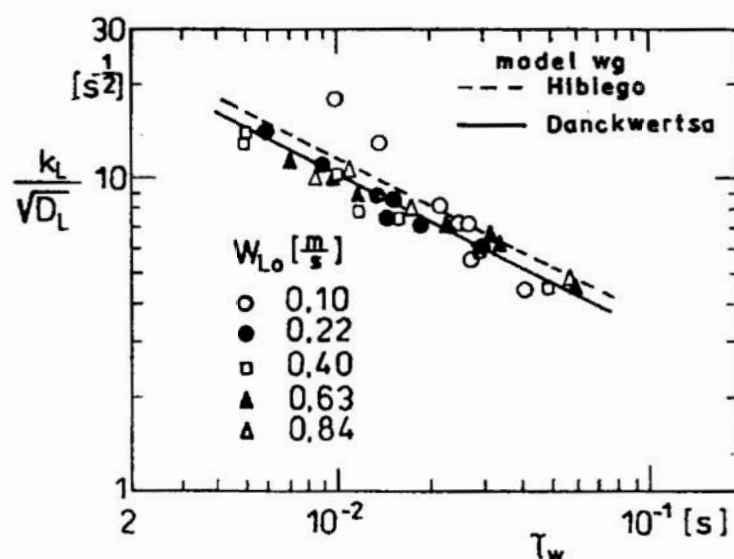


Rys. 4. Funkcje stochastyczne PDF i ACF dla struktur mieszaniny gaz-ciało stałe w złożu fluidalnym

Fig. 4. Stochastic functions: PDF and ACF for two-phase gas-solid flow structures in fluidized bed

5. Wymiana ciepła i masy przy przepływie mieszaniny gaz-ciecz

Opracowanie pewnych metod obliczeniowych dla potrzeb projektowania procesów transportu ciepła i masy przy przepływie wielofazowym stanowi cel dla wielu badaczy. W tradycyjnie prowadzonych badaniach, współczynnik wnikania ciepła wyznacza się na podstawie zmierzonego rozkładu temperatur. W pracach [5, 6] podjęto próbę wyznaczenia wartości współczynników wymiany ciepła i masy na podstawie oceny zjawisk hydrodynamicznych dla przepływu pierścieniowego w rurze z zafalowaną powierzchnią międzyfazową. Wykorzystując model penetracji zaproponowany w pracach [7,8], wartości współczynnika wymiany masy oblicza się w oparciu o częstość występowania fal, zakładając że fale jako struktury periodyczne mają związek z cyklicznym odnawianiem powierzchni kontaktu faz. Wyniki badań przedstawione na rys. 5 potwierdzają bardzo dobrą zgodność tego modelu. Pomiar częstości własnej występowania fal na powierzchni międzyfazowej sprowadza się do wyznaczenia funkcji autokorelacji sygnału dla sondy pomiarowej, którą rejestruje się fluktuacje grubości filmu cieczy płynącego po ścianie rury. Również w pracy [5] stwierdzono związek pomiędzy wartościami współczynnika wnikania ciepła a częstością występowania fal. Model penetracji może być zastosowany z powodzeniem do praktycznie wszystkich przypadków przepływu wielofazowego, wykazujących w skali mikro periodyczny charakter przepływu. Badania charakterystyk czasowych (funkcja interkorelacji) czy też częstotliwościowych (funkcja gęstości widmowej mocy) umożliwiają zastąpienie bardzo żmudnych i technicznie trudnych badań transportu ciepła i masy badaniami zjawisk przepływowych. Badania metodami korelacyjnymi mają z punktu widzenia metrologicznego ogromną zaletę nad innymi badaniami-kalibrowanie nie jest warunkiem niezbędnym poprawności uzyskanych wyników.



Rys. 5. Współczynnik wnikania masy dla przepływu pierścieniowego w rurze pionowej wyznaczony w oparciu o częstość występowania fal, wg [6]

Fig. 5. Mass transfer coefficient for annular flow in vertical pipe in relation of frequency of wave, in [6]

OZNACZENIA

- ACF - funkcja autokorelacji
CCF - funkcja interkorelacji
 D_L - współczynnik dyfuzji, cm^2/s
 k_L - współczynnik wnikania masy, cm/s
 ΔP - spadek ciśnienia, Pa
P - prawdopodobieństwo
PDF - funkcja gęstości prawdopodobieństwa
 R_{xx} - autokorelacja
 R_{xy} - interkorelacja
U - napięcie, V
w - prędkość, m/s
t - czas, s
 α - udział objętościowy faz
 τ - opóźnienie, s

LITERATURA

- [1] Ulbrich R.: Identyfikacja przepływu dwufazowego gaz-ciecz w kanale, Zeszyty Naukowe WSI Opole, s. Monografie z.32, Opole 1989.
- [2] Ulbrich R., Mewes D.: Gas void fraction in two-phase flow across tube bundle, 3rd Int. Symp. of Multiphase Flow and Heat Transfer, Xi'an 1994.
- [3] Hestroni G.: Handbook of Multiphase Systems, Hemisphere Publ. Corp., Washington 1982.
- [4] Szmolke N., Ściażko M., Ulbrich R.: Investigations of hydrodynamics of fluidized bed with circulation, 3rd Int. Symp. of Multiphase Flow and Heat Transfer, Xi'an 1994.
- [5] Hagiwara Y., Suzuki K., Sato T.: Studies on thin liquid film of annular-mist two-phase flow, Memoires of the Faculty of Engineering Kyoto University, 1982, vol. 44, no.2, s.309-328.
- [6] Tomita T.: Relation between wave characteristics and mass transfer coefficient in gas-liquid two-phase cocurrent upward flow, IEC Research, 1987, vol.26, no.7, s.1472-1475.
- [7] Dankwerts P.U.: Ind. Eng. Chem. 1951, vol.43, s.1460, wg [4].
- [8] Higbie R.: Trans. Am. Inst. Eng. 1935, vol.31, s.365, wg [4].

Application of Stochastic Process to Analysis of Transport Phenomena for Multiphase Flow

Summary

Results of experimental investigations for multiphase flow (gas-liquid, gas-solid) using stochastic parameters and functions have been presented. The method of flow pattern detection based on the analysis of fluctuations of chosen process parameters has been suggested.

MODELOWANIE MATEMATYCZNE IMPULSOWEJ OBRÓBKI ELEKTROCHEMICZNEJ (PECM)

W referacie przedstawione zostało zagadnienie powiązań kształtów elektrody roboczej i powierzchni obróbjonej impulsową obróbką elektrochemiczną (PECM). Podano model matematyczny procesu "idealnego" PECM oraz omówiono wyniki weryfikacji doświadczalnej.

1. Wprowadzenie

Obróbka elektrochemiczna (ECM), opierająca się na uzyskiwaniu kształtu przedmiotu obrabianego (PO) przez anodowe rozwarzanie elektrochemiczne (naddatku), zajmuje obecnie trwałą pozycję w wykonawstwie części maszyn i narzędzi. Na powyższe składają się zalety ECM, takie jak: możliwość uzyskiwania skomplikowanych kształtów, nie wprowadzanie „własnych” zmian w stanie warstwy wierzchniej (np. naprężeń wewnętrznych czy zmian mikrotwardości), nie zużywanie się elektrody roboczej - narzędzia (ER), niezależność wskaźników użytkowych obróbki od własności mechanicznych materiału obrabianego, wysoka wydajność obróbki przy stosunkowo wysokiej jakości obróbjonej powierzchni (przykładowo uzyskuje się $2 \cdot 10^3 \text{ mm}^3/\text{min}$ i $Ra=0.64 \mu\text{m}$).

Szczególnie ECM jest efektywna przy wytwarzaniu części i narzędzi z materiałów trudnych do obróbki mechanicznej, a więc ze stopów żarowytrzymałych i żaroodpornych, stopów tytanowych, berylowych oraz stopów o wysokiej wytrzymałości.

Cechy negatywne wiążą się z trudnościami:

- w uzyskiwaniu złożonych kształtów z tolerancjami wymiarów poniżej $100 \mu\text{m}$,
- w projektowaniu elektrody roboczej, zwłaszcza przy złożonym kształcie wyrobu, gdyż kształty powierzchni roboczej narzędzia i powierzchni obróbjonej różnią się,
- w zabezpieczaniu przed zwarciami elektrycznymi prowadzącymi do uszkodzeń elektrody roboczej i przedmiotu obrabianego.

Inne wady mają charakter ogólny, gdyż dotyczą wszystkich sposobów obróbki elektrochemicznej a należą do nich przede wszystkim zwiększenie kosztów związanych z ochroną środowiska oraz zabezpieczaniem przed korozją.

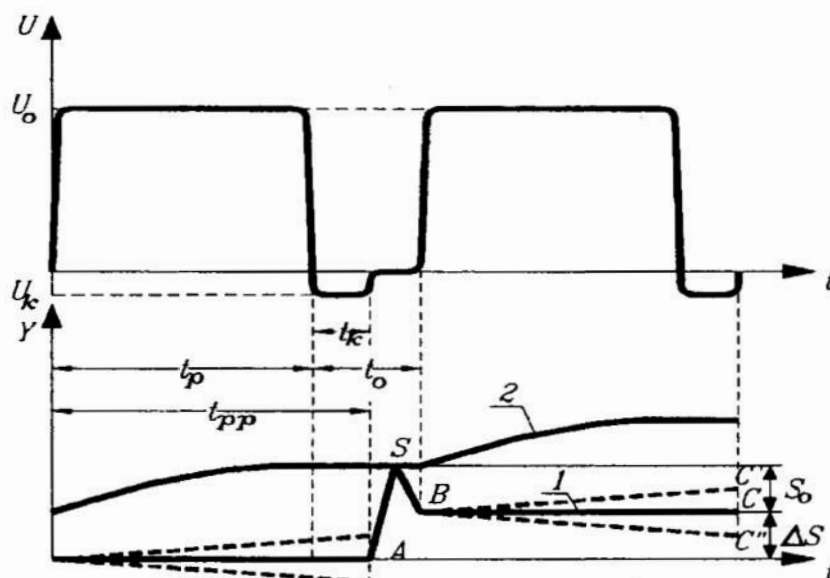
Wśród różnych sposobów zmniejszania wymienionych ujemnych cech obróbki ECM, czołowe miejsce zajmuje impulsowa obróbka elektrochemiczna (PECM - Pulse Electrochemical Machining). W PECM rozwarzanie elektrochemiczne zachodzi w określonych odcinkach czasu, najczęściej o długości poniżej 5 ms.

W procesie tym na elektrody (ER - PO) podawany jest impuls napięcia roboczego charakteryzujący się określonym przebiegiem w czasie, amplitudą U_0 oraz czasem

^a Non-traditional Manufacturing Research Cenetr, University of Nebraska-Lincoln, USA

^b Instytut Technologii Mechanicznej PW, Al. Niepodległości 222, 00-663 W-wa

impulsu t_p (rys. 1). W przerwach t_o napięcie równe jest zero (proces unipolarny) lub przyjmuje wartość ujemną U_k w odniesieniu do przedmiotu obrabianego (proces bipolarny) w ciągu określonego czasu t_k . Ruch elektrody roboczej (ER) jest zmienny w czasie, przy czym najczęściej realizuje się go jako nałożenie ruchu harmonicznego okresowego i postępowego lub jako ruchu ze skokowo zmieniającymi się prędkościami. W tym ostatnim przypadku możliwe staje się kontrolowanie i ustalanie szczeliny roboczej na początku impulsu lub grupy impulsów prądu, co ma duże znaczenie dla dokładności obróbki.



Rys. 1. Wykres zmiany napięcia oraz położenia elektrod w czasie:
1 - położenie elektrody roboczej-katody, 2 - położenie powierzchni obrobionej-anody

Ruch elektrody roboczej może być wówczas następujący (rys. 1):

- dosuw ER do styku z PO (odcinek AS),
- odsunięcie ER na ściśle określoną odległość S_0 (odcinek SB),
- ER nieruchoma (BC) lub porusza się z prędkością V_p lub $-V_p$ (BC' lub BC'').

Istotne znaczenie ma wzajemne powiązanie przebiegu impulsów prądowych z ruchem elektrody. Impuls napięcia najczęściej jest podawany na odcinku BC (BC' lub BC''), a w przypadku ruchu harmonicznego elektrody roboczej w półokresie $V_p < 0$.

Roztworzenie elektrochemiczne (usuwanie nadmiaru) odbywa się w czasie t_p , gdy na elektrody podawany jest impuls napięcia roboczego. O intensywności roztwarzania można wnioskować na podstawie natężenia prądu w impulsie (obrabiarki PECM z reguły zaopatrzone są w oscyloskopy dwukanałowe). W przerwie między impulsami ($t_o = t_{pp} - t_p$) lub dłuższej przerwie między grupami impulsów następuje usuwanie produktów obróbki (wodorotlenków, gazu, ciepła) oraz doprowadzenie do obszaru obróbki oczyszczonego elektrolitu.

Istotnym efektem zastosowania PECM jest możliwość realizacji procesu roztwarzania przy małych grubościach szczeliny (poniżej 0.1 mm). Praktyka i teoria obróbek elektrochemicznych wykazuje, że niedokładność obróbki jest proporcjonalna do odległości między elektrodami (ER i PO). Stosowanie w PECM grubości szczelin poniżej 0.1 mm pozwala na kilkakrotne podwyższenie dokładności (przykładowo błąd kształtu matrycy zmniejsza się z 0.2 ± 0.3 mm do 0.01 ± 0.02 mm):

Dodatkowe pozytywne zjawiska, które towarzyszą impulsowemu procesowi ECM, wiążą się z następującymi czynnikami decydującymi o dokładności oraz o kosztach obróbki:

- a. w przypadku prowadzenia procesu PECM z parametrami spełniającymi warunek: $wt_p \ll L$, (gdzie: w - charakterystyczna średnia prędkość przepływu, L - charakterystyczna długość szczeliny wzdłuż przepływu), różnice w warunkach fizycznych wzdłuż szczeliny są małe, co zmniejsza błędy kształtu wywołane zmianą, np. konduktywności elektrycznej elektrolitu wskutek nagrzewania i wydzielania się gazu w szczelinie,
- b. przy zastosowaniu krótkich impulsów prądowych (poniżej 2 ms) występuje większa lokalizacja roztwarzania anodowego, wyrażająca się intensywnym roztwarzaniem tylko tych miejsc powierzchni anody, które znajdują się najbliżej elektrody roboczej - katody, a zanika roztwarzanie przy odległościach większych,
- c. w PECM wzrasta niezawodność procesu obróbki, co wiąże się ze znacznie łatwiejszym zapobieganiem stanom awaryjnym, m.in. zwarciom elektrycznym,
- d. istotne znaczenie techniczne PECM wiąże się również ze znacznie szerszymi możliwościami zastosowania przy procesie impulsowym sterowania adaptacyjnego z wykorzystaniem mikroprocesorów, prowadzącymi do istotnego podwyższenia wskaźników użytkowych obróbki.

Negatywne cechy PECM wiążą się przede wszystkim z obniżeniem wydajności obróbki, wywołanym przerwami w usuwaniu nadmiaru. W wyniku, prędkość drażenia obniża się do wartości $0.1 \div 0.2$ mm/min. Jednym ze sposobów redukcji wymienionej wady jest prowadzenie drażenia elektrochemicznego w dwóch zabiegach:

- I. obróbka zgrubna procesem ciągłym ECM z możliwie maksymalną prędkością V_f , z pozostawieniem nadmiaru niezbędnego dla obróbki wykańczającej,
- II. impulsowa obróbka wykańczająca, zapewniająca uzyskanie żądanej dokładności po usunięciu pozostawionego nadmiaru.

Mówi się tu o zabiegach, gdyż w przypadku zastosowania obrabiarki elektrochemicznej impulsowej, możliwe jest również prowadzenie na niej procesu prądem ciągłym. Obróbka może się więc odbywać przy jednym zamocowaniu przedmiotu i tą samą elektrodą roboczą.

Dotychczasowe badania teoretyczne i doświadczalne wyjaśniły szereg zachodzących w procesie zjawisk oraz stworzyły podstawy do wstępnego wyboru parametrów obróbki oraz prognozowania przebiegu procesu i jego wyników [1÷6]. Jednocześnie można stwierdzić, że istniejący poziom wiedzy o procesie impulsowej ECM, jest w dalszym ciągu niezadowalający, zwłaszcza w stosunku do aktualnych wymagań odnośnie dokładności obróbki.

Konieczne z powyższego względu zwiększenie efektywności PECM, jest bezpośrednio związane z dalszym rozszerzeniem zastosowań wspomagania komputerowego przy projektowaniu PECM w oparciu o modelowanie matematyczne procesu.

W referacie przedstawiono model matematyczny opisujący powiązania kształtów elektrod: narzędzia i przedmiotu obrabianego. Zagadnienie powyższe wiąże się bezpośrednio z wyznaczaniem kształtu elektrody roboczej dla PECM.

2. Model matematyczny procesu "idealnego" PECM

Ze względu na złożoność matematycznego opisu procesu PECM w jego najbardziej ogólnej postaci, celem jest tworzenie prostszych modeli cząstkowych, uwzględniających jedynie najważniejsze czynniki występujące w danych warunkach realizacji procesu, a więc opisujących proces w określonym przedziale czasów trwania impulsów oraz w określonych warunkach wyjściowych PECM. Jednocześnie na tyle adekwatnych ze względu na wybrane charakterystyki procesu rzeczywistego, że pozwalają na uzyskiwanie istotnych wyników dla potrzeb teorii i praktyki PECM.

Podobnie jak w przypadku procesu ciągłego ECM, szereg ważnych zależności dla PECM można uzyskać z tzw. modelu procesu idealnego obróbki elektrochemicznej, w którym zakłada się:

- stałość konduktywności elektrycznej ośrodka w czasie i przestrzeni, $\kappa = const$,
- nie uwzględnia się właściwości procesów elektrodowych i zakłada się stałość sumarycznej polaryzacji elektrod E oraz wydajności prądowej procesu roztwarzania η_A (lub inaczej współczynnika obrabialności elektrochemicznej: $k_v = const$).

Przyjęte założenia pozwalają ekstrapolować prawo Ohma w postaci różniczkowej na obszary bezpośrednio przylegające do powierzchni elektrod, które na mocy założenia $E_{a,k} = const$, stają się powierzchniami ekwipotencjalnymi.

Uwzględniając ponadto, że PECM jest obróbką wykańczającą prowadzoną przy grubościach szczelin poniżej 0.2 mm (często 0.1 mm), rozszerzają możliwość linearyzacji potencjału elektrycznego wzdłuż grubości szczeliny. Ułatwia to wyznaczenie gradientu potencjału na anodzie z dostateczną dokładnością, a mianowicie z zależności:

$$|\text{grad } u| = \frac{U - E}{S}, \quad (1)$$

gdzie: S - grubość szczeliny mierzona w kierunku normalnym do ER

$E = E_A - E_K$ - sumaryczna polaryzacja elektrod.

Dla zwiększenia dokładności obliczeń, na podstawie tego modelu, należy starać się stosować uśrednione wartości κ , E , k_v w czasie i przestrzeni przy danych warunkach procesu.

Podstawowym zastosowaniem modelu procesu idealnego jest wyznaczenie powiązań między kształtami elektrod: anody (PO) i katody (ER) przy danej kinematyce ruchu względnego elektrod.

Powyższe powiązania wyznacza się dla stanu quasiustalonego, zdefiniowanego dla procesu przerywanego ECM, w pracy [1].

Jak wykazuje symulacja komputerowa procesu przerywanego ECM, po pewnym czasie występuje tzw. stan quasiustalony, w którym zmiany stanu ośrodka w szczelinie oraz jej grubość stają się periodyczne w czasie.

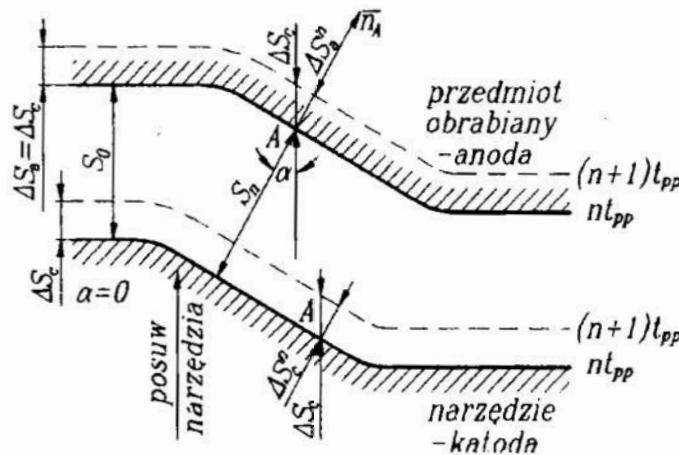
W stanie quasiustalonym, funkcja opisująca kształt przedmiotu obrabianego $z = F(x, y, t)$, jest funkcją okresową o okresie t_{pp} (rys. 1), co można zapisać:

$$F(x, y, t) = F(x, y, t + nt_{pp}) \quad \text{gdzie } n = \dots -2, -1, 0, 1, 2 \dots \quad (2)$$

W równaniu tym, (x, y, z) jest układem współrzędnych związanym z poruszającą się elektrodą-narzędziem. Niezbędnymi warunkami osiągnięcia tego stanu są wystarczająco długi czas obróbki i możliwość powtarzalnego podawania parametrów procesu takich jak kształt i parametry impulsu roboczego, parametry hydrodynamiczne i kinematyczne oraz własności elektrolitu. Aby określić geometrię przedmiotu obrabianego względem zadanego kształtu elektrody roboczej w stanie quasiustalonym, należy rozważyć początkowe i końcowe położenia elektrod. Jak pokazano na rys. 2, w czasie t_{pp} , katoda (elektroda robocza) przemieszcza się skokowo o ΔS_c , podczas gdy powierzchnia anody (przedmiot obrabiany) pod wpływem roztwarzania przesuwa się o ΔS_a w kierunku zgodnym z kierunkiem przemieszczania katody. W stanie quasiustalonym, grubość warstwy roztworzonej w rozpatrywanym punkcie A, będzie równa rzutowi ΔS_c na kierunek normalny do powierzchni anody:

$$\Delta S_{n(A)} = \Delta S_c \cos \alpha_A, \quad (3)$$

gdzie: α_A - kąt pomiędzy normalną do anody w punkcie A a kierunkiem dosuwu katody.



Rys. 2. Położenie elektrod w stanie quasi-ustalonym: n - numer impulsu, t_{pp} - okres impulsu, ΔS_c - przemieszczenie katody w kierunku posuwu, ΔS_c^n - przesunięcie powierzchni katody w kierunku normalnym, ΔS_a^n - grubość naddatku roztworzonego w czasie impulsu, S_0 - grubość szczeliny na początku impulsu

W czasie t_{pp} , grubość warstwy usuniętego elektrochemicznie materiału w punkcie A na anodzie, na podstawie I prawa Faradaya opisana jest wzorem:

$$\Delta S_{n(A)} = \frac{k}{\rho_m} \int_0^{t_p} \eta_A(t) i_A(t) dt, \quad (4)$$

lub:

$$\Delta S_{n(A)} = (k/\rho_m) q_A,$$

gdzie:

$$q_A = \int_0^{t_p} \eta_A(t) i_A(t) dt \quad (5)$$

- ładunek elektryczny przenoszony jonami materiału obrabianego w czasie t_p

gdzie: k - równoważnik elektrochemiczny,

ρ_m - gęstość właściwa materiału anody,

$\eta_A(t)$ - wydajność prądowa roztwarzania anodowego,

$i_A(t)$ - gęstość prądu na anodzie,

Wstawiając (3) do (4), otrzymujemy następujący warunek dla stanu quasi-ustalonego:

$$\int_0^{t_p} \eta_A(t) i_A(t) dt = (\Delta S_c / k_V) \cos \alpha_A, \quad (6)$$

lub:

$$q_A = q_0 \cos \alpha_A, \quad (7)$$

gdzie: $k_V = k / \rho_m$ i $q_0 = \Delta S_c / k_V$.

Dla porównania, podobna równość może być wyprowadzona dla obróbki ciągłej ECM w stanie ustalonym:

$$\eta_A i_A = (V_f / k_V) \cos \alpha_A. \quad (8)$$

Można wykazać, że obie równości niosą ze sobą te same informacje jeśli równość (6) jest przekształcona z użyciem dwóch średnich wartości:

$$\overline{\eta i_A} = \frac{1}{t_{pp}} \int_0^{t_p} \eta_A(t) i_A(t) dt, \quad (9)$$

oraz:

$$V_f = \frac{\Delta S_c}{t_{pp}}. \quad (10)$$

Równanie (6) przybierze wówczas postać analogiczną do (8):

$$\overline{\eta i_A} = \frac{V_f}{k_V} \cos \alpha_A, \quad (11)$$

lub:

$$V_{\kappa(A)} = V_f \cos \alpha_A, \quad (12)$$

gdzie:

$$V_{\kappa(A)} = \overline{\eta i_A} k_V \quad (13)$$

jest średnią prędkością roztwarzania w czasie t_{pp} .

Równości (7) i (12) mają znaczenie ogólne (tj. są słuszne nie tylko dla procesu idealnego). Mogą być one użyte do opisu wymiany ładunku elektrycznego i średniej ilości lokalnie usuniętego materiału w stanie quasiustalonym PECM oraz do obliczania rozkładu szczeliny w celu projektowania elektrody roboczej przez obliczenia lokalnych ubytków materiału dla zadanych parametrów wejściowych.

Rozpatrzmy bardziej szczegółowo często występujący w praktyce PECM ruchu przypadek ruchu ER, w którym w czasie impulsu napięcia roboczego jest ona nieruchoma (rys. 1 - odcinek BC), natomiast ruchy pomocnicze, np. cofanie dla

gdzie: k - równoważnik elektrochemiczny,

ρ_m - gęstość właściwa materiału anody,

$\eta_A(t)$ - wydajność prądowa roztwarzania anodowego,

$i_A(t)$ - gęstość prądu na anodzie,

Wstawiając (3) do (4), otrzymujemy następujący warunek dla stanu quasi-ustalonego:

$$\int_0^{t_p} \eta_A(t) i_A(t) dt = (\Delta S_c / k_V) \cos \alpha_A, \quad (6)$$

lub:

$$q_A = q_0 \cos \alpha_A, \quad (7)$$

gdzie: $k_V = k / \rho_m$ i $q_0 = \Delta S_c / k_V$.

Dla porównania, podobna równość może być wyprowadzona dla obróbki ciągłej ECM w stanie ustalonym:

$$\eta_A i_A = (V_f / k_V) \cos \alpha_A. \quad (8)$$

Można wykazać, że obie równości niosą ze sobą te same informacje jeśli równość (6) jest przekształcona z użyciem dwóch średnich wartości:

$$\overline{\eta i_A} = \frac{1}{t_{pp}} \int_0^{t_p} \eta_A(t) i_A(t) dt, \quad (9)$$

oraz:

$$V_f = \frac{\Delta S_c}{t_{pp}}. \quad (10)$$

Równanie (6) przybierze wówczas postać analogiczną do (8):

$$\overline{\eta i_A} = \frac{V_f}{k_V} \cos \alpha_A, \quad (11)$$

lub:

$$V_{\kappa(A)} = V_f \cos \alpha_A, \quad (12)$$

gdzie:

$$V_{\kappa(A)} = \overline{\eta i_A} k_V \quad (13)$$

jest średnią prędkością roztwarzania w czasie t_{pp} .

Równości (7) i (12) mają znaczenie ogólne (tj. są słuszne nie tylko dla procesu idealnego). Mogą być one użyte do opisu wymiany ładunku elektrycznego i średniej ilości lokalnie usuniętego materiału w stanie quasiustalonym PECM oraz do obliczania rozkładu szczeliny w celu projektowania elektrody roboczej przez obliczenia lokalnych ubytków materiału dla zadanych parametrów wejściowych.

Rozpatrzmy bardziej szczegółowo często występujący w praktyce PECM ruchu przypadek ruchu ER, w którym w czasie impulsu napięcia roboczego jest ona nieruchoma (rys. 1 - odcinek BC), natomiast ruchy pomocnicze, np. cofanie dla

po przekształceniach uzyskujemy zależność opisującą rozkład szczelin na początku impulsu:

$$S_n = k_w \frac{S_f}{\cos \alpha} - \frac{V_f t_{pp}}{2} \cos \alpha. \quad (21)$$

Ostatni człon po prawej stronie może być pominięty, gdyż dla warunków PECM: $V_f < 1/60 \text{ mm/s}$, $t_{pp} < 0.01 \text{ s}$. Mając zadany kształt PO oraz rozkład grubości szczelin z równania (20), możemy wyznaczyć współrzędne poszukiwanej ER bez większych trudności.

Wprowadzając wartość korekcji ER jako różnicę $e = S_n - S_n(\alpha=0)$, na podstawie zależności (20) otrzymujemy dla PECM:

$$e = 2k_w S_f \frac{\sin^2 \frac{\alpha}{2}}{\cos \alpha}. \quad (22)$$

Dla ECM $e = 2S_f \sin^2(\alpha/2)/\cos \alpha$ stąd można wnioskować, że zastosowanie PECM prowadzi do $1/k_w$ krotnego zmniejszenia korekcji ER. Najczęściej stosowane współczynniki wypełnienia zawierają się w granicach $k_w = 0.1 \div 0.2$.

Analiza innych wariantów ruchów ER w PECM, wykazuje, że również i dla tych przypadków uzyskuje się zależności podobne do otrzymanej. Różnica jest jedynie w postaci matematycznej ostatniego członu równania (20), który jest najczęściej znacznie mniejszy od pierwszego. Jest to uzasadnione poniższą analizą.

Powracamy w tym celu do równania stanu quasiustalonego PECM (11):

$$\overline{\eta i_A} = V_f \cos \alpha.$$

Przy stosowaniu krótkich impulsów poniżej 2 ms grubość szczeliny zmniejszy się w tym czasie nieznacznie, stąd dla procesu "idealnego" PECM średnia gęstość prądu wyrazi się w przybliżeniu:

$$i_A = \kappa \cdot \frac{U - E}{S_n}. \quad (23)$$

Po podstawieniu zależności (23) do równania (11) oraz wprowadzeniu wielkości S_f zdefiniowanej równaniem (20) otrzymamy po przekształceniu poszukiwany, przybliżony rozkład szczelin:

$$S_n \cong k_w \cdot \frac{S_f}{\cos \alpha}. \quad (24)$$

Wniosek odnośnie korekcji otrzymany dla rozpatrywanego szczegółowo przypadku, może więc być rozszerzony na inne przypadki PECM.

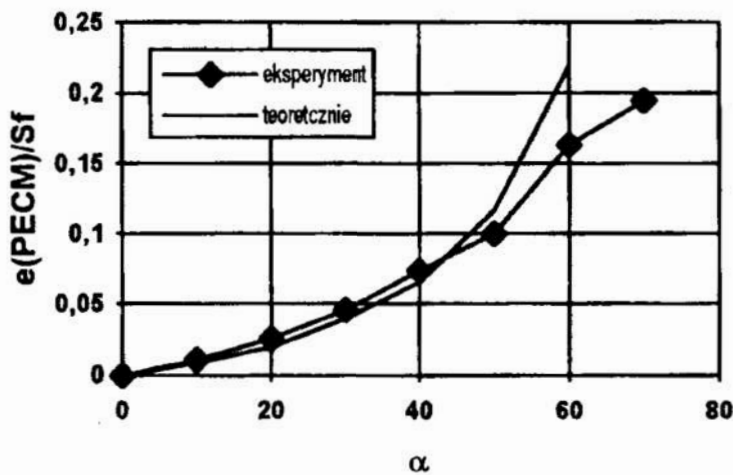
Oczywiście powyższy model procesu "idealnego" PECM nie uwzględnia z założeń, tak ważnych dla prowadzenia PECM, zjawisk jak nagrzewanie się elektrolitu, wydzielanie gazu i szereg innych związanych z procesami elektrodowymi. Uwzględnienie powyższych warunków w uściślonych modelach [1÷6] jest niezbędnym warunkiem przy doborze parametrów obróbki.

Dla wyznaczenia stopnia rozbieżności uzyskanych powyższych rozkładów grubości szczelin dla PECM, przeprowadzone zostały badania doświadczalne, w których wyznaczono względną wielkość korekcji:

$$\bar{e} = \frac{e}{S_f} = \frac{S_n - S_n(\alpha = 0)}{S_f}$$

Badania przeprowadzono na stanowisku badawczym IDEL 1 posiadającym programowany skokowy napęd ER (o skoku minimalnym $\Delta S_0 = 5.5 \mu\text{m}$) i regulowaną częstotliwością (od 0.5 do 25 Hz). Zasilanie prądem roboczym odbywa się z zasilacza prądu stałego przez tyrystorowy łącznik, który wytwarza impulsy o regulowanych parametrach: $t_p = 0.5 \text{ ms} \div 10 \text{ s}$, $t_o = 4 \text{ ms} \div 50 \text{ s}$, $U \leq 24 \text{ V}$, maksymalna amplituda natężenia prądu $I_m = 1000 \text{ A}$. Ciśnienie elektrolitu na wlocie do komory roboczej wynosi do 3 MPa [7].

Obróbkę prowadzono przy użyciu ER, której powierzchnią czynną była powierzchnia walca o $\phi 20 \text{ mm}$. Próbkę wykonano ze stali WNL zahartowanej do 55 HRC. Elektrolitem był roztwór wodny 15 % NaCl. Obróbka odbywała się z zastosowaniem impulsów $t_p = 2 \text{ ms}$, $t_{pp} = 10 \text{ ms}$, $U = 12 \text{ V}$, przy skoku $\Delta S_c = 5.5 \mu\text{m}$. Wyniki teoretyczne i doświadczalne przedstawione są na rys. 3. Jak wynika z przebiegu krzywych, zadowalającą zgodność wyników uzyskano dla kątów $\alpha < 50^\circ$. Dla wyższych wielkości kątów pochylenia różnice w przebiegach, a więc błąd modelu rośnie.



Rys. 3. Porównanie wartości korekcji teoretycznych z doświadczalnymi. Materiał obrabiany: stal WNL, elektrolit: 10 % NaCl, $t_p = 2 \text{ ms}$, $k_w = 0.2$, $U = 12 \text{ V}$, $p_{in} = 0.6 \text{ MPa}$

Główną przyczyną rozbieżności jest naruszenie założeń w stosunku do modelu oraz wzrost (przy $\alpha > 50^\circ$), różnicy kierunków normalnych do powierzchni elektrod, a mianowicie normalnej do ER „ \vec{n}_E ”, wzdłuż której mierzy się S i normalnej do anody w punkcie przecięcia się \vec{n}_E z anodą.

Pracę częściowo wykonano w ramach realizacji grantu PB 0256/P4/93/05.

Literatura

- [1] Kozak J.: *Kształtowanie powierzchni obróbką elektrochemiczną - bezstykową (ECM)*. Prace Naukowe PW, Mechanika Nr 41, 1976.
- [2] Petrov J.N. i inni: *Problemy povysenija točnosti elektrochimiceskogo formoobrazowanija*. Stinca, 1977.
- [3] Kozak J., Łubkowski K.: *Basic Investigation of Characteristic in PECM*. Proceed 20th MTDR Conf., Birmingham, 1979.

- [4] Kozak J., Lubkowski K., Perończyk J.: *Accuracy Problem of the Pulse Electrochemical Machining*. Proceed 22th MTDR Conf., Manchester, 1981.
- [5] Datta M., Landolt D.: *Electrochemical Dissolution under Cotinuous and Pulsed Current Conditions*. Proceed 6th ISEM, Kraków, 1980.
- [6] Kozak J., Lubkowski K., Mahboud A.: *Characteristic of PECM*. Proceed. of ASME, WAM-PED-vol 34, Chicago, 1989.
- [7] Kozak J., Dąbrowski L., Lubkowski K.: *Impulsowa obróbka elektrochemiczna*. Mat. konferencji EM'82, Bydgoszcz, 1982.

MATHEMATICAL MODELING OF THE PULSE ELECTROCHEMICAL MACHINING

Mathematical model of the pulse electrochemical machining and some problems of correction of the tool-electrode are described in this paper. Results of experimental verification is also presented.

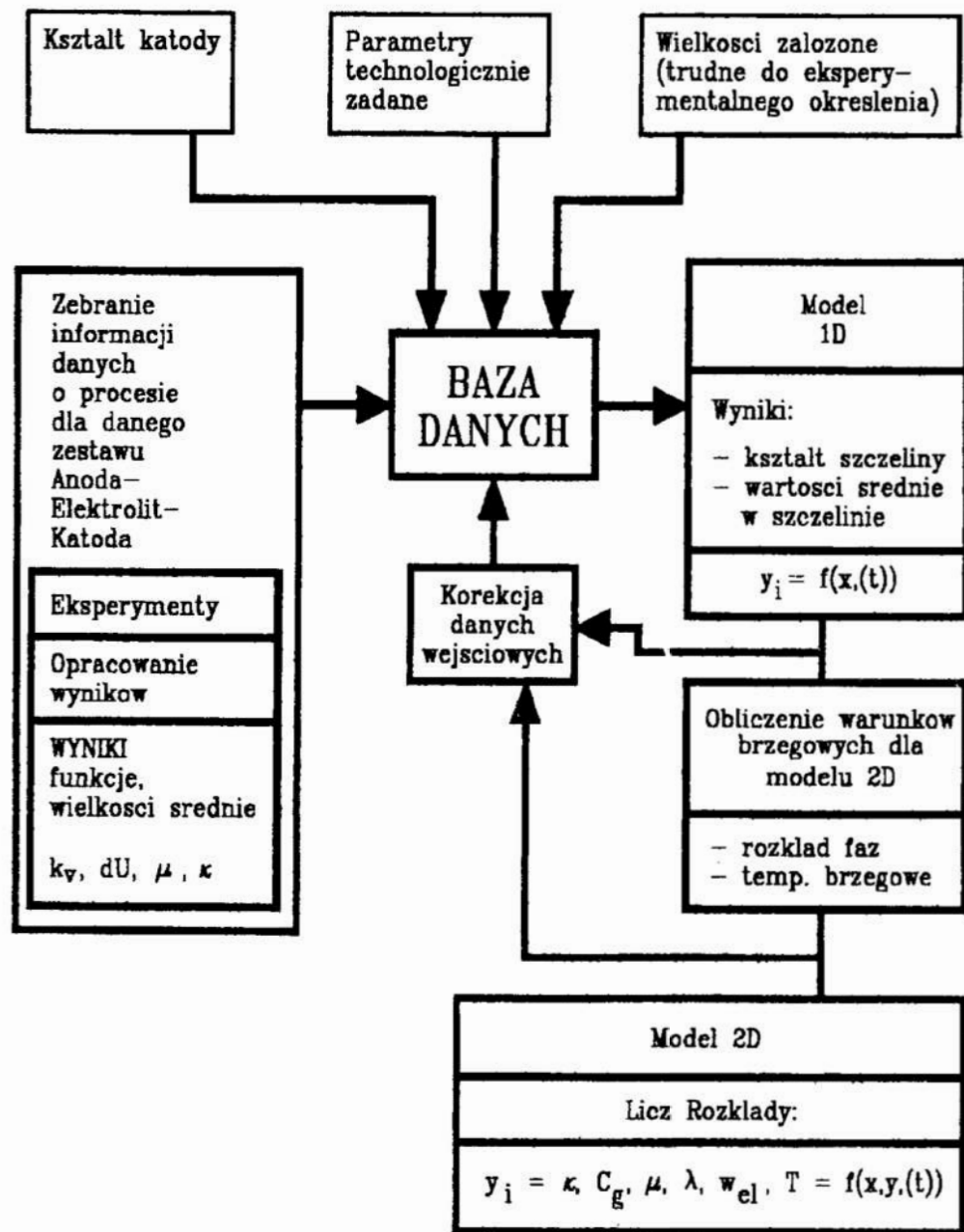
SYMULACJA KOMPUTEROWA DRAŻENIA ELEKTROCHEMICZNEGO ELEKTRODĄ KSZTAŁTOWĄ

W artykule przedstawiono główne zagadnienia symulacji komputerowej procesu obróbki elektrochemicznej. Przedstawiono symulację drażenia kształtowego. Podano model fizyczny i matematyczny, sposób rozwiązania i wybrane wyniki.

Wprowadzenie

Obróbka elektrochemiczna ECM wykorzystuje procesy elektrochemiczne towarzyszące przepływowi prądu elektrycznego w obwodzie składającym się z: elektrody roboczej (katody), przedmiotu obrabianego (anody) oraz zawartego między nimi elektrolitu. W wyniku tych procesów następuje roztwarzanie powierzchni anody czyli przechodzenie atomów metalu anody do roztworu, tworzenie tlenków i wodorotlenków metalu, które usuwane są przez przepływający elektrolit. Obróbka elektrochemiczna stosowana jest szeroko w wybranych gałęziach przemysłu, głównie do obróbki trudnoskrawalnych metali i stopów. Doświadczenia zebrane przez wiele lat stosowania obróbki elektrochemicznej wykazały, że proces ten stwarza szereg problemów nie występujących w obróbkach klasycznych a związanych ze strukturą wewnętrzną procesu elektrochemicznego roztwarzania przy wysokich gęstościach prądu. Problemy te obejmują: procesy elektrodowe, procesy wymiany masy, energii i przenoszenia ładunku elektrycznego, hydrodynamikę przepływu oraz zmienne własności ośrodka w szczeliny międzyelektrodowej. Wielowymiarowy, dynamiczny proces ECM wymaga zastosowania komputerowego systemu projektowania i sterowania obróbką. Systemy takie są realizowane, choć informacje o nich są skąpe [6,7]. Wstępem do opracowania nowoczesnej technologii ECM musi być poznanie natury zjawisk fizycznych zachodzących podczas obróbki oraz ograniczeń tkwiących w samym procesie roztwarzania elektrochemicznego hamujących lub ograniczających intensyfikację i podwyższanie efektów uzyskiwanych w wyniku tej obróbki. Dokładne ilościowe określenie tych ograniczeń umożliwi skonstruowanie modelu fizycznego i następnie matematycznego umożliwiając symulację [5,8] różnych wariantów obróbki, dobór optymalnych parametrów zapewniający wysokie wskaźniki użytkowe i ekonomiczne a równocześnie nie doprowadzające procesu ECM do tzw. stanów krytycznych w których następuje przerwanie obróbki (zwykle przez wyładowanie iskrowe i zwarcie) i uszkodzenie elektrod. W niniejszej pracy przedstawiono przykład symulacji jednowymiarowej drażenia elektrodą kształtową. Pełny schemat całego układu symulacji komputerowej procesu ECM wraz z koniecznymi badaniami doświadczalnymi pokazano na rys.1. W skład układu wchodzi: baza danych fizycznych i technologicznych oraz modele (jednowymiarowy i dwuwymiarowy) opisujące proces obróbki elektrochemicznej.

^a Instytut Technologii Mechanicznej PW, Al. Niepodległości 222 Warszawa



Rys.1. Schemat układu symulacji ECM

1. Badania doświadczalne konieczne dla symulacji

Aby wiarygodnie symulować proces ECM należy poznać wzajemne powiązania układu katoda-elektrolit-anoda ponieważ odgrywają one decydującą rolę w procesie obróbki. Powiązania te, to procesy elektrodowe (reakcje zachodzące na elektrodach), które łącznie stanowią o mechanizmie roztwarzania. Obejmują one:

- zjawiska zachodzące na anodzie; roztwarzanie powierzchni, wytwarzanie warstw tlenkowych i innych, wydzielanie gazów,
- zjawiska zachodzące na katodzie; głównie wydzielanie gazów i osadzanie produktów stałych,
- spadki potencjałów w warstwach przyelektrodowych,
- wydzielanie ciepła w warstwach tlenkowych i gazowych.

Wymienione powyżej, procesy wpływają na: wydajność i energochłonność obróbki, stan warstwy wierzchniej oraz dokładność wymiarowo-kształtową obrabianej powierzchni. Efekty tych procesów opisuje się obecnie dwiema charakterystykami podstawowymi: współczynnikiem obrabialności k_v lub sprawnością prądową η , oraz polaryzacją elektrod φ_A i φ_K (krzywe polaryzacyjne) w funkcji głównych parametrów obróbki dla konkretnego zestawu metal-elektrolit. Złożoność zjawisk zachodzących w czasie ECM powoduje, że na podstawie obecnego stanu wiedzy, niemożliwe jest prognozowanie cech danego układu metal-elektrolit, jak i teoretyczne wyznaczenie krzywych polaryzacyjnych $j(\varphi)$ i zależności $k_v(j)$ dla różnych warunków ECM. Powyższe charakterystyki można wyznaczyć jedynie metodami doświadczalnymi, którym poświęcone jest wiele prac, np [2]. Publikowane wyniki różnych autorów znacznie się różnią między sobą, w tej sytuacji konieczne jest więc wykonanie własnych badań doświadczalnych dla konkretnych zestawów metal-elektrolit. Próby modelowania komputerowego krzywych polaryzacyjnych przedstawił Devereux [3, 4] ale dla niskich gęstości prądu. Istotną sprawą jest również określenie rodzaju przepływu (laminarny czy turbulentny) oraz określenie strat hydraulicznych. Wszystkie te dane wprowadza się do bazy danych lub bezpośrednio do równań modelu jednowymiarowego, który jest podstawowym członem symulacji.

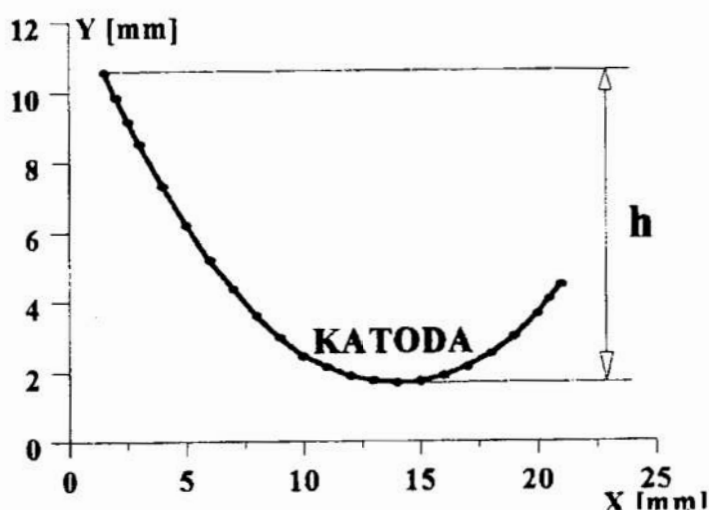
2. Symulacja jednowymiarowa procesu drażenia ECM

Podstawowym członem symulacji jest symulacja jednowymiarowa, w której na podstawie modelu jednowymiarowego 1D dla zadanego kształtu katody i parametrów obróbki obliczane są: rozkłady średnich parametrów w szczelinie wzdłuż jej długości oraz kształt powierzchni obrabianej. Poniżej przedstawiona jest symulacja drażenia ECM elektrodą kształtową płaską w tzw. stanie ustalonym, w którym prędkość roztwarzania anody jest równa prędkości posuwu katody.

2.1. Założenia modelu 1D

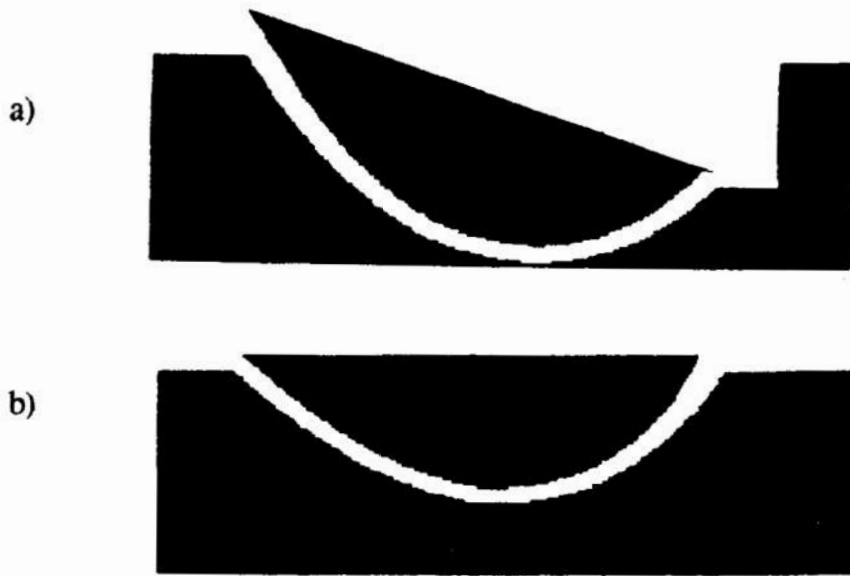
Podstawą opisu matematycznego jest przyjęty model fizyczny, w którym założono:

1-drażenie elektrodą kształtową płaską (kształt elektrody zadany jest punktami o współrzędnych; x_i, y_i , rys.2), ze stałą prędkością drażenia v_f w dwóch układach;

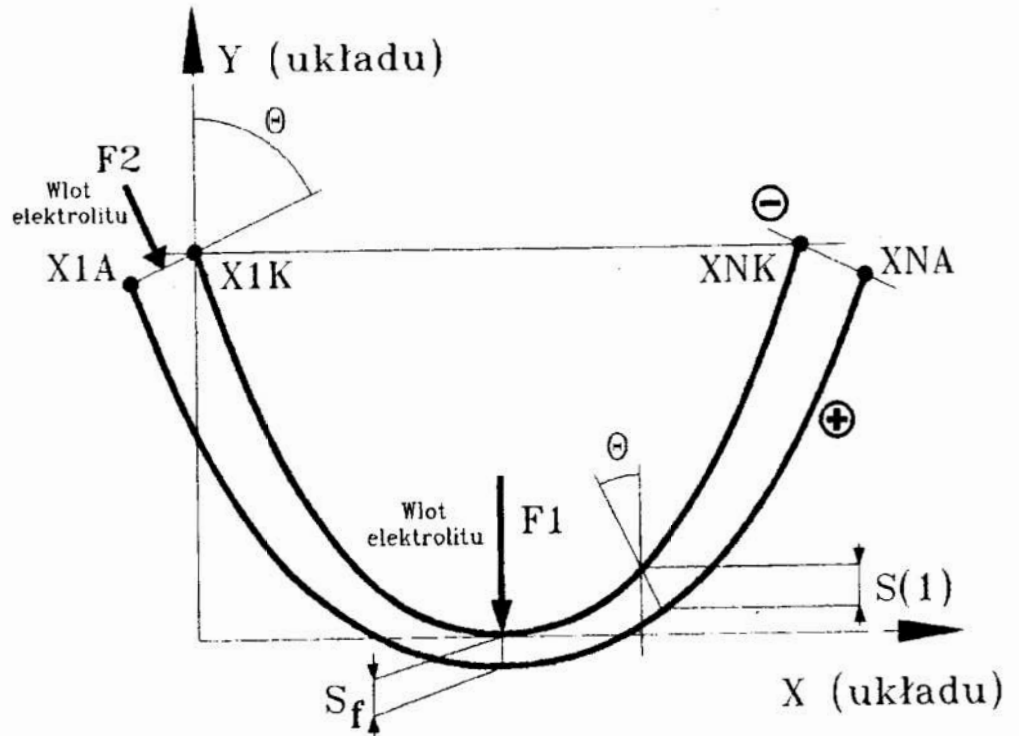


Rys.2 Rysunek profilu katody, h - wysokość profilu.

w układzie rysunkowym katody i w układzie obróconym korzystnym dla procesu ECM (rys.3). Założono dwa przypadki rozmieszczenia wlotu elektrolitu pokazane na przyjętym schemacie obróbki (rys.4).



Rys.3. Animacja drążenia elektrochemicznego (grubość szczeliny powiększona trzykrotnie); a- drążenie w układzie rysunkowym katody, b- układ obrócony tak aby wysokość profilu była minimalna.



Rys.4. Schemat układu obróbki (układ obrócony)

2- turbulentny, homogeniczny przepływ dwufazowy (ciecz-gaz), pomijane są zanieczyszczenia elektrolitu produktami roztwarzania oraz wpływ sił masowych, faza ciekła i gazowa ma te same prędkości, ciśnienia i temperatury, stałe w przekroju poprzecznym szczeliny.

3-pomija się odprowadzanie ciepła przez elektrody, rozszerzalność cieplną ośrodka oraz niejednorodności powierzchni z punktu widzenia procesów elektrochemicznych.

4-pomija się opisy procesów dyfuzji i kinetyki elektrodowej, ich skutki uwzględnia się przy pomocy funkcji opisujących zmianę współczynnika obrabialności k_v i spadku potencjałów w warstwach przyelektrodowych $\Delta U = \varphi_A + \varphi_K$ w funkcji głównych parametrów obróbki (w niniejszym opracowaniu przyjęto ΔU stałe).

5-prawo Ohma ekstrapoluje się na całą grubość szczeliny.

6-własności ośrodka dwufazowego określa się następująco:

gęstość $\rho = (1 - Cg) \rho_e$, gdzie: ρ_e - gęstość elektrolitu,

Cg - koncentracja objętościowa fazy gazowej,

konduktywność; $\kappa = \kappa(+ a\Delta T) (1 - Cg)^{1.5}$

gdzie: α -temperaturowy współczynnik konduktywności,

κ -konduktywność elektrolitu przy $T=20\text{ C}$,

lepkość dynamiczna; $\mu = \mu(1 + m Cg) e^{-b\Delta T}$

przyjęto; $m = 5.5, b=0.019$

Na podstawie powyższych założeń opisano proces następującym układem równań:

Lokalna grubość szczeliny międzyelektrodowej w stanie ustalonym zgodnie z pracą [9];

$$s \left(\frac{ds}{dx} \sin \vartheta + \cos \vartheta \right) = \frac{k_v \kappa (U - \Delta U)}{v_f} \quad (1)$$

Należy podkreślić, że powyższa zależność zdecydowanie lepiej "wiąże" kształt katody z kształtem anody w porównaniu z powszechnie stosowanym, np. w pracy [10] "prawem cosinus".

Lokalna gęstość prądu; $j = \kappa (U - \Delta U) / s$ (2)

Równanie ciągłości; $w s (1 - Cg) \rho_e = const$ (3)

gdzie: w - prędkość przepływu elektrolitu

Rozkład ciśnień w szczelinie;

$$\frac{dp}{dx} = - \frac{0.316 \rho_e w^2}{4s Re^{0.25}} - \frac{d(\rho_e w^2 s)}{s dx}, \quad (4)$$

gdzie: $Re = \frac{2sw}{\mu}$

Rozkład temperatur;

$$\frac{dT}{dx} = \frac{j^2}{\kappa \rho_e w s (1 - Cg) C} \quad (5)$$

Rozkład koncentracji fazy gazowej;

$$\frac{dCg}{dx} = \frac{\eta_H k_H j R_H T}{w s p} \quad (6)$$

gdzie: η_H - sprawność prądowa wodoru,

k_H - równoważnik elektrochemiczny wodoru,

R_H - indywidualna stała gazowa wodoru,

Zmienność współczynnika obrabialności określona z badań własnych [5,8],

$$k_v = 3.7 * 10^{-11} \exp(29 - 6.7 \ln T - 0.04 \ln j^2 + 0.208 \ln j \ln T) \text{ [m}^3/\text{C]} \quad (7)$$

Powyższy układ równań (1÷7) wraz z zależnościami z p.6 założeń należy rozwiązać przy następujących warunkach brzegowych;

$$\begin{aligned} \text{na wlocie:} \quad p &= p_{wlot} - \zeta_1 \rho_e w^2 / 2 \\ T &= T_{wlot}, \quad Cg = 0 \end{aligned} \quad (8)$$

$$\text{na wylocie:} \quad p_k = p_{wylot} + \zeta_2 \rho_e w^2 / 2$$

gdzie: ζ_1, ζ_2 - współczynniki strat lokalnych na wlocie i wylocie.

W powyższym układzie występują nieliniowe równania różniczkowe 1 stopnia, które rozwiązywane są metodą różnic skończonych (ulepszona metoda Eulera). W równaniach (4) (5) (6) zadawane są warunki brzegowe p, Cg, T na wlocie do szczeliny. Natomiast w równaniu (4) (3) nieznanym parametrem jest prędkość przepływu elektrolitu na wlocie do szczeliny. W obliczeniach zastosowano metodę strzału jednokrotnego, w której prędkość na wlocie dobierana jest wg specjalnej procedury tak aby był spełniony warunek :

$$\sup | p(n) - p_k | < \varepsilon_p \quad (9)$$

gdzie: $p(n)$ ciśnienie na wylocie ze szczeliny obliczone z modelu,

ε_p założona dokładność w obliczeniach rozkładu ciśnień.

Układ równań modelu rozwiązywany jest metodą iteracyjną kolejnych przybliżeń, obliczenia powtarzane są aż do osiągnięcia założonej dokładności obliczeń ε_s , schemat obliczeń pokazano na rys.5.

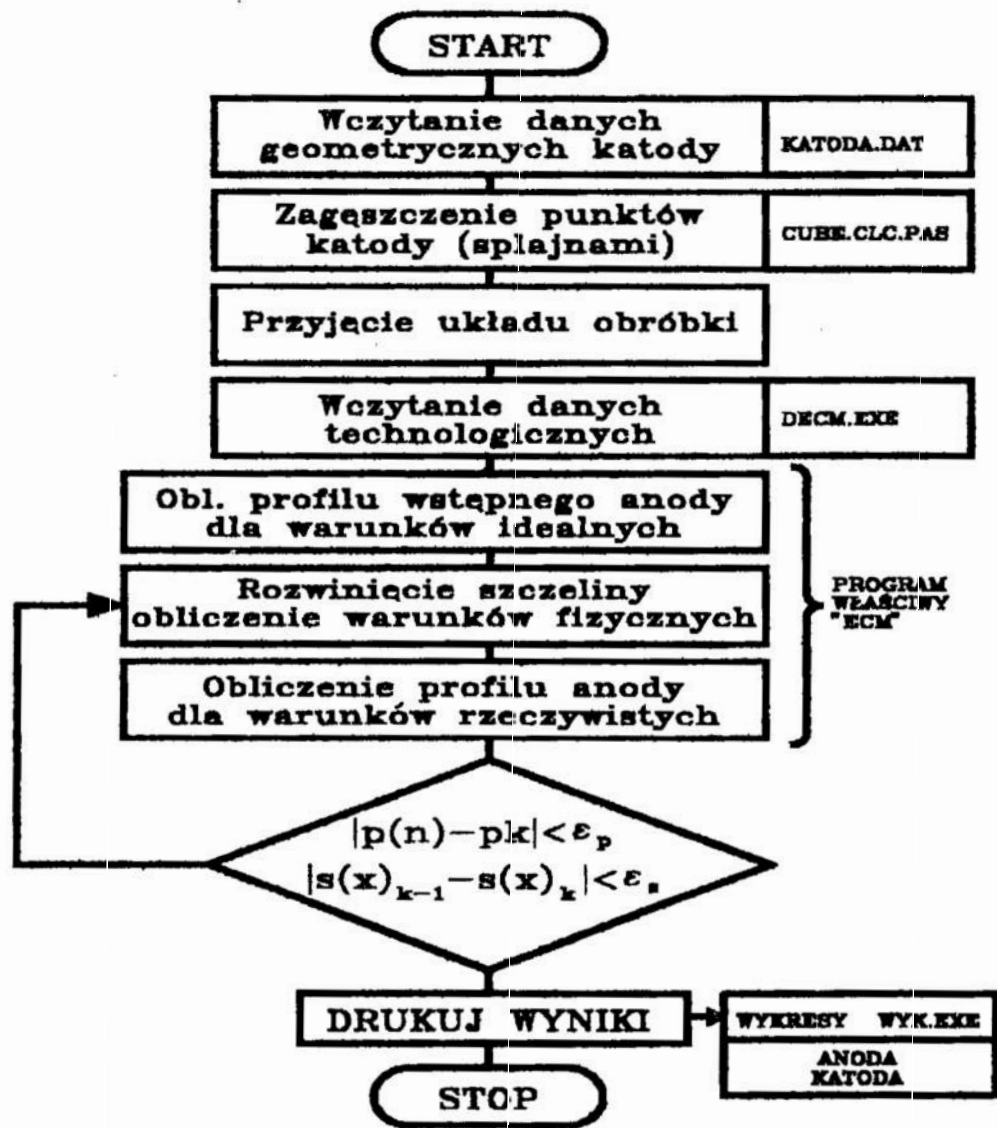
$$\sup | s(x)k-1 - s(x)k | < \varepsilon_s \quad (10)$$

gdzie: k - numer kolejnej iteracji.

Obliczenia przeprowadzono na mikrokomputerze PC przy pomocy programu komputerowego ECM. Pierwsza część programu oblicza kształt katody, "zagęszcza" punkty rysunkowe dodatkowymi punktami metodą splajnow (funkcje sklejące) [1]. Wykorzystano program Cube.clc.pas z biblioteki Numerical Methods Toolbox. Kształt katody zadany został 23 punktami (cięciwa profilu 21.7 mm), w celu zagęszczenia

profilu wprowadzono dodatkowe dwa skrajne punkty na końcach profilu podając w nich wartość pierwszej pochodnej. Pomiedzy kazdymi dwoma danymi punktami obliczane są cztery nowe punkty. Następnie po przyjęciu układu, dla zadanych parametrów obróbki obliczane są wartości parametrów średnich w szczelinie oraz kształt powierzchni obrabianej.

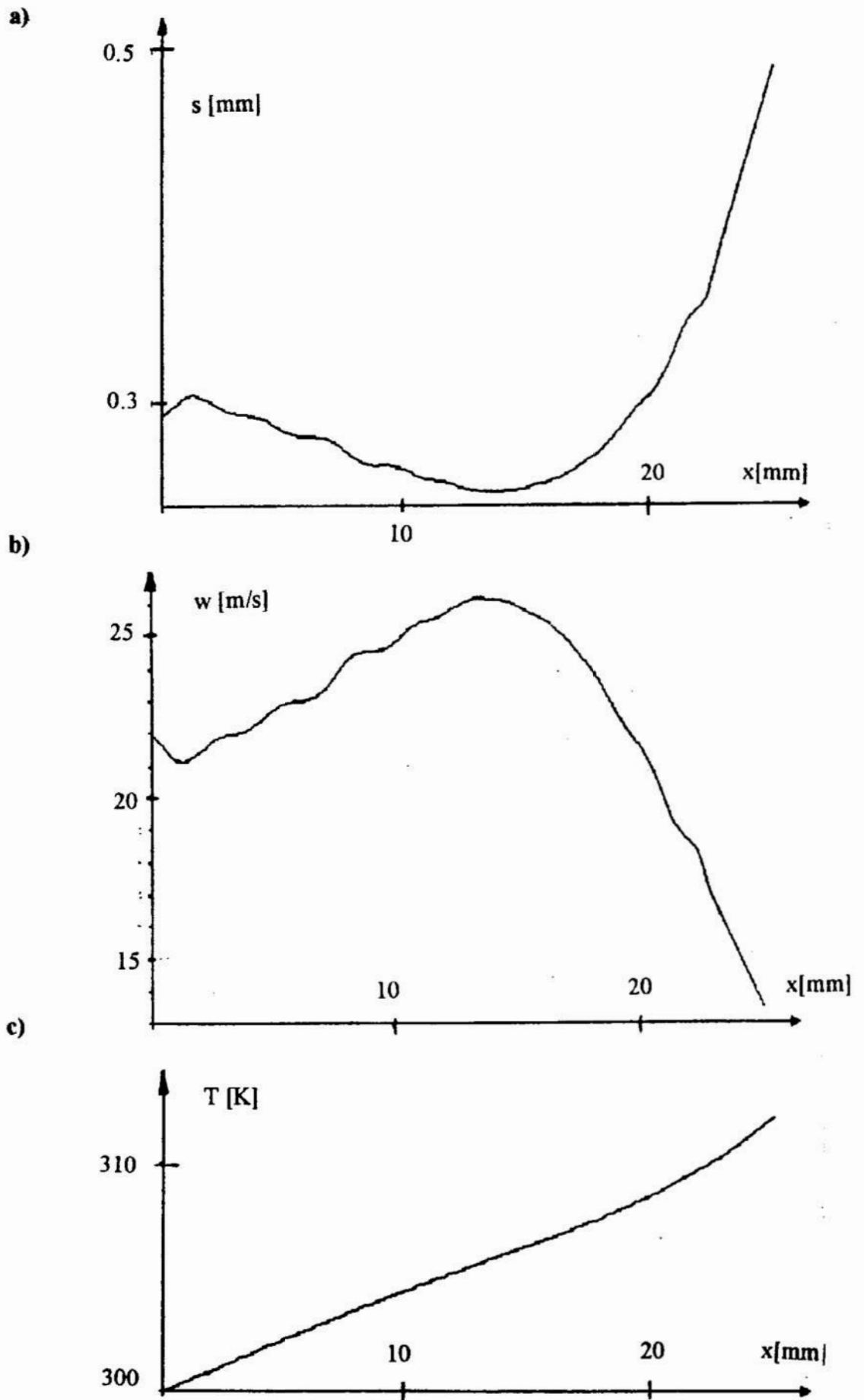
Na rys.6 i 7 przedstawiono wybrane wyniki symulacji drażenia elektrodą kształtową w układzie obróconym (wlot z lewej strony profilu), z prędkością posuwu katody 1 mm/min, w 10% wodnym roztworze NaNO₃, materiał obrabiany stal 45.



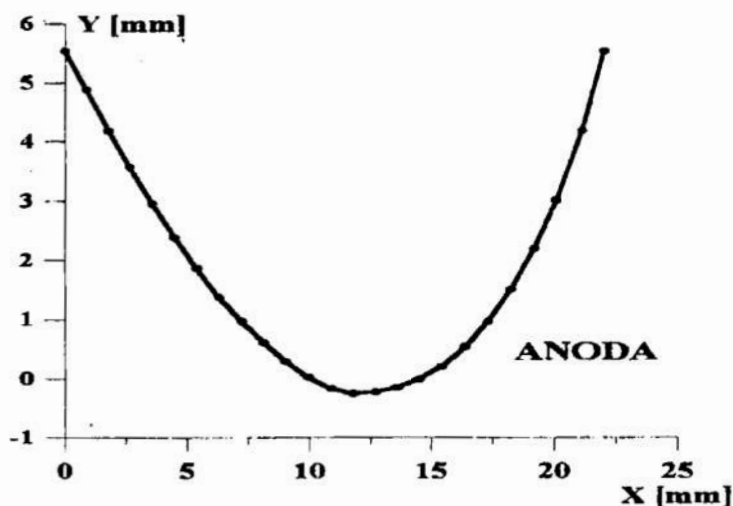
Rys. 5

Zakończenie

Powyższa symulacja może być stosowana do wstępnego określenia kształtu narzędzia (katody) do obróbki elektrochemicznej zadanego kształtu. Kształt katody w powyższej symulacji jest kształtem, który mamy otrzymać w wyniku obróbki ECM. Kształt anody uzyskany w wyniku symulacji jest więc pierwszym przybliżeniem kształtu narzędzia-katody dla zadanej geometrii układu i parametrów obróbki.



Rys.6. Rozkłady (wybranych) parametrów średnich w szczelinie.



Rys.7. Obliczony profil anody - powierzchni obrobionej.

Literatura

- [1] Cope R.J., Sawko F., Tickell R.G.: *Computer methods for civil engineers*. McGraw-Hill Book Company (UK) 1982r.
- [2] Dąbrowski L.: *Badania roztwarzania elektrochemicznego stopów metali o specjalnych właściwościach*. Sprawozdanie z Projektu Badawczego KBN 308729101, Warszawa 1993r.
- [3] Devereux O.F.: *Polarization curve-fitting by computer modelling*. *Corrosion* Vol.35, s.125÷129.
- [4] Devereux O.F.: *Computer modelling of polarization curves in passive systems*. *Corrosion* Vol.36, s.262÷263.
- [5] Kozak J., Łubkowski K.: *A computation of physical conditions and geometry of the gap in the steady state of electrochemical machining*. *Mat. 27 MATADOR Conference Manchester 1988r*, s.269÷274.
- [6] Larsson C.N.: *Adaptive control of ECM based on flow measurement*. *Mat. ISEM-6 Kraków 1980r*, s.298÷300.
- [7] Larsson C.N., Zhu S.: *Microcomputer control of an electrochemical machine*. *Mat. 22 MTDR Conference Manchester 1981r*, s.393÷399.
- [8] Łubkowski K.: *Określenie maksymalnej prędkości drążenia elektrochemicznego*. *Mat. IV Ogólnopolskiego sem. nauk. "Zastosowanie matematyki w budowie maszyn" Gorzów Wlkp. 1989r*, s.157÷169.
- [9] Thorpe J.F.: *A mathematical model of electrochemical machining process*. *Mat. CIRP Third Int. Seminar on Optimization of Manufacturing Systems. Piza 1971r*.
- [10] Tipton H.: *The dynamics of electrochemical machining*. *Mat. V Int. MTDR Conference, Birmingham 1964r*, s.509: 522.

SIMULATION OF ECM SINKING WITH THE FORM ELECTRODE

Summary

Some problems of computer simulation of ECM process are described. Simulation of ECM sinking with the form electrode is presented. The physical and the mathematical model, a method of solution and results of simulation are presented.

MODEL MATEMATYCZNY DRGAŃ POWSTAJĄCYCH PRZY SZLIFOWANIU ZĘBÓW KÓŁ ZĘBATYCH SPOSOBEM NILESA

W pracy przedstawiono model matematyczny drgań powstających podczas szlifowania zębów kół zębatych sposobem Nilesa. Wskazano na związki z trzpieniem mocującym koło zębate oraz na wrzeciono i umocowaną na nim ściernicę. Wreszcie przedstawiono model charakteryzujący wielkość drgań powstających podczas szlifowania zębów kół zębatych.

1. Wstęp

W pracy przeprowadzono rozważania dla wyodrębnionego układu P-N, modelując każdy z dwóch zespołów osobno. Analizowany układ dla przyjętego sposobu obróbki został przedstawiony w [1]. Nie uwzględniono wpływu otoczenia oraz zmiennej siły szlifowania, założono bowiem, że jej wpływ jest mały.

W pierwszym rzędzie przedstawione zostaną drgania trzpienia wraz z osadzonym na nim kołem zębatym oraz ściernicą zamocowaną na wrzecionie. Po wyznaczeniu równań opisujących ruch przedmiotu i narzędzia określona będzie wielkość drgań podczas obróbki.

2. Drgania trzpienia z osadzonym kołem zębatym

W dotychczasowej analizie drgań poprzecznych i skrętnych trzpienia upraszczano jak najbardziej schemat ujęcia ruchu. Gdy opisywano dany rodzaj drgań dopuszczano tylko jeden stopień swobody. Koło zębate zastępowano punktem materialnym o masie równej masie koła, natomiast biorąc pod uwagę wrzeciono ściernicy, to pomijano jego masę. Uwzględniano tym samym wyłącznie podatność więzi sprężystej. W większości przypadków takie uproszczenie jest dopuszczalne, gdyż otrzymane wyniki nie odbiegają w znaczny sposób od rozwiązania ścisłego, dla którego obraz zjawiska dość dobrze odpowiada rzeczywistości. Takie podejście jest wystarczające, dla takich przypadków, gdy masa koła jest dostatecznie duża w stosunku do masy trzpienia. Jednak, gdy jest ona porównywalna z masą więzi sprężystej powinno się ją uwzględnić [2].

W tym przypadku można skorzystać z metody przybliżonej Rayleigha. Wtedy częstość drgań koła wyznaczona zostanie z warunku zachowania energii. Założono, że masa trzpienia jest mniejsza od masy koła i nie zmienia charakteru drgań, że w czasie drgań poprzecznych linia ugięcia trzpienia przybiera w każdej chwili czasu postać geometrycznie podobną do statycznej linii ugięcia [1, 2] pod wpływem umownej siły zewnętrznej działającej w miejscu osadzenia koła (rys. 1). Można wówczas skorzystać z następującej proporcji wiążącej drgania poprzeczne z ugięciem statycznym [2]:

$$\frac{y}{y_{st}} = \frac{\omega}{f_t} \quad (1)$$

Korzystając z (1) można uzyskać wyrażenia określające przemieszczenie oraz prędkość wycinka trzpienia za pomocą ugięcia i prędkości masy skupionej koła zębatego [2]:

$$y = \frac{y_{st}(X)}{f_t} W, \quad \dot{y} = \frac{y_{st}(X)}{f_t} \dot{W} \quad (2)$$

Zastosowanie metody Rayleigha odpowiada reprezentacji ośrodka ciągłego, jako układu o jednym stopniu swobody. Z nieskończonej liczby postaci drgań trzpienia uwzględniono wyłącznie jego pierwszą postać, która odpowiada działaniu siły statycznej.

*) Instytut Badań i Ekspertyz Naukowych w Gorzowie Wlkp.

Przy formułowaniu zasady zachowania energii pominięto bardzo małą energię kinetyczną ruchu obrotowego elementów długości trzpienia. Nie uwzględniono także energii ruchu obrotowego koła. Wynika to z faktu, że przy analizie małych drgań trzpienia założono, iż osadzone na nim elementy nie doznają przemieszczeń kątowych. Odpowiada to sprowadzeniu ruchu koła do drgań punktu materialnego [1,3]. Wtedy energia kinetyczna drgającego układu składać się będzie z energii kinetycznej koła w ruchu postępowym E_k oraz energii kinetycznej trzpienia E_r :

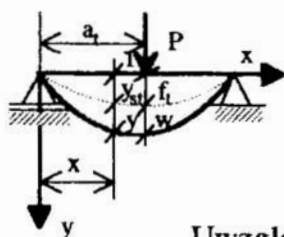
$$E = E_k + E_r \quad (3)$$

Przyjmując, że masa trzpienia jest rozłożona w sposób ciągły oraz korzystając z ogólnego wyrażenia na energię kinetyczną ($E = 1/2mv^2$) otrzymano:

$$E = 1/2m_k \dot{w}^2 + 1/2m_t/l_t \int y(x)^2 dx \quad (4)$$

Wprowadzając do powyższego równania proporcje (2) otrzymano energię całkowitą, która przyjęła następującą postać:

$$E = 1/2 \left\{ m_k + \frac{m_t}{l_t} \int [y_{st}(x)]^2 dx \right\} \dot{w}^2 \quad (5)$$



Rys. 1. Schemat działania siły zewnętrznej w miejscu osadzenia koła

Fig. 1. Scheme of the external force action in the place a gear is set

Uwzględniając odpowiadający temu przypadkowi przedział całkowania otrzymano [2]:

$$m_z = m_k + \frac{m_t}{l_t} \int_0^{l_t} [y_{st}(x)]^2 dx \quad (6)$$

gdzie statyczna linia ugięcia została już wyznaczona [2]:

$$y_{st}(x) = \begin{cases} y_1 - (2)_I & \text{dla } x \leq a_t \\ y_2 - (2)_{II} & \text{dla } x \geq a_t \end{cases} \quad (7)$$

Podstawiając do tak otrzymanej zależności na masę zredukowaną sprowadzoną do punktu osadzenia koła otrzymano równania linii ugięcia [1,3], wtedy po obliczeniu całki oznaczonej masa zastępcza przyjmie postać ostateczną:

$$m_z = m_k + \frac{(l_t + a_t)^2 (2l_t^2 - 3a_t^2) + 6a_t^4}{105a_t^2 (l_t - a_t)^2} m_t \quad (8)$$

Dla tak określonej masy układu zastępczego częstość własna drgań poprzecznych będzie następująca:

$$\omega_k = \sqrt{\frac{k_t}{m_z}} \quad (9)$$

Pozostałe związki na drgania poprzeczne koła zębatego będą takie same, jak przy analizie ruchu bezmasowego trzpienia [4]. Ponieważ w tym przypadku trzpień jest osadzony pionowo, więc pominięto wpływ siły ciężkości, a opuszczając ugięcie statyczne z równania ruchu otrzymano:

$$\begin{aligned} x_k &= A_x \cos(\omega_k t - \psi_x), \\ y_k &= A_y \cos(\omega_k t - \psi_y) \end{aligned} \quad (10)$$

W przypadku drgań skrętnych, stosując metodę Rayleigha, uwzględniono zastępczy masowy moment bezwładności. Częstość kołową drgań skrętnych uzyskano analogicznie, jak przy analizie drgań poprzecznych zastępując dotychczasowy moment bezwładności koła B_k bezwładnością zastępczą:

$$\omega_s = \sqrt{\frac{k_s}{B_z}} \quad (11)$$

gdzie bezwładność zredukowana, uwzględniająca wpływ więzi sprężystej określona jest następująco:

$$B_z = B_k + 1/3B_t \quad (12)$$

Przy wyprowadzaniu powyższej zależności skorzystano z proporcji wynikającej

z teorii skręcania, że kąt skręcania trzpienia zależy od położenia rozpatrywanej części elementu i jest wprost proporcjonalny do jej odległości od punktu utwierdzenia w zabieraku. Amplitudę oraz fazę drgań skrętnych, składającą się na rozwiązanie ruchu obrotowego, wyznaczono z zależności [1,2].

3. Drgania wrzeciona ściernicy

Przy analizie ruchu wrzeciona założono, że jest ono osadzone jednostronnie w sztywnych łożyskach [2]. Przyjęto, że wpływ podpór łożyskowych oraz przenoszonych za ich pośrednictwem sił i wymuszeń kinematycznych pochodzących od innych zespołów obrabiarki na dynamikę wrzeciona jest pomijalnie mały. Wirujące wrzeciono osadzone jest w poziomych łożyskach. Zatem należy uwzględnić wpływ siły ciężkości. Mogą bowiem powstać drgania drugiego rzędu. Są one wywoływane tak, jak drgania pierwszego rodzaju, a więc niewyrównoważeniem wrzeciona. Kinetostatyczna linia ugięcia wiruje wokół statycznej linii ugięcia. Wspomniane drgania zachodzą z częstością równą połowie częstości giętnych drgań własnych układu. Jak wykazano w [1,2,3] siła ta powoduje nieznaczne zaburzenia ruchu obrotowego. Założono, że prędkość obrotowa wrzeciona jest wtedy stała.

Rozpatrując drgania ściernicy przyjęto, że wrzeciono ściernicy jest bezmasowe. Posiada wyłącznie sztywność. Takie podejście podyktował stosunek masy ściernicy wraz z elementami mocującymi do masy wrzeciona. Jest ona na tyle duża, że można pominąć masę więzi sprężystej.

Uwzględniono natomiast obroty ściernicy w płaszczyznach ugięcia [1,2,3] z uwagi na występujące momenty żyroskopowe. W takim przypadku wirująca na wrzecionie ściernica posiada w rzeczywistości sześć stopni swobody. Ruch środka masy opisują znane równania Newtona, które można wyrazić w zapisie wektorowym:

$$m \frac{d\vec{v}_s}{dt} = \vec{p} \quad (13)$$

Uwzględnić jeszcze należy równania ruchu obrotowego dookoła środka masy, które opisuje znane z mechaniki równanie krętu:

$$\frac{d\vec{K}_s}{dt} = \vec{M} \quad (14)$$

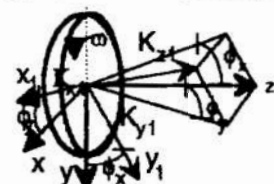
Wynika z niego, że pochodna względem krętu układu punktów materialnych, czy też jak w tym przypadku bryły sztywnej względem dowolnego nieruchomego bieguna, równa jest sumie momentów wszystkich sił zewnętrznych działających na ten biegun. Powyższa zależność prowadzi do trzech odpowiednich równań skalarnych:

$$\frac{dK_x}{dt} = M_x, \quad \frac{dK_y}{dt} = M_y, \quad \frac{dK_z}{dt} = M_z \quad (15)$$

z których uwzględniono jedynie pierwsze dwa, gdyż trzecie równanie opisuje z założenia jednostajny ruch obrotowy ze stałą prędkością kątową $\omega = \text{const}$.

Wprowadzając równania ruchu obrotowego wrzeciona przyjęto, że ściernica jest wyrównowazona, zarówno statycznie jak i dynamicznie, a więc jej środek ciężkości znajduje się na osi obrotu, która jest równocześnie główną centralną osią bezwładności.

W trakcie ruchu obrotowego oś wrzeciona wychyla się z położenia wyznaczonego przez prostą łączącą sztywne podpory i tworzy z płaszczyznami nieruchomego układu współrzędnych "xz" i "yz" małe kąty: ϕ_x oraz ϕ_y (rys. 2).



Rys. 2. Wychylenia osi wrzeciona w trakcie ruchu obrotowego

Fig. 2. Displacement of a spindle axis during rotational movement

Wprowadzono ruchomy układ współrzędnych zaczepiony w środku ciężkości i przemieszczający się wraz z ruchem środka ściernicy "x₁y₁z₁".

Wtedy ściernica obraca się wokół osi "z₁" z prędkością kątową równą ω , natomiast prędkości kątowe względem pozostałych osi układu ruchomego określane są jako pochodne odpowiednich kątów: $d\phi_x/dt$ i $d\phi_y/dt$.

Zgodnie z definicją krętu, kręt ciała materialnego wokół osi obrotu równy jest iloczynowi momentu bezwładności i prędkości kątowej względem tej osi. Dla ciała symetrycznego, jakim jest ściernica wirująca dookoła osi symetrii nie wystąpią odśrodkowe momenty bezwładności:

$$B_{x_1} = B_{y_1} = B = \text{const.}, \quad B_{z_1} = B_0 = \text{const.}, \quad B_{x_1y_1} = B_{y_1z_1} = B_{x_1z_1} = 0 \quad (16)$$

wtedy składowe krętu wokół osi układu ruchomego będą określone zależnościami:

$$K_{x_1} = B \frac{d\phi_x}{dt} = B\dot{\phi}_x, \quad K_{y_1} = B \frac{d\phi_y}{dt} = B\dot{\phi}_y, \quad K_{z_1} = B_0\omega \quad (17)$$

By wyznaczyć równania ruchu obrotowego powinno się jednak uwzględnić kręty względem nieruchomych osi "x" i "y". Ich wartość uzyskano rzutując te osie składowe krętu w układzie ruchomym (rys. 2). Skorzystano przy tym z założenia, że dla małych kątów obrotu wrzeciona w punkcie osadzenia ściernicy spełnione są następujące zależności:

$$\sin \phi_x \cong \phi_x, \quad \sin \phi_y \cong \phi_y, \quad \cos \phi_x \cong \phi_y \cong 1 \quad (18)$$

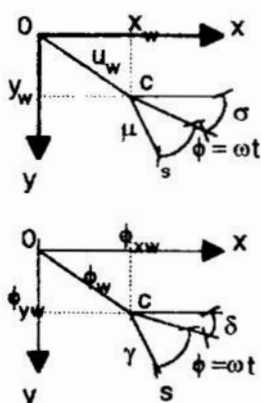
a są one słuszne, gdy układ wykonuje małe drgania. Biorąc pod uwagę powyższe założenia otrzymano:

$$K_x = K_{x_1} + K_{z_1}\phi_y = B\dot{\phi}_x + B_0\omega\phi_y, \quad K_y = K_{y_1} - K_{z_1}\phi_x = B\dot{\phi}_y - B_0\omega\phi_x \quad (19)$$

Można zatem przejść do ostatecznego zapisu równań ruchu bryły sztywnej. Tym samym ogólne równania ruchu ściernicy: (13) oraz (15) prowadzą do poniższych równań równowagi dynamicznej:

$$m\ddot{x}_s = P_x, \quad B\ddot{\phi}_{ys} - B_0\omega\dot{\phi}_{xs} = M_y, \quad m\ddot{y}_s = P_y, \quad B\ddot{\phi}_{xs} + B_0\omega\dot{\phi}_{ys} = M_x \quad (20)$$

z których dwa pierwsze opisują ruch w płaszczyźnie "xz", natomiast kolejne dwa w płaszczyźnie "yz". Z sześciu ogólnych równań ruchu uwzględniono wyłącznie cztery. Pominięto bowiem równania ruchu w kierunku osi obrotu. Po lewej stronie układu różniczkowego (20) występują nieznanne siły zewnętrzne, które pochodzą od oddziaływania wrzeciona na ściernicę. Są one związane ze sztywnością rozpatrywanego układu. Przyjęto określone obciążenie wrzeciona siłami umownymi [2]. Stosując zasadę superpozycji, niezależności działania sił określono przemieszczenia oraz obroty wrzeciona w punktach działania sił [2] dla każdej z płaszczyzn. Wyznaczono elementy macierzy podatności [2] oraz określono siły zewnętrzne działające na wrzeciono [2]. Wreszcie sprowadzono uzyskany układ do przemieszczeń opisujących ruch punktu "c" (rys. 3). Wykorzystano w tym celu przyjęte zależności wiążące ruch środka masy z ruchem środka wrzeciona w punkcie osadzenia ściernicy (rys. 3) (21):



Rys. 3. Związki ruchu środka masy z ruchem środka wrzeciona

Fig. 3. Relations between centre of mass movement and a spindle centre movement

W uzyskanych związkach występują parametry określające niewyrównoważenie statyczne i momentowe ściernicy [3,5]. Wpływ mimośrodowości środka ciężkości można sobie łatwo wyobrazić (rys. 3a), gdy idzie natomiast o związki między obrotem ściernicy, a obrotem wrzeciona w jej płaszczyźnie, to sytuacja ta nie jest już

tak jednoznaczna (rys. 3b).-Na drugim rysunku punkty "c" oraz "s" zaznaczono analogicznie, jak w przypadku niewyrównoważenia statycznego. Dla niewyrównoważenia momentowego punkty te pokrywają się. Jednak dla przejrzystości

na rysunku zaznaczono je oddzielnie.

$$\begin{aligned} X_s &= X_w + \mu \cos(\omega t + \sigma), & Y_s &= Y_w + \mu \sin(\omega t + \sigma), & \phi_{xs} &= \phi_{xw} + \gamma \cos(\omega t + \delta) \\ \phi_{ys} &= \phi_{yw} + \gamma \sin(\omega t + \delta) \end{aligned} \quad (21)$$

Uwzględniając zależności (21) otrzymano następujący niejednorodny układ równań różniczkowych:

$$\begin{aligned} \bar{m} \ddot{X}_w + k_{11} \dot{X}_w + k_{12} \phi_{yw} &= m \mu \omega^2 \cos(\omega t + \sigma) \\ B \ddot{\phi}_{yw} - B_0 \omega \dot{\phi}_{xw} + k_{21} \dot{X}_w + k_{22} \phi_{yw} &= (B - B_0) \gamma \omega^2 \sin(\omega t + \delta) \end{aligned} \quad (22)$$

$$\bar{m} \ddot{Y}_w + k_{11} \dot{Y}_w - k_{12} \phi_{xw} = m \mu \omega^2 \sin(\omega t + \sigma)$$

$$\bar{B} \ddot{\phi}_{xw} + B_0 \omega \dot{\phi}_{yw} - k_{21} \dot{Y}_w + k_{22} \phi_{xw} = (B - B_0) \gamma \omega^2 \cos(\omega t + \delta)$$

Rozwiązanie tego układu równań równowagi, opisującego drgania wrzeciona w punkcie "C" składać się będzie z rozwiązania ogólnego układu jednorodnego i całek szczególnych równań niejednorodnych. Przyjęto następujące rozwiązania ogólne:

$$\begin{aligned} X_{wj} &= X_{wj} \cos(\alpha_j t - \beta_j), & \phi_{xwj} &= F_{xj} \sin(\alpha_j t - \beta_j) \\ Y_{wj} &= Y_{wj} \sin(\alpha_j t - \beta_j), & \phi_{ywj} &= F_{yj} \cos(\alpha_j t - \beta_j) \end{aligned} \quad (23)$$

gdzie wskaźnik "j" odpowiada j-tej częstości drgań swobodnych. Z pierwszych trzech równań układu (22) otrzymano związki między amplitudami drgań własnych:

$$\begin{aligned} \frac{Y_{wj}}{X_{wj}} = \chi_{1j} &= \frac{k_{12} k_{21} - (k_{22} - B \alpha_j^2) (k_{11} - m \alpha_j^2)}{(k_{11} - m \alpha_j^2) B_0 \omega \alpha_j}, & \frac{F_{xj}}{X_{wj}} = \chi_{2j} &= \frac{k_{12} k_{21} - (k_{22} - B \alpha_j^2) (k_{11} - m \alpha_j^2)}{k_{12} B_0 \omega \alpha_j}, \\ \frac{F_{yj}}{X_{wj}} = \chi_{3j} &= \frac{m \alpha_j^2 - k_{11}}{k_{12}} \end{aligned} \quad (24)$$

natomiast z czwartego równania następujące równanie częstości:

$$m B \alpha^4 - m B_0 \omega \alpha^3 - (k_{22} m + k_{11} B) \alpha^2 + k_{11} B_0 \omega \alpha + k_{11} k_{22} - k_{12}^2 = 0 \quad (25)$$

które rozwiązano w sposób przybliżony na drodze obliczeń numerycznych [2,5].

Zanim wyznaczono rozwiązania ogólne całek szczególnych układu równowagi dynamicznej dokonano zapisu zespolonego [2] w wyniku którego otrzymało:

$$\begin{aligned} (k_{11} - m \omega^2) e^{i\sigma} U - i k_{12} e^{i\delta} F &= m \mu \omega^2 e^{i\sigma}, \\ i k_{21} e^{i\sigma} U + [k_{22} + (B_0 - B) \omega^2] e^{i\delta} F &= (B - B_0) \gamma \omega^2 e^{i\delta} \end{aligned} \quad (26)$$

Po rozwiązaniu powyższego układu dwóch równań [2] stwierdzono, że rozwiązania szczególne są kombinacją liniową pod wpływem każdego z dwóch przypadków niewyrównowazenia z osobna [2] zaś amplitudy drgań wymuszonych mają określoną wartość [2]. Ponadto, stwierdzono, że gdy prędkość kątowna ω osiąga wartości pierwiastków równania częstości własnych (25), to amplitudy drgań wymuszonych rosną nieograniczenie. Układ przechodzi przez rezonans. Zatem, gdy wirującą ściernicę charakteryzuje określony stosunek momentów bezwładności B_0/B dla którego $\Theta > 45^\circ$, to posiada ona cztery częstości własne drgań giętych oraz jedną prędkość krytyczną precesji prostej od niewyrównowazenia ($\alpha = \omega$) dla każdej prędkości obrotowej wrzeciona. Prędkość krytyczna precesji współbieżnej jest przy tym większa od wartości przybliżonej ω_k otrzymanej w przypadku, gdy nie uwzględniono skoszeń ściernicy i efektu żyroskopowego. Dla mniejszych od jedności wartości stosunku momentów bezwładności, gdy $\Theta < 45^\circ$ pojawia się dodatkowa prędkość krytyczna związana z przyjętym jako dodatkowy stopniem swobody [2].

Po wyznaczeniu rozwiązań niejednorodnego układu równań: (23), (24) i [2] różniczkowych równowagi wrzeciona w punkcie osadzenia ściernicy można opisać ruch wrzeciona jako superpozycję drgań własnych oraz wymuszonych. Po wykorzystaniu wzoru Eulera rozwiązanie przyjmie postać [3,5] zaś otrzymany układ rozwiązań na przemieszczenia wrzeciona będzie zawierał dwa nowe parametry określające amplitudę drgań wymuszonych za pomocą liczb rzeczywistych [2]:

$$\bar{U}_\gamma = -i U_\gamma, \quad \bar{F}_\mu = i F_\mu \quad (27)$$

W równaniu tym wystąpi osiem niewiadomych parametrów opisujących drgania własne: $X_{w1}, X_{w2}, X_{w3}, X_{w4}$ oraz B_1, B_2, B_3, B_4 . Przyjmując określone warunki początkowe (dla $t=0$) w punkcie "C" otrzymano osiem potrzebnych równań nieliniowych. Równania te można podzielić na dwa układy równań liniowych ze względu na nowe zmienne, tj. $X_{wj}\sin\beta_j$ i $X_{wj}\cos\beta_j$:

$$\left\{ \begin{array}{l} \sum_{j=1}^4 \alpha_j X_{wj} \sin \beta_j = x_{w0} + (U_\mu \sin \sigma + \bar{U}_\gamma \cos \delta) \omega \\ \sum_{j=1}^4 \chi_{1j} X_{wj} \sin \beta_j = U_\mu \sin \sigma + \bar{U}_\gamma \cos \delta - y_{w0} \\ \sum_{j=1}^4 \chi_{2j} X_{wj} \sin \beta_j = \bar{F}_\mu \sin \sigma + F_\gamma \cos \delta - \phi_{xw0} \\ \sum_{j=1}^4 \chi_{3j} \alpha_j X_{wj} \sin \beta_j = \phi_{yw0} - (\bar{F}_\mu \sin \sigma + F_\gamma \cos \delta) \omega \\ \sum_{j=1}^4 X_{wj} \cos \beta_j = x_{w0} + \bar{U}_\gamma \sin \delta + U_\mu \cos \sigma \\ \sum_{j=1}^4 \chi_{1j} \alpha_j X_{wj} \cos \beta_j = y_{w0} + (\bar{U}_\gamma \sin \delta - U_\mu \cos \sigma) \omega \\ \sum_{j=1}^4 \chi_{2j} \alpha_j X_{wj} \cos \beta_j = \phi_{xw0} + (F_\gamma \sin \delta - \bar{F}_\mu \cos \sigma) \omega \\ \sum_{j=1}^4 \chi_{3j} X_{wj} \cos \beta_j = \phi_{yw0} - F_\gamma \sin \delta + \bar{F}_\mu \cos \sigma \end{array} \right. \quad (28)$$

Uzyskane układy równań rozwiązano w sposób analityczny za pomocą metody Cramera. Następnie w wyniku prostych działań amplitudy oraz fazy drgań swobodnych będą znane [2,5].

Interesujące są jednak drgania środka ciężkości niewyrównoważonej ściernicy. Aby przejść do środka masy "S" wykorzystano związki między przemieszczeniami (21). Wtedy:

$$\left\{ \begin{array}{l} x_s = \sum_{j=1}^4 X_{wj} \cos(\alpha_j t - \beta_j) + \tilde{U}_\mu \cos(\omega t + \sigma) - \bar{U}_\gamma \sin(\omega t + \delta) \\ y_s = \sum_{j=1}^4 \chi_{1j} X_{wj} \sin(\alpha_j t - \beta_j) + \tilde{U}_\mu \sin(\omega t + \sigma) + \bar{U}_\gamma \cos(\omega t + \delta) \\ \phi_{xs} = \sum_{j=1}^4 \chi_{2j} X_{wj} \sin(\alpha_j t - \beta_j) + \bar{F}_\mu \sin(\omega t + \sigma) + \bar{F}_\gamma \cos(\omega t + \delta) \\ \phi_{ys} = \sum_{j=1}^4 \chi_{3j} X_{wj} \cos(\alpha_j t - \beta_j) - \bar{F}_\mu \cos(\omega t + \sigma) + \bar{F}_\gamma \sin(\omega t + \delta) \end{array} \right. \quad (29)$$

gdzie:

$$\tilde{U}_\mu = U_\mu + \mu, \quad \bar{F}_\gamma = F_\gamma + \gamma$$

Układ ten opisuje współbieżne drgania wrzeciona i ściernicy. Wprowadzając dodatkowe warunki swobody zwiększono liczbę częstości własnych, natomiast drgania wymuszone zachodzą zarówno pod wpływem niewyrównoważenia statycznego jak i momentowego. Przy czym są to drgania sprzężone, bowiem każde z niewyrównoważeń wpływa nie tylko na odpowiadające mu przemieszczenia ściernicy, ale na każde z przemieszczeń [2], tj. ugięcie i kąt obrotu równocześnie.

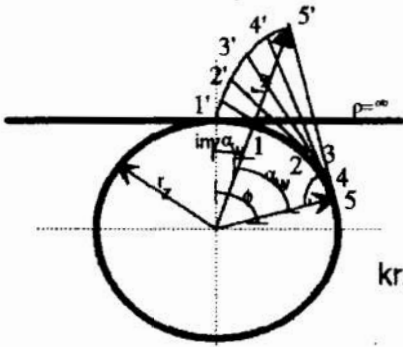
4. Drgania występujące podczas obróbki

Przy analizie ruchu drgającego dokonano rozdziału w układzie P-N. Określając wielkość drgań podczas obróbki zęba uwzględniono zarówno drgania koła, jak i ściernicy [1,3]. Przyjęto, że środek ciężkości ściernicy jest również jej środkiem geometrycznym, a główna centralna oś bezwładności osią symetrii. Z uwagi na to, że przyjęto jako możliwe obroty wrzeciona w miejscu osadzenia ściernicy należy także uwzględnić jej obroty wokół osi "x". Przy wyprowadzeniu wzoru na wielkość drgań podczas obróbki pominięto składowe drgań w kierunku stycznym do zarysu zęba oraz zmianę kąta ϕ_{ys} , ponieważ dla normalnej szerokości wieńca obrabianego koła ramię jego działania jest znacznie mniejsze od promienia ściernicy.

Z uwagi na to, że podczas szlifowania koło zębate wykonuje ruch odtaczania oraz związany z nim ruch posuwisty przyjęto, że układ współrzędnych określający drgania koła jest z nim związany na stałe, a więc jest układem ruchomym. W takiej sytuacji inne też będą kąty określające

położenie punktu styku narzędzia z przedmiotem oraz usytuowanie składowych drgań koła [1,2,3]. Tak więc należy ustalić położenie punktu współpracy względem ruchomego koła. Aby tego dokonać należy opisać geometrię zarysu zęba.

Ogólnie biorąc kształt zarysu boku zęba może być dowolny. W praktyce najkorzystniejsze okazały się jednak zarysy utworzone przez krzywe cykliczne, z których najczęściej spotykaną jest ewolwenta [6]. Jest ona szczególnym przypadkiem krzywych cyklicznych, a powstaje, gdy promień koła odtaczanego jest nieskończony $p=\infty$. Oznacza to, że prosta toczy się po torze kołowym, określanym inaczej jako koło zasadnicze (rys. 4).



Rys. 4. Powstawanie ewolwenty w oparciu o koło zasadnicze

Fig. 4. An involute in strength of base circle formation

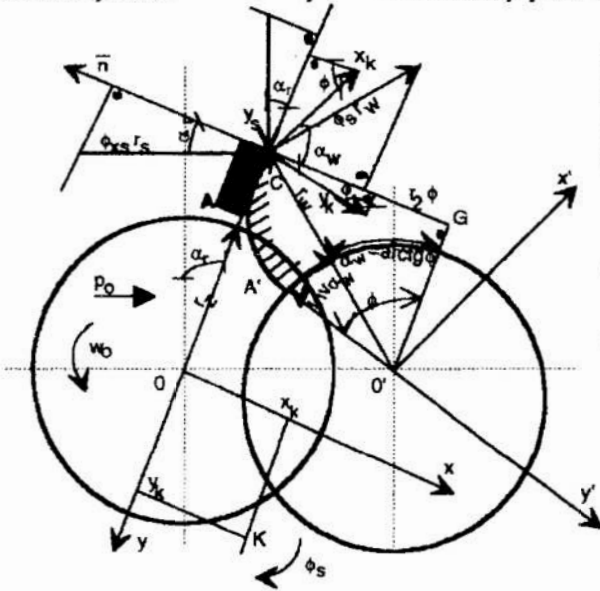
Jest ona styczna do koła zasadniczego z którego jest odwijana, a jednocześnie normalna do zarysu kreślonej ewolwenty. Dla tak uzyskanej krzywej można wyprowadzić następujące związki określające położenie dowolnego punktu ewolwenty:

$$r_w = r_z \sqrt{1 + \phi^2}, \quad \alpha_w = \arctg \phi \quad (30)$$

gdzie:

r_z – promień koła zasadniczego, ϕ – kąt odtoczenia ($\phi = \omega_0 t$).

Przy szlifowaniu sposobem Nilesa ściernica wykonuje ruch obrotowy wokół swojej osi oraz posuwisto zwrotny wzdłuż boku zęba. Przedmiot obrabiany wykonuje natomiast ruch odtaczania i sprzężony z nim ruch posuwisty [6]. Ruchy te w konsekwencji odwzorowują współpracę obrabianego koła z ściernicą zaś bokowi zęba nadawany jest zarys ewolwentowy (rys. 5).



Rys. 5. Nadawany zarys ewolwentowy bokowi zęba podczas szlifowania

Fig. 5. An involute profile given to a tooth side during process of grinding

r_w – promień wodzący punktu współpracy

α_r – obróbczo-toczny kąt zarysu, kąt nachylenia krawędzi ściernicy do prostej prostopadłej do kierunku posuwu p_0

α_w – kąt nachylenia promienia wodzącego do stycznej w punkcie współpracy

r_s – promień ściernicy

W punkcie styku narzędzia z przedmiotem obrabianym, ruch wypadkowy ściernicy względem zarysu zęba powstaje w wyniku złożenia ruchów składowych. Przy czym są one odzwierciedleniem drgań w omawianym układzie. Rzutując wyliczone składowe drgań: (10), [1], (29) na normalną do zarysu boku zęba (rys. 5) otrzymano następujące wyrażenie określające wiel –

kość drgań podczas obróbki:

$$\delta_r(\phi) = x_k \cos \phi + y_k \sin \phi + \phi_s r_w \cos \alpha_w - y_s \sin \alpha_r + \phi_{xs} r_s \cos \alpha_r \quad (31)$$

Gdyby założyć, że oś "y" pokrywa się z kierunkiem pionowym należałoby uwzględnić nie drgania ściernicy w kierunku tej osi, ale w kierunku osi "x", gdyż trzpień jest osadzony pionowo. Z uwagi jednak na to, że ugięcie statyczne nie wpływa na dokładność obróbki osie układu współrzędnych mogą być dowolnie położone w płaszczyźnie ściernicy. Więc przyjęto je tak, jak przedstawiono na rys. 5.

Definiując wielkość drgań podczas obróbki przyjęto, że $\delta_r > 0$, gdy grubość warstwy szlifowanej jest mniejsza od wartości założonej.

Podczas kształtowania uzębienia prawego boku zęba zmianie ulegną wyłącznie znaki przy trzech składowych wyrażenia (31). Tak więc:

$$\delta_r(\phi) = \pm x_k \cos \phi + y_k \sin \phi \pm \phi_s r_w \cos \alpha_w - y_s \sin \alpha_r + \phi_{xs} r_s \cos \alpha_r \quad (32)$$

co odpowiada szlifowaniu zarysu lewego, jak i prawego boku zęba.

5. Zakończenie

Przedstawiony model matematyczny może być wykorzystany zarówno do oceny wielkości drgań powstających podczas szlifowania zębów kół zębatych jak i do analizy grubości warstwy skrawanej. Może też służyć do diagnostyki wibroakustycznej procesu szlifowania kół zębatych.

6. Literatura

1. Maciejewski J., Zaborowski T.: Algorytm do analizy drgań przy szlifowaniu zębów kół zębatych. ODE Gorzów Wlkp. IV Ogólnopolskie seminarium naukowe nt. Zastosowanie matematyki w budowie maszyn—ZMBM'89. Studia i Materiały t. VIII, nr 1–2, Seria: Referaty, 1989
2. Zaborowski T.: Drgania tworzące się podczas szlifowania zębów kół zębatych sposobem Nilesa. Studia i Materiały t. XI, nr 2, Seria: Monografie (w druku)
3. Zaborowski T.: Analiza przyczyn wpływających na niedokładność wykonania zębów kół zębatych szlifowanych sposobem Nilesa powodujących brak stabilizacji procesu w wyniku drgań. III Ogólnopolska konf. nauk.–tech. nt. Obróbka ścierna kół zębatych. ODE Gorzów Wlkp., Politechnika Poznańska, Gorzów Wlkp.–Głębokie k/Międzyrzecz 1989
4. Zaborowski T.: Analysis of Vibration in Gear Teeth Grinding Process. IBEN Gorzów Wlkp. Studia i Materiały t. X, nr 2, Seria: Referaty, 1991
5. Zaborowski T., Maciejewski J.: Numeryczna analiza drgań występujących w procesie szlifowania zębów kół zębatych. GOBiEN Gorzów Wlkp., Studia i Materiały t. VII, nr 2, Seria: Monografie, 1988
6. Ochęduszko K.: Koła zębate, t. I. WNT Warszawa 1974

MATEMATICAL MODEL OF VIBRATION BEING GENERATED DURING GEAR TEETH GRINDING WITH NILES' METHOD

Summary

A mathematical model of vibration being generated during gear teeth grinding with Niles' method have been presented in the paper. Relations with a grinding arbor on which a gear is set have been indicated as well as a spindle and a grinding wheel fixed on it. Eventually the model describing a quantity of vibration being generated during gear teeth grinding has been provided.

MATEMATYCZNA ANALIZA KRZYWYCH UDZIAŁU NOŚNEGO LINIOWEGO

Analiza wyników pomiarów parametrów profilu powierzchni przy użyciu współczesnej techniki obliczeniowej. Metoda aproksymacji wyników pomiarów t_p za pomocą wielomianów ortogonalnych. Wykorzystanie otrzymanego wielomianu aproksymującego do matematycznej i funkcjonalnej analizy wyników pomiaru parametrów profilu powierzchni

1. Wprowadzenie

Ostatnie lata cechują się stałym postępem technologii stosowanych w budowie maszyn. W ślad za tym zwiększają się wymagania dotyczące technik pomiarowych stosowanych w pomiarach parametrów stereometrii warstwy wierzchniej. Postęp wymusił na konstruktorach i technologach wprowadzenie i coraz powszechniejsze stosowanie obszerniejszych, zawierających więcej parametrów, kompleksowych charakterystyk profilu chropowatości powierzchni wytworzonych wyrobów. Przy obecnym intensywnym rozwoju metod obróbki, wprowadzeniu nowoczesnych technologii oraz przy dążeniu do maksymalnego zwiększenia okresu eksploatacji maszyn coraz powszechniej stosuje się do oceny obrobionej powierzchni, oprócz parametrów R_a , R_m [1, 2, 3] oraz tracącego na znaczeniu parametru R_z , krzywe nośności liniowej t_p . Udział nośny staje się jednym z ważniejszych kryteriów charakteryzujących własności użytkowe wytwarzanych powierzchni, szczególnie w wypadku powierzchni obciążonych siłami zewnętrznymi współpracującymi ze sobą w spoczynku lub w ruchu.

Znajomość przebiegu krzywych udziału nośnego nabiera wówczas dużego znaczenia, gdyż rzeczywisty styk obu powierzchni jest znacznie mniejszy od nominalnego obszaru styku, a o jego wartości można wnioskować znając przebieg krzywej t_p . Największe znaczenie ma przy tym przebieg krzywej udziału nośnego dla wartości $0 \div 0,25 R$, gdyż te fragmenty nierówności zostają usunięte w trakcie wstępnego docierania współpracujących części w początkowym okresie eksploatacji. Przebieg krzywej nośności w tym obszarze determinuje przebieg procesu docierania, pozwala w pewnym stopniu przewidzieć cechy użytkowe powierzchni w trakcie właściwej eksploatacji (po okresie docierania) danego wyrobu.

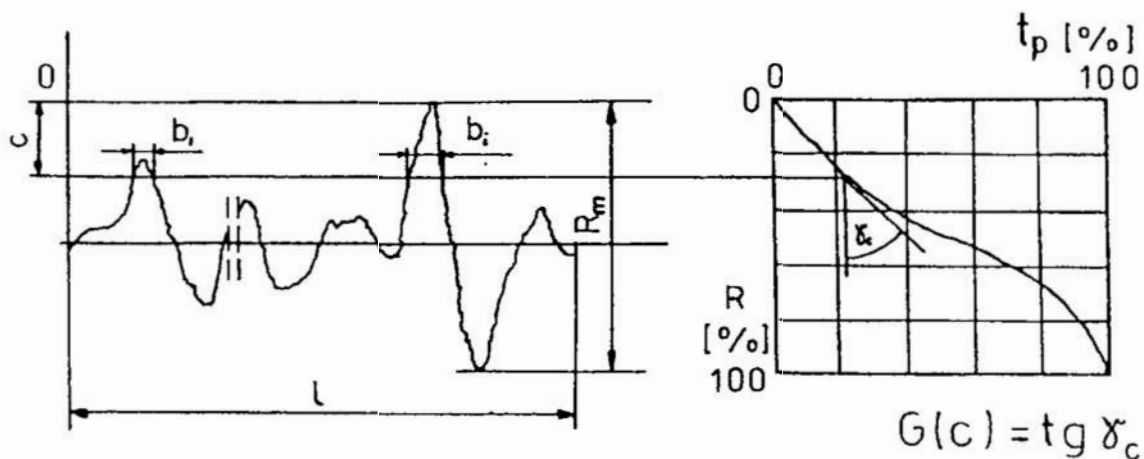
2. Znaczenie krzywych nośności

Sposób powstawania krzywej t_p objaśnia rys.1. Przecinając profil o długości l linią leżącą w odległości $R = c/R_m\%$ od najwyższego wierzchołka profilu otrzymuje się odcinki b_1, b_2, \dots, b_n , których suma, zwana długością nośną, stanowi podstawę do określenia udziału nośnego liniowego profilu:

$$t_p = \frac{1}{l} \sum_{i=1}^n b_i \quad (1)$$

Wartość R przyjmuje [3] wartości: 5, 10, 15, 20, 25, 30, 40, 50, 60, 70, 80 i 90%.

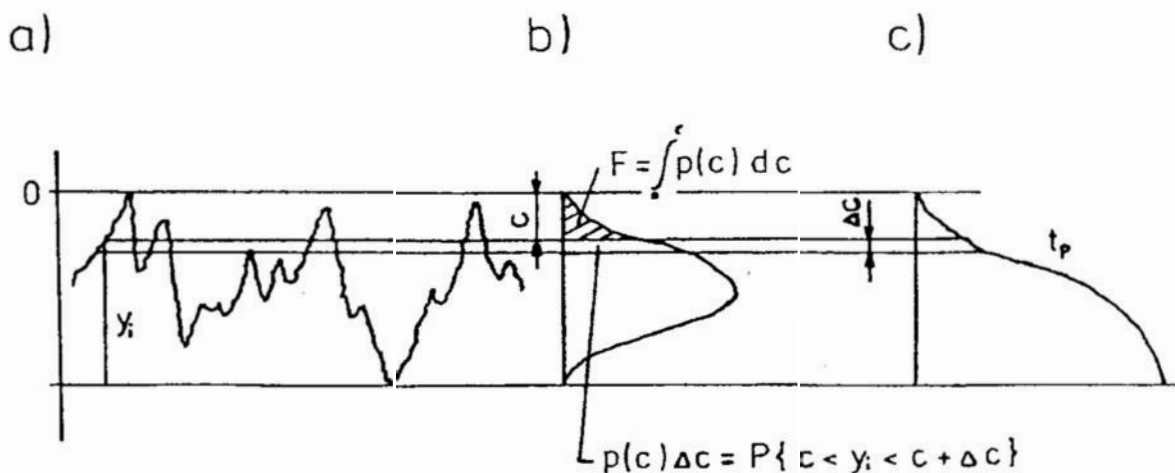
* Katedra Technologii Budowy Maszyn. Politechnika Częstochowska
42 - 200 Częstochowa ul. Armii Krajowej 21



Rys. 1. Sposób tworzenia krzywej udziału nośnego liniowego profilu chropowatości
 Fig. 1. The way of build the bearing linear curve with roughness profile

Dodatkową wielkością charakteryzującą przebieg udziału nośnego liniowego w wybranych punktach jest wartość parametru

$$G(x) = \frac{dt_p}{dR} \quad \text{dla } x = R \quad (2)$$



Rys. 2. Związek pomiędzy rozkładem rzędnych (b) i przebiegiem krzywej udziału nośnego (c) powstaje na podstawie zdjętego profilu chropowatości (a)

Znajomość analitycznej postaci krzywej $t_p = f(x)$ pozwoliłoby w łatwy sposób określić przebieg rozkładu rzędnych rozpatrywanego profilu. W tym celu (rys. 2) wystarczy wykorzystać fakt że rozkład udziału nośnego profilu t_p w funkcji zbliżenia

$x = R$ stanowi dystrybuantę gęstości prawdopodobieństwa rzędnej profilu $p(x)$. Rozkład rzędnych profilu $p(c)$ $c = x$ określa prawdopodobieństwo wystąpienia profilu (prawdopodobieństwo wystąpienia rzędnej profilu równej y_i) z przedziału $(c, c + \Delta c)$. Pomiędzy rozkładem krzywej nośności i rozkładem rzędnych profilu zachodzi związek:

$$t_p = \int_0^x p(x) dx \quad (3)$$

Ponieważ gęstość prawdopodobieństwa rzędnych profilu $p(x)$ jest równa pierwszej pochodnej funkcji dystrybuanty t_p ($t_p \approx f(x)$), można w przybliżeniu zapisać:

$$p(x) \approx f'(x) = G(x) \quad (4)$$

Tak więc prześledzenie przebiegu krzywej nośności i przybliżone wyznaczenie na tej podstawie przebiegu rozkładu rzędnych profilu pozwala analizować badane powierzchnie pod kątem ich konstytuowania. Oszacowanie rozkładu $p(x)$ pod względem jego asymetrii (skośności) oraz ekscesu (określającego stopień "rozmycia" rozkładu od centralnej wartości) umożliwia szybką identyfikację sposobu obróbki i przewidywanie niektórych własności eksploatacyjnych mierzonych powierzchni.

3. Aproksymacja krzywych udziału nośnego liniowego

Aproksymacja otrzymanych punktów pomiarowych $[x_p, t_{pi}]$ krzywej udziału nośnego liniowego może zostać przeprowadzona wieloma metodami (np. metoda najmniejszych kwadratów, interpolacji czy też analizy harmonicznej). W ich wyniku można otrzymać różne postacie równania aproksymującego. Literatura [2] proponuje aproksymację początkowej części krzywej nośności równaniem w postaci:

$$f(x) = ax^b \quad (5)$$

czy też

$$f(x) = ax^b + c \quad (6)$$

można również posłużyć się odpowiednio dobraną sumą krzywych wykładniczych, trygonometrycznych lub logarytmicznych.

W trakcie opisanych tu badań autorzy doszli do wniosku, że w zupełności zadowalające wyniki daje aproksymacja wielomianem stopnia n - tego

$$f(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n \quad (7)$$

Ta postać funkcji aproksymującej jest dogodna do dalszych przekształceń, pozwala na łatwą analizę uzyskiwanych funkcji i nie przysparza większych problemów w trakcie aproksymacji.

Po wyborze stopnia wielomianu n można znaleźć odpowiednie współczynniki $a_0 \dots a_n$, stosując metodę najmniejszych kwadratów. Jednak w metodzie tej w momencie zmiany stopnia wielomianu wszystkie współczynniki $a_0 \dots a_n$ trzeba wyznaczyć ponownie. Proces aproksymacji można jednak znacznie [4, 5, 6] uprościć stosując metodę wielomianów ortogonalnych Czebyszewa. Istota tej metody polega na tym, że aproksymującego wielomianu szuka się nie bezpośrednio w postaci składników a_ix^n , lecz w postaci kombinacji wielomianów. Zatem poszukiwany wielomian (7) przyjmuje postać:

$$f(x) = a_0 + a_1 \Phi_1(x) + a_2 \Phi_2(x) + \dots + a_n \Phi_n(x) \quad (8)$$

gdzie $\Phi_k(x)$ są wielomianami k -tego stopnia o postaci

$$\Phi_k(x) = x^k + \alpha_k^{(1)} x_{k-1} + \dots + \alpha_k^{(k)} \quad (9)$$

W metodzie tej aproksymującego wielomianu (7) poszukuje się wyznaczając sumy iloczynów wielomianów $\Phi_k(x)$ i odpowiednich współczynników a_k (8), przy czym dodanie nowych składników w równaniu (8), a więc zwiększenie stopnia wielomianu aproksymującego, nie zmienia współczynników przy składnikach wielomianu niższego stopnia wyznaczonych wcześniej. Wielomiany (9) są tak dobrane, by spełnione były następujące warunki:

$$\sum_{i=1}^n \Phi_k(x_i) \Phi_l(x_i) = 0 \quad \text{dla } k \neq l \quad (10)$$

$$\sum_{i=1}^n [\Phi_k(x_i)]^2 \neq 0 \quad (11)$$

Warunek (11) oznacza, że w żadnym z punktów x_i wielomian $\Phi_k(x_i)$ nie jest zerem (warunek ortogonalności). Wykorzystując warunki (10) i (11) można wyznaczyć współczynniki a_l ($a_0 \dots a_n$) równania (8)

$$a_l = \frac{\sum_{i=1}^n t_{pi} \Phi_l(x_i)}{\sum_{i=1}^n [\Phi_l(x_i)]^2} \quad l = 0, 1, 2, \dots, n \quad (12)$$

Posługując się warunkami (9, 10, 11) można wyznaczyć ogólną postać równania (9):

$$\Phi_{k+1}(x) = (x - \beta_{k+1}) \Phi_k(x) + \gamma_{k+1} \Phi_{k-1}(x) \quad (13)$$

gdzie

$$\beta_k = \frac{\sum_{i=1}^n x_i [\Phi_k(x_i)]^2}{\sum_{i=1}^n [\Phi_k(x_i)]^2} \quad (14)$$

$$\gamma_k = \frac{\sum_{i=1}^n x_i \Phi_{k-1}(x_i) \Phi_k(x_i)}{\sum_{i=1}^n [\Phi_{k-1}(x_i)]^2} \quad (15)$$

Tak wyznaczone wielomiany $\Phi_k(x)$ spełniają nierówność (11). Są to więc wielomiany ortogonalne.

Próby przeprowadzone przez autorów wykazały, że aby uzyskać zadowalające przybliżenie krzywej $f(x)$ należy zastosować wielomiany aproksymujące (8) o dość wysokich, trudnych do praktycznego wykorzystania stopniach. Wadę tę można wyeliminować stosując aproksymację wielomianami ortogonalnymi omówionymi w literaturze [5]. W metodzie tej przyjęto nieco inną niż (13) postać ogólną równania (9):

$$\Phi_k(x) = x^k - \alpha_k^{(1)} \Phi_{k-1}(x) - \alpha_k^{(2)} \Phi_{k-2}(x) - \dots - \alpha_k^{(k)} \quad (16)$$

gdzie:

$$\alpha_k^{(l)} = \frac{\sum_{i=1}^n x_i^l \Phi_{k-1}(x_i)}{\sum_{i=1}^n [\Phi_{k-1}(x_i)]^2} \quad l = 0, 1, 2, \dots, k \quad (17)$$

Metoda ta jest na tyle szybko zbieżna, że w zupełności wystarczające wydaje się przyjęcie do aproksymacji wielomianu (8) stopnia czwartego. Odchyłki wartości funkcji aproksymującej od wartości pomiarowych nie powinny na ogół przekraczać kilku procent wartości funkcji t_p w poszczególnych punktach, a otrzymany wielomian dość wiernie odtwarza przebieg wykresu pomiarowego. Oczywiście w razie potrzeby zawsze istnieje możliwość zwiększenia stopnia wielomianu do stopnia piątego i wyższych.

Dla przyjętej aproksymacji wielomianem stopnia czwartego wielomiany aproksymujące (16) przyjmują postać:

$$\Phi_0(x) = 1 \quad (18)$$

$$\Phi_1(x) = x - \frac{1}{n} \sum_{i=1}^n x_i \quad (19)$$

$$\Phi_2(x) = x^2 - \frac{\sum_{i=1}^n x_i^2 \Phi_1(x_i)}{\sum_{i=1}^n [\Phi_1(x_i)]^2} \Phi_1(x) - \frac{1}{n} \sum_{i=1}^n x_i^2 \quad (20)$$

$$\Phi_3(x) = x^3 - \frac{\sum_{i=1}^n x_i^3 \Phi_2(x_i)}{\sum_{i=1}^n [\Phi_2(x_i)]^2} \Phi_2(x) - \frac{\sum_{i=1}^n x_i^3 \Phi_1(x_i)}{\sum_{i=1}^n [\Phi_1(x_i)]^2} \Phi_1(x) - \frac{1}{n} \sum_{i=1}^n x_i^3 \quad (21)$$

Można zauważyć, że wielomiany (18 ÷ 21) spełniają warunek ortogonalności (11) jedynie dla sygnału centrowanego ($\bar{x} = 0$). W przypadku, gdy przedział argumentów x zostanie wprowadzony tak, że wartość $\bar{x} \neq 0$ (co jest regułą przy podawaniu wartości argumentu dla krzywej udziału nośnego liniowego), wykonuje się operację centrowania. W tym celu wprowadza się pomocniczą wartość parametru:

$$x_{pom(i)} = x_i - \bar{x} \quad i = 1, 2, \dots, n \quad (22)$$

i uzyskuje się nowy zbiór wartości $\{x_{pom(i)}\}$, dla których można przeprowadzić aproksymację (gdyż $\bar{x}_{pom} = 0$). Otrzymuje się wówczas wielomian

$$f(x_{pom}) = a_0 + a_1 x_{pom} + a_2 x_{pom}^2 + a_3 x_{pom}^3 \quad (23)$$

który łatwo można doprowadzić (wykorzystując (22)) do postaci ostatecznej:

$$f(x) = t_p(x) = ax^3 + bx^2 + cx + d \quad (24)$$

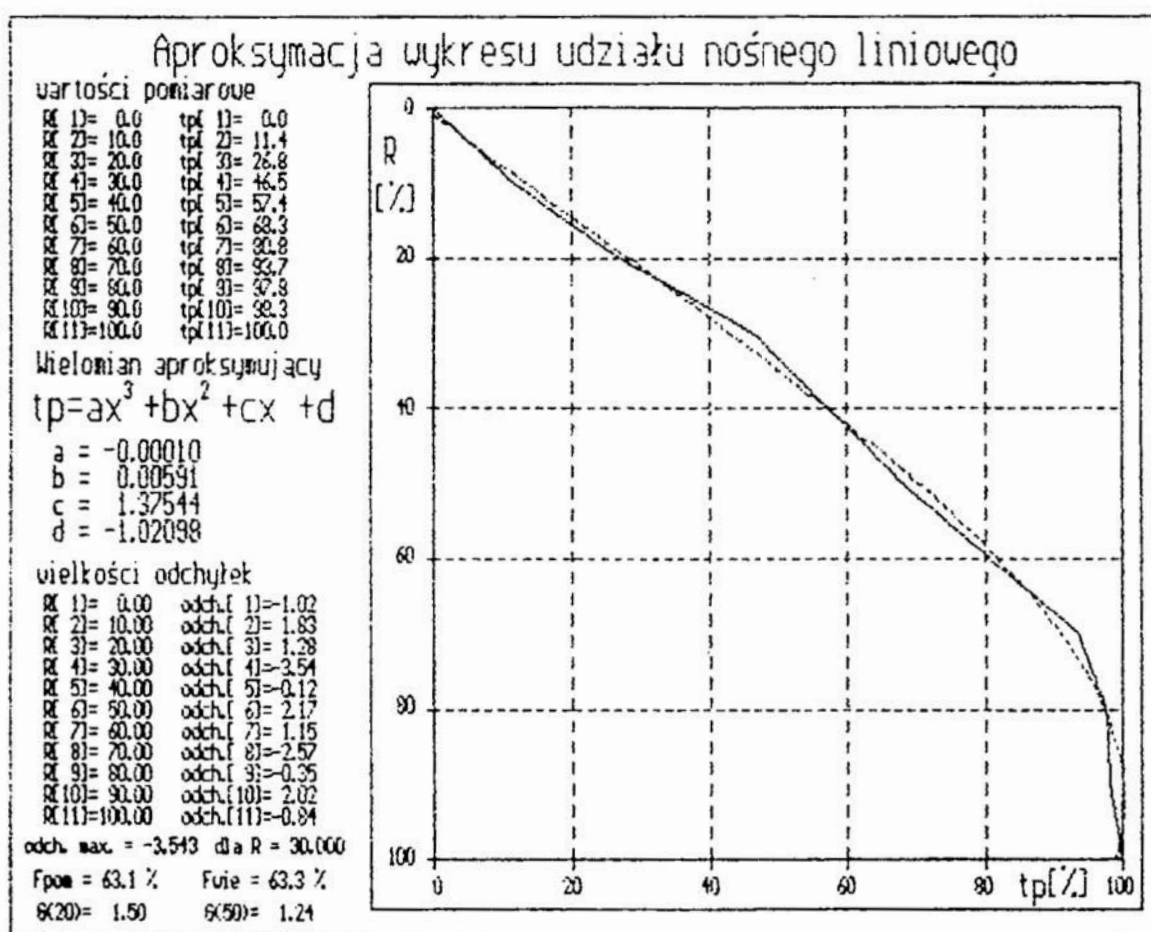
4. Praktyczne wykorzystanie metody aproksymacji krzywych nośności

Wyznaczenie wielomianu aproksymującego przedstawiającego krzywą t_p w postaci (24) umożliwi natychmiastową jej analizę. Możliwym staje się również otrzymanie pełnego jej przebiegu mimo nieznaności lub niedokładnego wyznaczenia jej części na podstawie kilku punktów pomiarowych. Jest to szczególnie istotne, gdy dysponuje się typowym profilografometrem nie pozwalającym na wyznaczenie wartości

t_p dla wartości $R = 5, 15$ i 25% wymaganych przez normy [3]. Opisana metoda umożliwia wyznaczenie orientacyjnych ich wartości. Przedstawiona metoda pozwala na wyznaczenie pełnego przebiegu $G_{(x)}$ i rozkładu rzędnych profilu $p(x)$. Znajomość ich przebiegu ma duże znaczenie przy badaniach własności użytkowych powierzchni, można bowiem na ich podstawie w pewnym stopniu przewidywać zachowanie się badanych powierzchni w trakcie przyszłej eksploatacji.

W trakcie wykonywania obliczeń aproksymacyjnych wyznaczone zostają dodatkowo wartości $G_{(20)}$ i $G_{(50)}$ oraz wielkości pól pod krzywą nośności (dla przebiegu rzeczywistego i aproksymowanego) charakteryzujące przebieg i kształt badanych przebiegów krzywej t_p .

Przykładowe wyniki obliczeń analizy krzywej t_p powierzchni obrabianej nagniataniem krążkiem o średnicy zewnętrznej $R_z = 40$ mm, promieniu krzywizny zarysu $R_k = 2,5$ mm z parametrami: siła nagniatania $F = 0,5$ kN, $f = 0,102$ mm/obr, $v = 8,8$ m/min, przedstawiono na rys.3.



Rys.3. Analiza przebiegu krzywej nośności liniowej na podstawie programu opisującego profil powierzchni próbki nagniatanej krążkiem $R_k = 2,5$ mm, $R_z = 40$ mm z parametrami $F = 0,5$ kN, $f = 0,102$ mm/obr, $v = 8,8$ m/min
 Fig.3. Analysis of the course bearing linear curve based on the programme describing the roughness profile of sample burnished by roll $R_k = 2,5$ mm, $R_z = 40$ mm with parameters $F = 0.5$ kN, $f = 0.102$ mm/obr, $v = 8.8$ m/min

5. Podsumowanie

Przedstawiona w niniejszym opracowaniu metoda nie może konkurować pod względem uniwersalności z współczesnymi metodami kompleksowej analizy stereometrii powierzchni przy użyciu współczesnej techniki obliczeniowej (profilografometrów sprzężonych z komputerami). Przeprowadzenie zamiany otrzymanego przebiegu profilu (najczęściej sygnał elektryczny) na sygnał cyfrowy, to znaczy przeprowadzenie jego dyskretyzacji za pomocą przetworników analogowo - cyfrowych, pozwala analizować cyfrowo nie wyniki wcześniejszych pomiarów lecz sam przebieg profilu powierzchni. Możliwym staje się wyznaczenie wszystkich jego parametrów na drodze obliczeń numerycznych jak również jego analiza wybranymi funkcjami (np. metodą funkcji autokorelacji, czy gęstości widmowej mocy [1, 2]). Metody te można stosować jednak tylko tam, gdzie dysponuje się odpowiednim, skomputeryzowanym sprzętem pomiarowym. Szersze rozpowszechnienie tego typu zestawów pomiarowych pomimo błyskawicznie postępującej komputeryzacji musi jednak trochę potrwać.

Tym samym zaproponowana przez autorów metoda analizy przebiegu krzywej udziału nośnego liniowego, aczkolwiek zdecydowanie mniej uniwersalna (ograniczona jedynie do parametru t_p i otrzymanych przebiegów $G_{(x)} \approx p(x)$) może również znaleźć zastosowanie. Przy użyciu najprostszego sprzętu liczącego staje się możliwe wykonanie błyskawicznej analizy i wykreślenie krzywej t_p na podstawie kilku (co najmniej czterech) punktów pomiarowych, co może być wykorzystane w praktyce.

6. Literatura

1. Górecka R., Polański Z.: Metrologia warstwy wierzchniej. WNT, Warszawa 1983.
2. Nowicki B.: Struktura geometryczna - chropowatość i falistość powierzchni. WNT, Warszawa 1990.
3. PN - 87/M - 04250, PN - 87/04251, PN - 87/M - 04252/02
4. Guter R. S., Owczyński W.: Matematyczne opracowanie wyników doświadczeń. PWN, Warszawa 1983.
5. Sakulewicz F. J., Miszyn Ł. K., Kubinowa E. J.: Metod analiticzeskowo wyrażenija opornych kriwych. Nowyje metody ispytaniija i obrabotki materiałow. Nauka i miechanika. Mińsk 1976.
6. Tubielewicz K., Zaborski A.: Metoda matematycznego modelowania krzywych udziału nośnego liniowego. Przegląd Mechaniczny 19/1993.

MATHEMATICAL ANALYSIS OF BEARING LINEAR CURVES

Summary

This paper presents the analysis of results of measurement parameters of surface profile with modern estimate technology. The method of approximation of measurement results t_p with using orthogonal polynomial was shown. This approximation polynomial is used for mathematical and functional results analysis of measurement parameters of surface layer.

WZÓR NA UMOWNY KĄT ŚCINANIA DLA TOCZENIA WIBRACYJNEGO

Wyprowadzono teoretyczne równania opisujące reakcje składowych siły skrawania na dynamiczne zmiany warunków skrawania. Pozwala to na ocenę wpływu parametrów drgań narzędzia na umowny kąt ścinania oraz umożliwia wyliczenie zarówno wartości liczbowej składowych siły skrawania czynnej na powierzchni natarcia, jak i przesunięcia fazowego między siłą skrawania a grubością warstwy skrawanej.

1. Wstęp

Spośród różnych sposobów wpływania na kształt i rozdrabnianie wióra skrawanie wibracyjne należy do tej grupy, której stosowanie zapewnia niezawodne dzielenie wióra na odcinki niezależnie od materiału obrabianego i w dowolnych warunkach skrawania. Nałożenie drgań na ruch posuwowy wywołuje okresowe zmiany głównie grubości warstwy skrawanej i kątów narzędzia w układzie roboczym. Ponadto wychodzenie ze strefy skrawania narzędzia prowadzi do wzrostu - w porównaniu ze skrawaniem zwykłym - intensywności procesów adsorpcji i reakcji chemicznych powierzchni stykowych narzędzia, przedmiotu obrabianego i spływającego wióra z atomami lub molekułami otaczającego ośrodka: gazu, cieczy lub ciała stałego. Dzięki temu stopień odkształcenia plastycznego warstwy skrawanej przekształcanej w wiór podlega okresowej zmianie. Podobnej zmianie ulega również wartość umownego kąta ścinania. Stąd też podjęto próbę wyprowadzenia tego ważnego, dla procesu powstawania wióra, wzoru.

2. Obliczenie sił działających na ostrze

Aby uprościć rozważania, założono płaski układ sił działających na ostrze w warunkach skrawania swobodnego prostokątnego. Uwzględniając, iż podczas obróbki stali napężenie styczne w płaszczyźnie ścinania τ_{shi} ma w przybliżeniu stałą wartość [1], z zależności czysto geometrycznych wynikających z rys. 1 znaleźć można siłę styczną w płaszczyźnie ścinania:

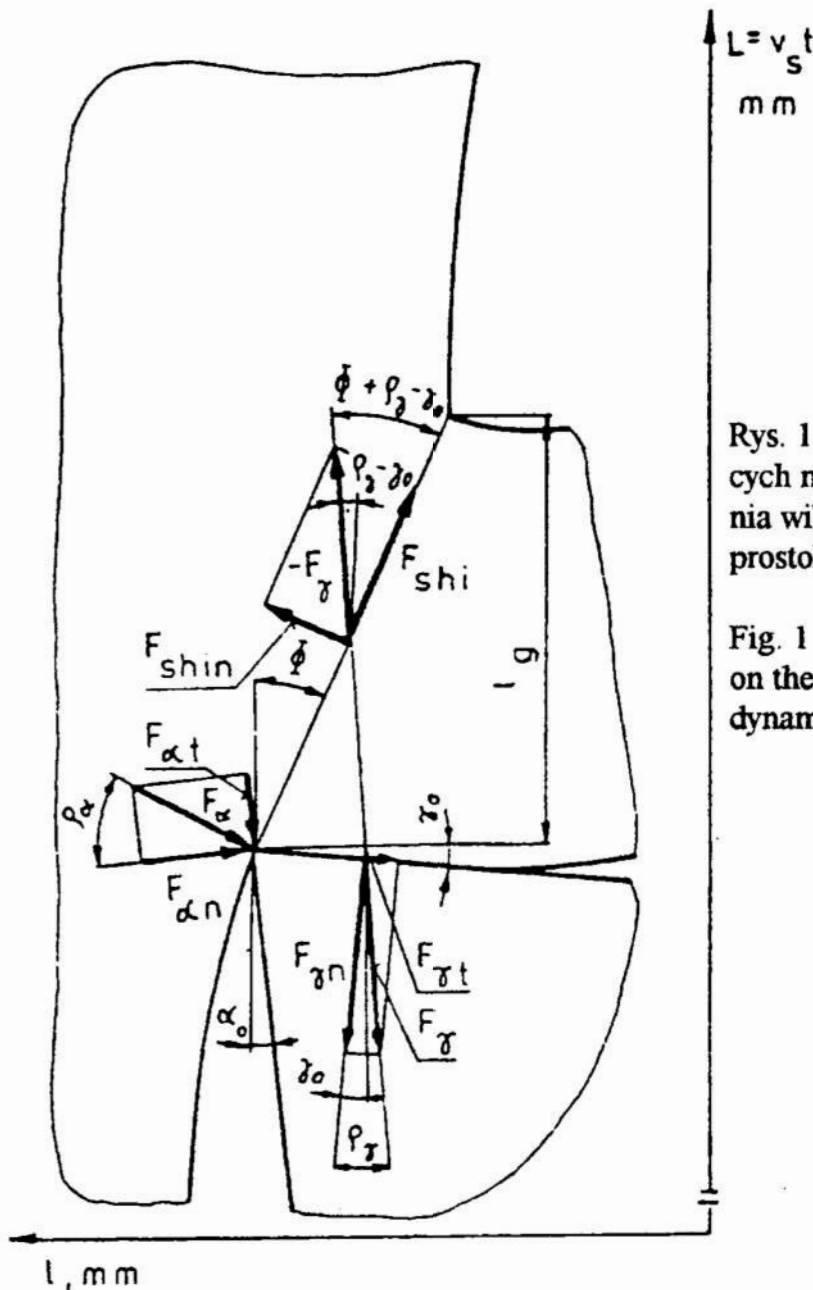
$$F_{shi} = \tau_{shi} b \frac{l_g}{\cos \phi} \quad (1)$$

a także składowa siły skrawania działająca na powierzchni natarcia:

$$F_\gamma = \frac{F_{shi}}{\cos(\phi + \rho_\gamma - \gamma_0)} \quad (2)$$

w których: b - szerokość warstwy skrawanej, l_g - długość rzutu linii ścinania na kierunek ruchu głównego, γ_0 - główny kąt natarcia, ρ_γ - kąt tarcia wióra o powierzchnię natarcia, ϕ - kąt ścinania.

*Instytut Technologii Maszyn i Automatykacji Politechniki Wrocławskiej, Wrocław



Rys. 1 Schemat rozkładu sił działających na ostrze w warunkach toczenia wibracyjnego, swobodnego, prostokątnego

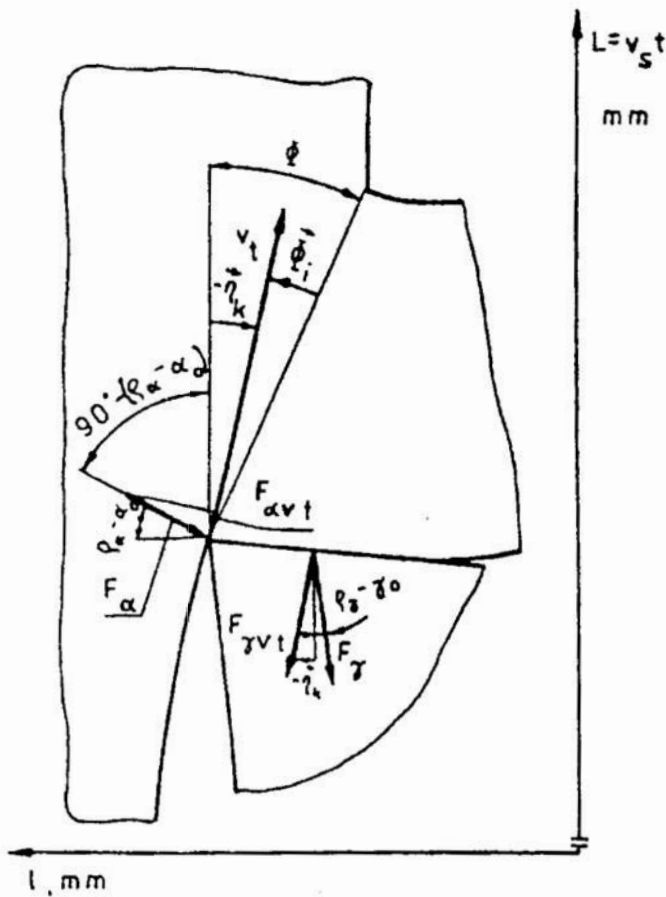
Fig. 1 Distribution of forces acting on the edge in free orthogonal dynamic turning

Aby obliczyć składowe $F_{\gamma c}$ i $F_{\gamma f}$ siły skrawania, należy wyprowadzić najpierw zależność umownego (chwilowego) kąta ścinania od parametrów drgań wymuszonych narzędzia. Można w tym celu posłużyć się hipotezą, według której umowny kąt ścinania powinien przyjmować takie wartości, aby praca skrawania osiągała minimum. Z uwagi na to, że w toczeniu wibracyjnym siła skrawania jest zmienna a ruch krzywoliniowy, elementarna praca skrawania ΔE stanowi iloczyn siły wypadkowej F_{vt} działającej wzdłuż wektora chwilowej prędkości skrawania v_t i elementarnego przesunięcia noża Δl :

$$\Delta E = F_{vt} \Delta l \quad (3)$$

raktując Δl jako stałą, minimum pracy skrawania wyraża się wzorem:

$$\frac{\delta F_{vt}}{\delta \phi_i} = 0 \quad (4)$$



Rys.2. Rozkład sił czynnych na powierzchni natarcia i przyłożenia ostrza na składowe działające w kierunku wektora chwilowej prędkości skrawania

Fig. 2. Distribution of active forces on the tool and relief faces into components acting along a vector of instantaneous cutting speed

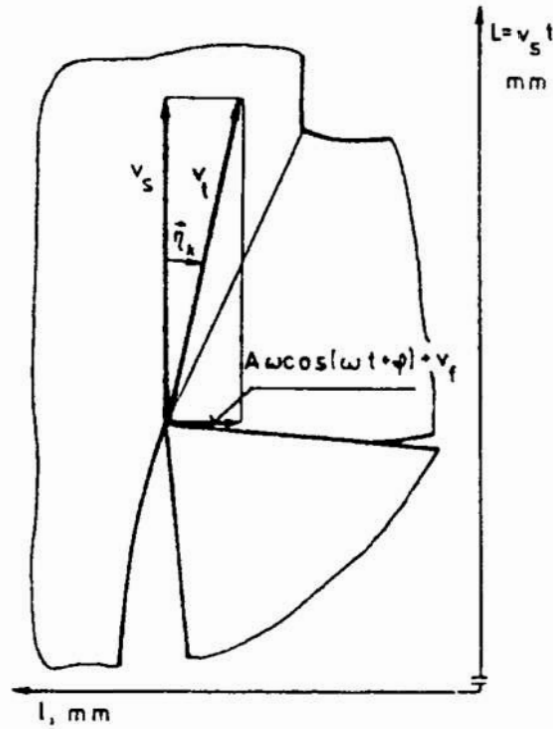
Z zależności trygonometrycznych wynikających z rys. 2 można znaleźć składową siły skrawania skierowaną wzdłuż wektora chwilowej prędkości skrawania v_f :

$$F_{vt} = F_\gamma \cos(p_\gamma - \gamma_o - \bar{\eta}_k) + F_\alpha \sin(p_\alpha + \bar{\eta}_k - \alpha_o) \quad (5)$$

gdzie F_α - składowa siły skrawania działająca na powierzchni przyłożenia. Po uwzględnieniu związków (1) i (2) oraz wykonaniu odpowiednich obliczeń, równanie (5) przyjmie postać:

$$F_{vt} = \tau_{shi} b \frac{l_g \cos(p_\gamma - \gamma_o - \bar{\eta}_k)}{\cos \phi \cos(\phi + p_\gamma - \gamma_o)} + F_\alpha \sin(p_\alpha + \bar{\eta}_k - \alpha_o) \quad (6)$$

Ponieważ kąt kierunku ruchu wypadkowego η_k może przybierać dodatnie i ujemne wartości w zależności od zgodności lub niezgodności zwrotów prędkości posuwu i drgań, przeto staje się on kątem skierowanym. W przyjętym prawoskrętnym układzie współrzędnych prostokątnych początkowe ramię kąta $\bar{\eta}_k$ stanowi wektor prędkości ruchu głównego v_s . Miarę stopniową tego kąta można otrzymać z prostej zależności wynikającej z rys. 3:



Rys.3. Zmienność kąta kierunku ruchu wypadkowego w toczeniu wibracyjnym

Fig.3. Variation of the resultant motion direction in dynamic turning

$$\operatorname{tg} \bar{\eta}_k = \frac{A \omega \cos(\omega t + \varphi) + v_f}{v_s} \wedge \bar{\eta}_k \in \left(-\frac{\pi}{2}, \frac{\pi}{2} \right) \quad (7)$$

gdzie: A - amplituda wymuszonych drgań narzędzia, v_f - prędkość posuwu, φ - przesunięcie fazowe między dwoma powierzchniami skrawania, które powstają w dwóch kolejnych obrotach przedmiotu, ω - częstość kołowa drgań.

Uwzględniając, że kąt ścinania stanowi sumę lub różnicę kątów skierowanych $\bar{\phi}_1$ i $\bar{\eta}_k$ (rys.2):

$$\phi = \bar{\phi}_1 - \bar{\eta}_k \wedge \phi \in \left(0, \frac{\pi}{2} \right) \quad (8)$$

równanie (6) można zapisać następująco:

$$F_{vt} = \tau_{shi} b \frac{l_g}{\cos(\bar{\phi}_1 - \bar{\eta}_k) \cos(\bar{\phi}_1 - \bar{\eta}_k + \rho_\gamma - \gamma_o)} \cos(\rho_\gamma - \gamma_o - \bar{\eta}_k) + F_\alpha \sin(\rho_\alpha + \bar{\eta}_k - \alpha_o) \quad (9)$$

Z ostatniego wzoru, po zróżniczkowaniu względem $\bar{\phi}_1$ wynika zależność [2]:

$$\frac{d}{d\bar{\phi}_1} \cos(\bar{\phi}_1 - \bar{\eta}_k) \cos(\bar{\phi}_1 - \bar{\eta}_k + \rho_\gamma - \gamma_o) + l_g \sin(2\bar{\phi}_1 - 2\bar{\eta}_k + \rho_\gamma - \gamma_o) = 0 \quad (10)$$

Aby obliczyć występującą we wzorze (10) pochodną cząstkową funkcji l_g względem $\bar{\phi}_1$ trzeba najpierw znaleźć równanie tej funkcji. Można je wyznaczyć z następujących zależności trygonometrycznych wynikających z rys. 4 i 5.

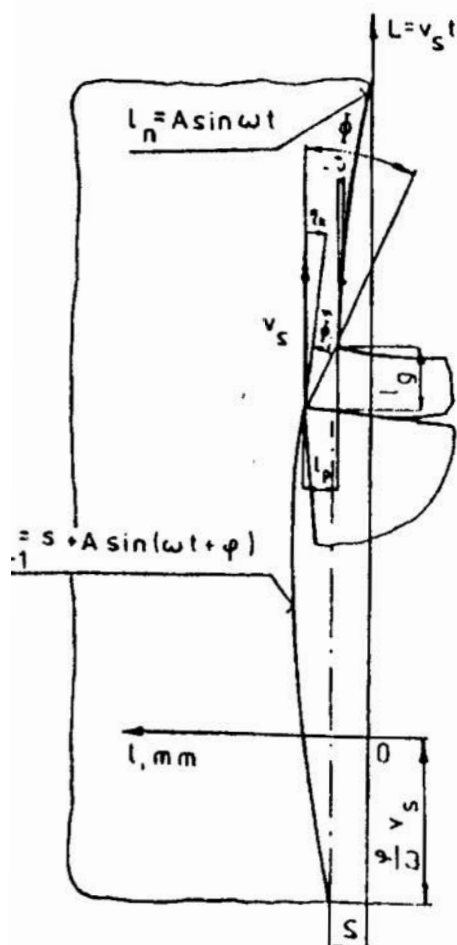
$$l_g = l_p \operatorname{ctg}(\bar{\phi}_i - \bar{\eta}_k) \quad (11)$$

oraz

$$l_p = s + A \sin(\omega t + \varphi) - A \sin \omega \left(t + \frac{l_g}{v_s} \right) \quad (12)$$

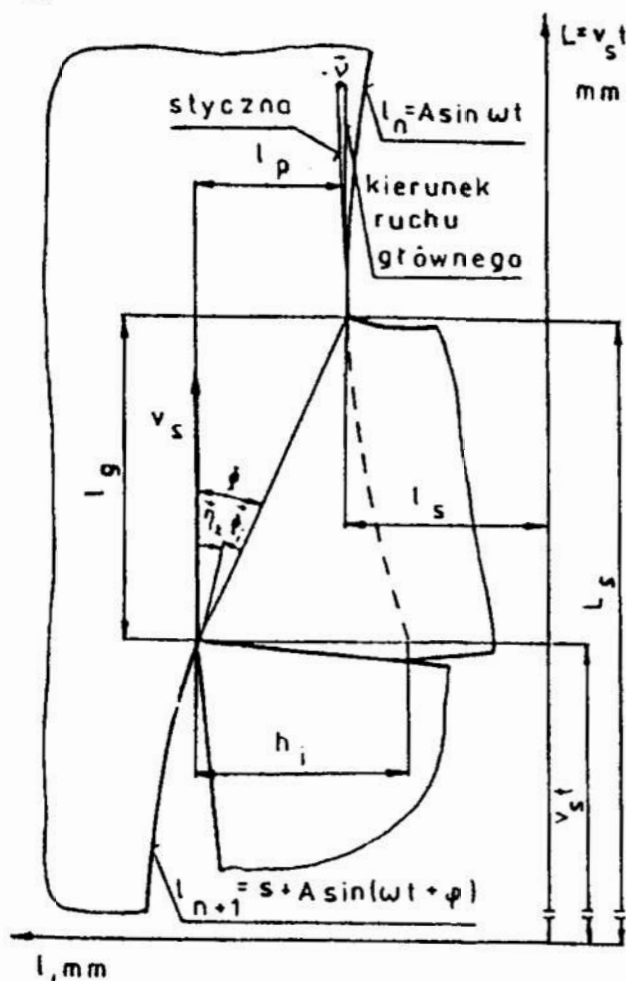
w których: s - posuw na obrót, l_p - długości rzutu linii ścinania na kierunek ruchu osuwowego. Korzystając z różnicy sinusów związek (12) można zapisać w postaci:

$$l_p = s + 2A \cos \left[\omega \left(t + \frac{l_g}{2v_s} \right) + \frac{\varphi}{2} \right] \sin \left(\frac{\phi}{2} - \frac{\omega l_g}{2v_s} \right) \quad (13)$$



Rys.4. Rysunek pomocniczy do wyznaczania rzutu l_g linii ścinania na kierunek ruchu głównego

Fig.4. Scheme used in determining projection l_g of the shearing line onto the principal motion direction



Rys.5. Rysunek pomocniczy do wyznaczania kąta \bar{V}

Fig.5. Scheme for determining angle \bar{V}

Podstawiając ostatni wzór w równanie 11 uzyskuje się zależność, o którą chodziło:

$$l_g = \left\{ s + 2A \cos \left[\omega \left(t + \frac{l_g}{2v_s} \right) + \frac{\varphi}{2} \right] \sin \left(\frac{\varphi}{2} - \frac{\omega l_g}{2v_s} \right) \right\} \operatorname{ctg}(\bar{\phi}_i - \bar{\eta}_k) \quad (14)$$

W wyniku zrózniczkowania funkcji l_g względem $\bar{\phi}_i$ dostaje się wyrażenie [1]:

$$\frac{d l_g}{d \bar{\phi}_i} = \frac{-l_p}{\left[A \frac{\omega}{v_s} \cos \omega \left(t + \frac{l_g}{v_s} \right) \operatorname{ctg}(\bar{\phi}_i - \bar{\eta}_k) + 1 \right] \sin^2(\bar{\phi}_i - \bar{\eta}_k)} \quad (15)$$

Równanie (15) wstawione do wzoru (10) daje [1]:

$$A \frac{\omega}{v_s} \cos \omega \left(t + \frac{l_g}{v_s} \right) - \operatorname{ctg}(2\bar{\phi}_i - 2\bar{\eta}_k + \rho_\gamma - \gamma_o) = 0 \quad (16)$$

Wzór ten można uprościć, jeśli uwzględni się, że pierwszy składnik uzyskanego równania stanowi pochodną funkcji opisującej ślad powierzchni obrabianej w punkcie przecięcia się go z linią ścinania. Łatwo sprawdzić, że jeśli ruch tego punktu zmienia się według prawa (rys5.):

$$l_s = A \sin \omega \frac{L_s}{v_s} \quad (17)$$

to

$$(l_s)'_{L_s} = A \frac{\omega}{v_s} \cos \omega \frac{L_s}{v_s} \quad (18)$$

w których L_s - odcięta punktu przecięcia się śladu powierzchni obrabianej z linią ścinania, l_s - rzędna punktu przecięcia się śladu powierzchni obrabianej z linią ścinania. Mając na uwadze, że odcięta rozważanego punktu wyraża się wzorem:

$$L_s = v_s t + l_g \quad (19)$$

można również napisać:

$$(l_s)'_{L_s} = \operatorname{tg} \bar{v} = A \frac{\omega}{v_s} \cos \omega \left(t + \frac{l_g}{v_s} \right) \wedge \bar{v} \in \left(-\frac{\pi}{2}, \frac{\pi}{2} \right) \quad (20)$$

przy czym występujący w ostatniej zależności symbol \bar{v} oznacza kąt, jaki styczna do śladu powierzchni obrabianej w punkcie przecięcia się go z linią ścinania tworzy z dodatnim kierunkiem osi odciętych (rys.5), zgodnym z kierunkiem ruchu głównego. Po wstawieniu zależności (20) do (16) otrzymuje się:

$$\operatorname{tg} \bar{v} - \operatorname{ctg}(2\bar{\phi}_i - 2\bar{\eta}_k + \rho_\gamma - \gamma_o) = 0 \quad (21)$$

a stąd po wprowadzeniu funkcji sinus i cosinus równanie (16) wyrazi się formułą:

$$\cos \bar{v} \cos(2\bar{\phi}_i - 2\bar{\eta}_k + \rho_\gamma - \gamma_o) - \sin \bar{v} \sin(2\bar{\phi}_i - 2\bar{\eta}_k + \rho_\gamma - \gamma_o) = 0 \quad (22)$$

Ponieważ lewa strona równania (22) odpowiada cosinusowi sumy dwóch kątów, przeto można je zapisać w postaci równania trygonometrycznego elementarnego względem sumy kątów: $\bar{v} + 2\bar{\phi}_i - 2\bar{\eta}_k + \rho_\gamma - \gamma_o$. a mianowicie:

$$\cos(\bar{v} + 2\bar{\phi}_i - 2\bar{\eta}_k + \rho_\gamma - \gamma_o) = 0 \quad (23)$$

Rozwiązanie tego równania wyraża się związkiem:

$$\bar{v} + 2\bar{\phi}_i - 2\bar{\eta}_k + \rho_\gamma - \gamma_o = \pm \frac{\pi}{2} + k\pi \quad (24)$$

gdzie: k jest dowolną liczbą całkowitą. Biorąc w rachubę rozwiązanie podstawowe równania trygonometrycznego dochodzi się do miary głównej umownego kąta ścinania:

$$\bar{\phi}_i = \frac{\pi}{4} - \frac{1}{2}(\rho_\gamma - \gamma_o) + \bar{\eta}_k - \frac{1}{2}\bar{v} \quad (25)$$

Wzór ten po uprzednim przekształceniu

$$\bar{\phi}_i = \frac{\pi}{4} - \frac{1}{2}(\rho_\gamma - \gamma_o) + \text{arctg} \frac{v_f + A \omega \cos(\omega t + \varphi)}{v_s} - \frac{1}{2} \text{arctg} A \frac{\omega}{v_s} \cos \left(t + \frac{l_g}{v_s} \right) \quad (26)$$

pozwała ocenić wpływ parametrów drgań wymuszonych narzędzia i parametrów skrawania na umowny kąt ścinania.

3. Podsumowanie

W wyniku teoretycznych rozważań dynamiki toczenia wibracyjnego, znaleziono ważną dla przebiegu obróbki zależność (26). Otrzymany wynik:

a) potwierdza rezultaty prób skrawania Fielda i Merchanta [3], [4], którzy wykazali, że zmiana falistości powierzchni obrabianej wywołuje zmiany kąta ścinania;

b) pozostaje w zgodności z wynikami rozważań Albrechta [3] odniesionymi do toczenia normalnego, w związku z oddziaływaniem zewnętrznej wyjściowej powierzchni obrabianej na kąt ścinania;

c) jest zbieżny z opinią Chakraborti i Bhattacharyyi [4] o wpływie prędkości skrawania na kąt ścinania.

Mając na uwadze, że w toczeniu wibracyjnym chwilową grubość warstwy skrawanej h_t wyznacza ekstrapolacja krzywoliniowego śladu powierzchni obrabianej (rys.5) oraz jeśli się uwzględni złożony wpływ tej grubości na pole płaszczyzny ścinania A_{shi} , od którego zależy składowa siły skrawania, działająca na powierzchni natarcia F_γ , to dochodzi się do przekonania, że rozważane drgania narzędzia mogą być przyczyną powstawania przesunięcia fazowego ψ między siłą skrawania a grubością warstwy skrawanej. Można je wyliczyć z proporcji:

$$\frac{l_g}{\lambda} = \frac{\psi}{2\pi}$$

a ponieważ odległość λ , na jaką określona faza drgań rozchodzi się w czasie jednego okresu wyraża się wzorem:

$$\lambda = v_s T = v_s \frac{2\pi}{\omega}$$

wiec przesunięcie fazowe:

$$\psi = \frac{l_g \omega}{v_s}$$

4. Literatura

- [1] LOLADZE T.N., Pročnost' i iznosostojkost' režuščego instrumenta, Moskva, Mašinostroenie, 1982.
- [2] CHOROSZY B., Mechanizm zużywania się ostrzy podczas skrawania wibracyjnego. Pr. nauk. Inst. Technol. Masz. i Aut. PWr. Nr 48, Seria: Monografie Nr 13, Wrocław 1991.
- [3] ALBRECHT P., Dynamics of the Metal Cutting Process, Trans. ASME 1965, Nr 4.
- [4] CHAKRABORTI S. K., BHATTACHARYYA A., The Cutting Characteristics of Oscillating Cutting Tool. I. Instn. Engrs. Mech. Engng. Div., 1968, 48, No. 9.

IMAGINARY SCHEAR ANGLE DURING VIBRATORY CUTTING

Summary

In the paper, describing reactions of component cutting forces, theoretical equations were derived. It allows to evaluate an effect of tool vibration parameters for conventional shear angle and makes possibilities for evaluation as well as numerical value of component of cutting forces active at friction face, and phase shift between cutting force and thickness of cutting layer.

WYBORY ZŁOŻONE I OCENY ROZMYTE

W pracy zebrano określenia i definicje algebry rozmytej. Wykazano, że istnieje ścisły związek pomiędzy sposobem wyznaczania oceny złożonej a definiowaniem algebry rozmytej. Wykazano ponadto, że sumowanie ocen odpowiada operacji splotu i wskazano na komplikacje obliczeniowe.

Wstęp

W procesie projektowania istnieją fazy wyboru wariantów. Należy wybrać wariant najlepszy ze skończonego zbioru możliwości. Przykładami mogą być: wybór koncepcji konstrukcji, wybór oferty licencyjnej, wybór dostawcy podzespołów na podstawie danych katalogowych. Zwykle należy uwzględnić wiele kryteriów, zwanych też kryteriami oceny, niezależnie od ograniczeń nierównościowych, nałożonych na właściwości obiektu wyboru. Występują w procesie projektowania kryteria, takie jak wygoda, estetyka czy bezpieczeństwo, których w kategoriach mierzalnych czy probabilistycznych wyrazić nie można. Kryteria takie mają charakter rozmyte.

Istnieją trzy ujęcia problemu polioptymalizacji [6]:

1. Najogólniejsze to takie, kiedy staramy się wyznaczyć w zbiorze wariantów poprawnych podzbiór wariantów polioptymalnych a więc równoważnych względem przyjętego kryterium wektorowego. Wybór tak dokonany nie jest jednoznaczny a czasami gdy zbiór polioptymalny jest liczny także nieprzydatny.

2. Aby uniknąć niejednoznaczności rozwiązania można sformułować zadanie optymalizacji wielokryterialnej;

- tworzy się zastępcze skalarnie kryterium optymalizacji F:

$$F = F(k_1, k_2, \dots, k_I) \quad (1)$$

k - ocena.

Często jest to kryterium sztuczne, nie mające interpretacji fizycznej. Można przyjąć, że najbardziej przydatna jest liniowa postać kryterium:

$$F = \sum_{i=1}^I w_i \cdot u_i(k_i) \quad (2)$$

u - użyteczność, w - waga..

*- Instytut Elektroniki Wyższej Szkoły Inżynierskiej w Koszalinie
ul. Partyzantów 17 75-711 Koszalin

3. Najczęstszym i najbardziej arbitralnym sposobem jest wybranie kryterium "najważniejszego" i rozwiązanie zadania optymalizacji jednokryterialnej przy zamianie pozostałych kryteriów na ograniczenia.

Traktowanie oceny jako liczby rozmytej przyjmuje się za naturalne (najbardziej zbliżone do naturalnego odczuwania). Mówimy, że ocena jest parą

$$O = \langle n, \mu \rangle \quad (3)$$

gdzie : n - nośnik oceny, μ - funkcja przynależności oceny.

Nośnik oceny to zbiór tych parametrów użytkowych konstrukcji, dla których funkcja przynależności do oceny jest większa od zera. Nośnik pojedynczej oceny nie wypełnia zwykle całkowicie zbioru wszystkich możliwych wartości ocenianej cechy (własności, parametru). Sens oceny polega na istnieniu w zbiorze wartości ocenianej cechy jeszcze innych ocen. Ocenianie powinno posiadać walor porządkujący. Struktura tworzona przez relację oceniania w zbiorze ocen nie może być cykliczna. Dla ocen ostrych takich, że nośnik jest jednoelementowy a funkcja przynależności równa 1 stosuje się dobrze opracowane metody porządkowania i normowania ocen. Dla ocen rozmytych, których nośnik jest zbiorem będącym dziedziną funkcji przynależności, brak takiej jednoznaczności. Opracowanie zawiera przegląd pojęć i metod stosowanych przy ocenianiu rozmytym. Wskazuje się szczególnie na możliwość i sposób wyznaczania oceny wypadkowej będącej sumą ważoną ocen cząstkowych i oceny wypadkowej z użyciem kryterium maxmin.

1. Podstawowe pojęcia i ich definicje

Podstawowe dla każdej algebry definicje przecięcia i sumy w przypadku teorii zbiorów rozmytych nie są zbyt kategorycznie określone. Różne koncepcje tych definicji można sprowadzić do dwóch krańcowych, "twardej" i "miękkiej" oraz pośrednich [7]:

-definicja "twarda" - ZADECHA:

$$\begin{aligned} \mu_{A \cap B}(x) &= \mu_A(x) \otimes \mu_B(x) = \min[\mu_A(x), \mu_B(x)] \\ \mu_{A \cup B}(x) &= \mu_A(x) \oplus \mu_B(x) = \max[\mu_A(x), \mu_B(x)] \end{aligned} \quad (4)$$

-definicja "miękką" - BELLMANA - ZADECHA:[1]

$$\begin{aligned} \mu_{A \cap B}(x) &= \mu_A(x) \otimes \mu_B(x) = \mu_A(x) \cdot \mu_B(x) \\ \mu_{A \cup B}(x) &= \mu_A(x) \oplus \mu_B(x) = \mu_A(x) + \mu_B(x) - \mu_A(x) \cdot \mu_B(x) \end{aligned} \quad (5)$$

-liczba rozmyta:

$$A = \sum_{x \in R}^M \mu_A(x) / x \sum_{x \in R}^M - \text{suma mnogościowa } \mu_A(x) / x - \text{para } \langle \mu_A(x), x \rangle \quad (6)$$

-moc liczby rozmytej A:

$$M_A = \sum_{x \in A} \mu_A(x) \quad (7)$$

-wartość środkowa liczby rozmytej A:

$$m_A = \frac{\sum_{x \in A} x \cdot \mu_A(x)}{M_A} \quad (8)$$

-moc liczby rozmytej o ciągłej i całkowalnej funkcji przynależności:

$$M_A = \int_{c_A}^{d_A} \mu_A(x) \cdot dx \quad (9)$$

-wartość środkowa liczby rozmytej o ciągłej i całkowalnej funkcji przynależności:

$$m_A = \frac{\int_{c_A}^{d_A} \mu_A(x) \cdot dx}{M_A} \quad (10)$$

-wysokość liczby rozmytej:

$$h_A = \max_{x \in A} \mu_A(x) \quad (11)$$

subnormalna liczba rozmyta - wtedy gdy $h_A < 1$.

-normalizowana liczba rozmyta A:

$$\mu_A^N = \frac{\mu_A(x)}{h_A} \quad (12)$$

-mnożenie liczby rozmytej przez stałą p:

$$x \in A \wedge z = px \Rightarrow \forall_{z \in B} \mu_B(z) = \mu_A\left(\frac{x}{p}\right) \quad (13)$$

-funkcja przynależności sumy liczb rozmytych A+B:[4]

$$\mu_{A+B}(z) = \bigoplus_{z \in B} \{ \bigvee_{x \in A} [\mu_B(z-x) \otimes \mu_A(x)] \} \quad (14)$$

2. Wybory rozmyte

Ilustracją pewnej idei stosowania ocen rozmytych spełniającej zasadę rozszerzenia będą przykłady stosowania ocen rozmytych dla dwóch metod wyboru:

- metody maximin i ważony maximin,
- metody sumy ważonej.

2.1. Metoda maximin

W metodzie tej przyjmuje się, że poszczególne warianty rozwiązań konstrukcyjnych reprezentowane są przez wartość ich najsłabszego parametru. Metoda zakłada porównywalność parametrów a więc wymaga liniowego uporządkowania ocen i ich normalizacji. Sumaryczną użyteczność rozwiązania konstrukcyjnego określa się na podstawie najgorszego parametru konstrukcyjnego.

$$F_i = \min_j k_{i,j}^* \quad (15)$$

Za wariant najlepszy uważa się wariant o najwyższej ocenie najsłabszego parametru.

$$A^* = \{ a_i : \max_i \min_j k_{i,j}^* \} \quad (16)$$

gdzie A^* - wariant optymalny,

$a_i = \{k_{i,1}, \dots, k_{i,j}\}$ - warianty rozwiązania,
 $k_{i,j}^*$ - oceny znormalizowane.

Metoda maximin reprezentuje strategię pesymisty. Jest ona wykorzystywana w sytuacji gdy ogólna ocena wariantu zależy od najsłabszego parametru. Metoda może być uzupełniona o stosowanie wag.

$$A^* = \{ a_i : \max_i \min_j (w_j \cdot k_{i,j}^*) \} \quad (17)$$

gdzie: w_j - waga j-tego parametru.

W obydwu przypadkach wariant optymalny reprezentuje najwyższą sumaryczną w sensie minimaxowym użyteczność. Takiemu podejściu odpowiada definicja ZADECHA (twarda) przecięcia i sumy. Sumaryczna rozmyta użyteczność wariantu wynikająca z definicji sumy algebraicznej liczb rozmytych będzie także liczbą rozmytą o funkcji przynależności:

$$\mu_{A+B}(z) = \max_{z \in B} \{ \forall_{x \in A} \min[\mu_B(z-x); \mu_A(x)] \} \quad (18)$$

2.2. Metoda sumy ważonej.

W metodzie sumy ważonej przyjmuje się skalarną funkcję użyteczności wariantu konstrukcyjnego w postaci:

$$F_i = \sum w_j \cdot k_{i,j}^* \quad (19)$$

a najlepszy wariant określany jest jako:

$$A^* = \{ a_i : \max_i F_i \} \quad (20)$$

Metodę stosuje się wtedy, gdy pogorszenie jednego z parametrów może być zrekompenrowane przez polepszenie innego a współczynniki wagi są stałe na całym obszarze oceny [6]. Jeśli oceny parametrów konstrukcyjnych są rozmyte, wtedy sumaryczną użyteczność należy wyznaczać stosując we wzorze na sumę algebraiczną liczb rozmytych (14) operacje \otimes \oplus zgodne z definicją BELLMANA-ZADECHA (miękką) przecięcia i sumy.

$$\begin{aligned} \mu_A(x) \otimes \mu_B(x) &= \mu_A(x) \cdot \mu_B(x) \\ \mu_A(x) \oplus \mu_B(x) &= \mu_A(x) + \mu_B(x) - \mu_A(x) \cdot \mu_B(x) \end{aligned} \quad (21)$$

co można zapisać następująco:

$$\mu_{A+B}(z) = \oplus_{z \in B} [\forall_{x \in A} \mu_B(z-x) \cdot \mu_A(x)] \quad (22)$$

Stosując wybór oparty na sumie ważonej i miękkiej definicji przecięcia i sumy można wykorzystać uproszczenia dla ciągłej i całkwalnej funkcji przynależności

$$\mu_{A+B}(z) = \frac{\int_{c_A}^{d_A} \mu_B(z-x) \cdot \mu_A(x) \cdot dx}{\text{maximum} \left[\int_{c_A}^{d_A} \mu_B(z-x) \cdot \mu_A(x) \cdot dx \right]} \quad (23)$$

Spotyka się reprezentację "trapezową" liczby rozmytej.

$$A = \langle a_A, b_A, c_A, d_A \rangle \quad (24)$$

gdzie: c, d - odpowiednio lewa i prawa granica nośnika liczby A .

$\text{supp } A = [c_A, d_A]$ a, b - lewa i prawa granice przedziału, że:

$$x \in [a, b] \Rightarrow \mu_A(x) = 1$$

Dla reprezentacji trapezowej można zastosować następujące oszacowania:

$$m_A = \frac{(c_A + d_A)}{2}; \quad \mu_A(x) = \frac{1}{1 + p \cdot (x - m_A)^q} \quad (25)$$

dobierając p i q tak by $\mu_A(c_A) < 0,01$ i $\mu_A(d_A) < 0,01$ i obliczać funkcję przynależności sumy dwóch liczb rozmytych stosując całkę splotową (23).

Istnieje także reprezentacja liczby rozmytej wg DUBOIS-PRADE [2]:

$$A = \langle m_A, \alpha_A, \beta_A \rangle \quad (26)$$

gdzie: - m_A - wartość środkowa liczby rozmytej A ,

 - α_A - dyspersja lewostronna,

 - β_A - dyspersja prawostronna.

Wtedy suma algebraiczna takich liczb rozmytych $A+B$:

$$\begin{aligned} A &= \langle m_A, \alpha_A, \beta_A \rangle; \quad B = \langle m_B, \alpha_B, \beta_B \rangle \\ A + B &= \langle m_A + m_B, \alpha_A + \alpha_B, \beta_A + \beta_B \rangle \end{aligned} \quad (27)$$

Zakończenie

Zbadano możliwość zastosowania ocen rozmytych do wyborów wielokryterialnych. Stwierdzono, że jest możliwe rozszerzenie znanych metod wielokryterialnych wyborów o oceny rozmyte. Analiza wykazała, że nie jest celowe wskazanie czy też opracowanie procedur uniwersalnych do rozmytych wyborów wielokryterialnych. Istnieje silna zależność pomiędzy stosowaną metodą wyboru wielokryterialnego (co do postaci i sensu skalarnego kryterium wyboru) i samą definicją liczby rozmytej służącej jako ocena.

Podstawowe dla każdej algebry definicje przecięcia i sumy w przypadku teorii zbiorów rozmytych nie są zbyt kategorycznie określone. Różne koncepcje tych definicji można sprowadzić do dwóch krańcowych: "twardej" i "miękkiej" oraz pośrednich. Przyjęcie każdej z tych definicji ma kardynalne konsekwencje obliczeniowe zwłaszcza przy określaniu funkcji przynależności sumy algebraicznej dwu liczb rozmytych. Operacja ta jest operacją splotu w sensie całki SUGENO [5]. Należy ją definiować odrębnie dla każdej z metod polioptymalizacji.

Można wykazać, że zaproponowane procedury sumowania użyteczności rozmytych zachowują swe właściwości także w stosunku do użyteczności ostrych dla wyborów opartych o sumę ważoną. Oczywiście, jeżeli oceny ostre traktować się będzie jako rozmyte o zerowym rozmyciu. To znaczy wtedy, gdy użyteczności ostrej przypisze się funkcje przynależności w postaci:

$$\begin{aligned} \mu(u) &= 1 \text{ dla } u = u_i \\ \mu(u) &= 0 \text{ dla } u \neq u_i \end{aligned} \tag{28}$$

Ważną rolę w praktycznych zastosowaniach wydaje się odgrywać trapezowa reprezentacja liczby zespolonej. Stosowanie jej ułatwia fazę ekspertowego definiowania ocen. Postać ta ze względu na niegładką funkcję przynależności jest niewygodna przy obliczaniu całek splotowych. Przedstawione w opracowaniu całkowanie oszacowanie trapezowej funkcji przynależności nie wydaje się w pełni zadowalające.

Literatura:

1. Tarnowski W.: Wspomaganie wyboru i optymalizacji wielokryterialnej za pomocą mikrokomputera. Prace Naukowe Instytutu Konstrukcji i Eksploatacji Maszyn Politechniki Wrocławskiej. zeszyt 56, ss.398-401, Wrocław 1989. (Materiały Międzynarodowej Konferencji Naukowej MASAD88).
2. Zadeh L. A. : Fuzzy sets. Inf. and Control. 1965 vol.8 ss.338-353. Zadeh L. A. : Fuzzy logic and approximate reasoning. "Synthese". 1975 vol. 30 pp.407-428.

3. Bellman R. Zadeh L.A. "Local and fuzzy logic". w: Modern uses of multiple valued logic. Epstein D. (red) Boston: Reidel 1977.
4. Kacprzyk J.: Zbiory rozmyte w analizie systemowej Państwowe Wydawnictwo Naukowe 1986
5. Dubois D. and Prade. H.: "Fuzzy sets and systems" Academic Press, New York 1980.
6. Sugeno M., Imaoka H. Generalised truth value in truth qualification. 4-th Report Working Group on Fuzzy Sets. Tokyo 1978.
7. Gupta M.N. Ragade R.K. Yager R.R. (red) Advances in fuzzy set theory and applications. Amsterdam: North-Holland 1979.

The complex variants and fuzzy gauges Summary

The definitions and qualifications of the fuzzy algebra are presented in the paper. There have been proved that there is exact connection between the way of marking the complex value determining fuzzy algebra. It was also proved that the value addition conforms with the tangle operation and the calculation complications are pointed act.

WPROWADZENIE DO WNIOSKOWANIA ROZMYTEGO W ZASTOSOWANIU DO ANALIZY CECH PROCESÓW SZLIFOWANIA

W inteligentnych systemach, wśród wielu wzorców wnioskowania takich jak: wnioskowanie probabilistyczne, rozmyte, indukcyjne i dedukcyjne, na szczególną uwagę zasługuje wnioskowanie rozmyte. Dla wielu właściwości układu technologicznego i cech charakteryzujących procesy szlifowania nie można określić wyraźnych granic klasyfikacji tych cech np. granic klasyfikacji jakościowej. Zbiorami, które nie mają ostro określonych granic są np.: zbiór dopuszczalnych parametrów szlifowania, okres zachowania określonych właściwości narzędzia ściernego, zbiór parametrów optymalnych procesu szlifowania. Zbiory te są zbiorami rozmytymi i dlatego naturalnym staje się dążenie do wykorzystania teorii takich zbiorów.

1. Probabilistyczne i rozmyte cechy procesów szlifowania

W procesach szlifowania, zwłaszcza precyzyjnego, liczne zjawiska o małym często znaczeniu w obróbce skrawaniem nabierają znaczenia decydującego o przebiegu i efektach procesu. Takimi zjawiskami są [2, 3]: nieciągłość procesu i zmienna częstość tworzenia mikrowiórów, ciepłne i mechaniczne sublokalne odkształcenia narzędzi i materiału obrabianego, zwłaszcza w narzędziach o podwyższonej podatności, losowe zmiany topografii powierzchni aktywnej narzędzia, zmienność warunków mikroskrawania wzdłuż drogi pracy każdego z ostrzy stykających się z obrabianym materiałem.

We wszystkich sposobach obróbki zagłębienie naroża ostrza w materiał jest zmienne wzdłuż toru skrawania. Jest to skutek wielu czynników, takich jak [2]: zmienność zagłębienia nominalnego, zależna od kinematycznych cech metody, nierówność powierzchni przedmiotu w strefie obróbki, podatność lokalna materiału obrabianego i ziaren ściernych, drgania narzędzia i ziaren ściernych, znaczne lokalne zróżnicowanie przyrostów temperatur, lokalne zróżnicowanie i strefowa zmienność właściwości materiału obrabianego oraz mikronieciągłość procesu tworzenia wiórów. Te cechy procesu są przyczyną nie tylko dużej losowości procesu, ale także i tego, iż dostępna informacja o procesie jest nieściśła.

W klasycznym podejściu, gdy określa się właściwości układu technologicznego - przedmiotu, obrabiarki, narzędzia, płynu obróbkowego i oprzyrządowania, to rozumie się pod tym pojęciem ściśle określone parametry, pozwalające jednoznacznie oddzielić jeden zbiór właściwości od innych.

* Wyższa Szkoła Inżynierska w Koszalinie, ul. Raławicka 15/17, 75-620 Koszalin

W interpretacji matematycznej oznacza to, iż pewna właściwość określa zbiór A obiektów spełniających tę właściwość. Zbiór A posiada jednoznaczne granice oddzielające elementy należące do tego zbioru od innych. Można to zapisać następująco:

$$\mu_A(x) = \begin{cases} 1, & x \in A \\ 0, & x \notin A \end{cases}$$

Dla wielu właściwości trudno jest określić granicę dzielącą elementy spełniające jej warunki od elementów, które ich nie spełniają. Przykładami takich zbiorów mogą być [2]: zbiór dopuszczalnych parametrów szlifowania, okres zachowania określonych właściwości narzędzia ściernego, zbiór parametrów optymalnych procesu szlifowania. Zbiory te są zbiorami rozmytymi i dlatego naturalnym staje się dążenie do wykorzystania teorii takich zbiorów.

2. Cechy i zasady wnioskowania przybliżonego

Zbiór rozmyty jest charakteryzowany przez funkcję $\mu_A(x)$, nazywaną funkcją przynależności, przekształcającą elementy przestrzeni X w liczby z zakresu $0..1$ (znormalizowany zbiór rozmyty), odpowiadające prawdopodobieństwu przypisania elementów przestrzeni X do właściwości (cechy) odpowiadającej zbiorowi A .

Jeżeli jakość narzędzia ściernego można zawrzeć w dziesięciostopniowej przestrzeni X i oznaczyć

$$X = \{x_1=1, x_2=2, x_3=3, x_4=4, x_5=5, x_6=6, x_7=7, x_8=8, x_9=9, x_{10}=10\},$$

gdy A jest zbiorem narzędzi średniej jakości, to funkcja $\mu_A(x)$ może przyjmować np. takie wartości:

$$\mu_A(1)=0, \mu_A(2)=0,1, \mu_A(3)=0,4, \mu_A(4)=0,8, \mu_A(5)=1,$$

$$\mu_A(6)=1, \mu_A(7)=0,8, \mu_A(8)=0,4, \mu_A(9)=0,1, \mu_A(10)=0,$$

co w specjalnej notacji wprowadzonej przez Zadeha [4] wyraża się następująco:

$$A = \mu_A(x_1)/x_1 + \mu_A(x_2)/x_2 + \mu_A(x_3)/x_3 + \mu_A(x_4)/x_4 + \mu_A(x_5)/x_5 +$$

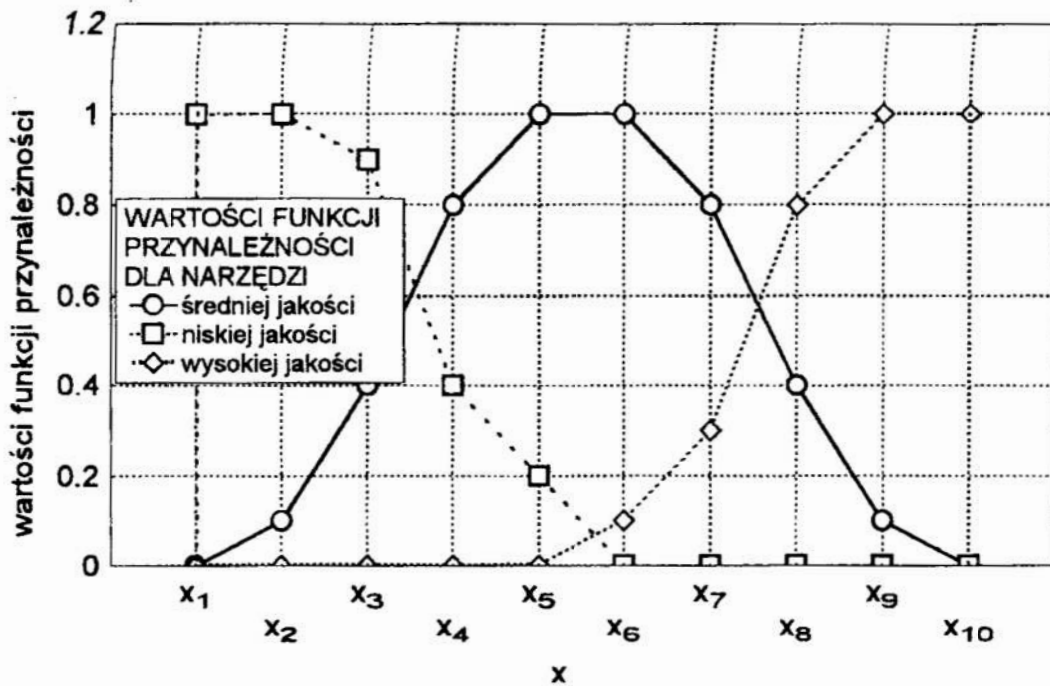
$$\mu_A(x_6)/x_6 + \mu_A(x_7)/x_7 + \mu_A(x_8)/x_8 + \mu_A(x_9)/x_9 + \mu_A(x_{10})/x_{10},$$

czyli $A = 0,1/2 + 0,4/3 + 0,8/4 + 1/5 + 1/6 + 0,8/7 + 0,4/8 + 0,1/9$ (rys.1).

Dla narzędzi niskiej jakości (zbiór B) oraz narzędzi wysokiej jakości (zbiór C) może być np.:

$$B = 1/1 + 1/2 + 0,9/3 + 0,4/4 + 0,2/5, \quad C = 0,1/6 + 0,3/7 + 0,8/8 + 1/9 + 1/10 \text{ (rys.1)}.$$

Na zbiorach rozmytych mogą być przeprowadzane operacje takie jak: sumowanie, przecięcie i uzupełnienie, co opisano poniżej.



Rys. 1. Wykresy przykładowych funkcji przynależności cechy rozmytej
 Fig. 1. Diagrams of exemplary membership function of fuzzy feature

Suma C dwóch zbiorów rozmytych A i B o funkcjach przynależności odpowiednio $\mu_A(x)$ i $\mu_B(x)$ jest zbiorem rozmytym o funkcji przynależności

$$\mu_C(x) = \max(\mu_A(x), \mu_B(x)) \text{ (rys. 2).}$$

Przecięcie C dwóch zbiorów rozmytych A i B jest zbiorem rozmytym o funkcji przynależności: $\mu_C(x) = \min(\mu_A(x), \mu_B(x))$.

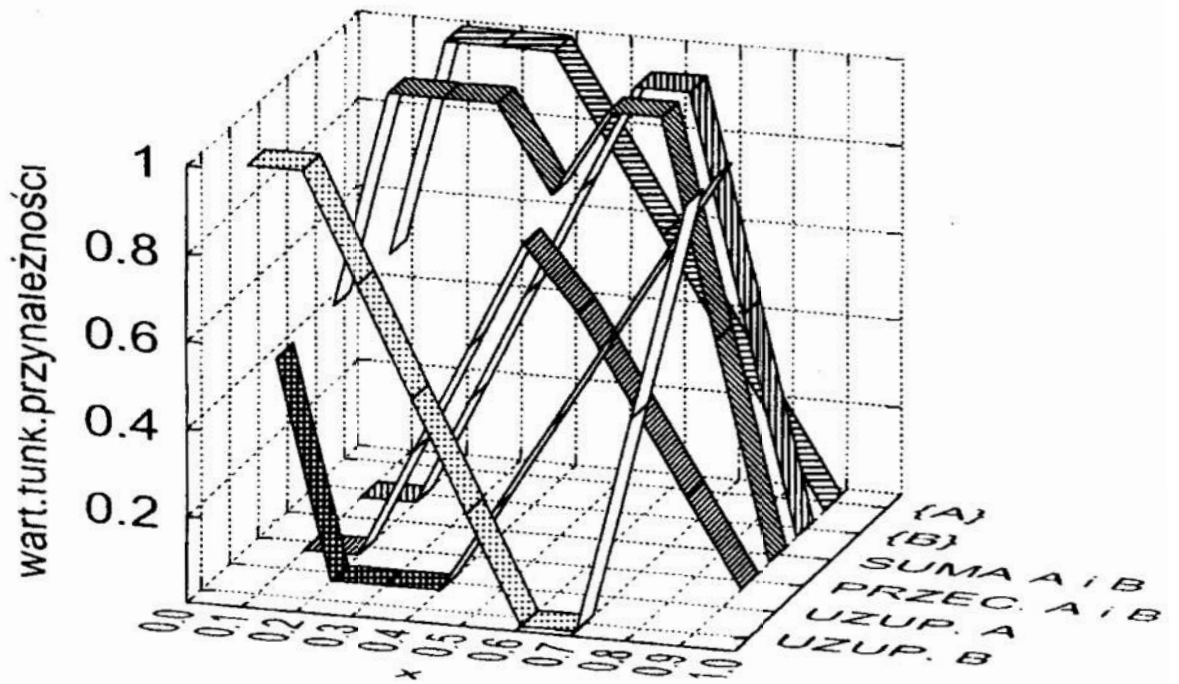
Uzupełnienie A' zbioru rozmytego A jest zbiorem rozmytym o funkcji przynależności: $\mu_{A'}(x) = 1 - \mu_A(x)$.

Iloczyn kartezjański zbioru A z przestrzeni X i zbioru B z przestrzeni Y określa się następująco: $A \times B = \min(\mu_A(x), \mu_B(y))$ (rys. 3).

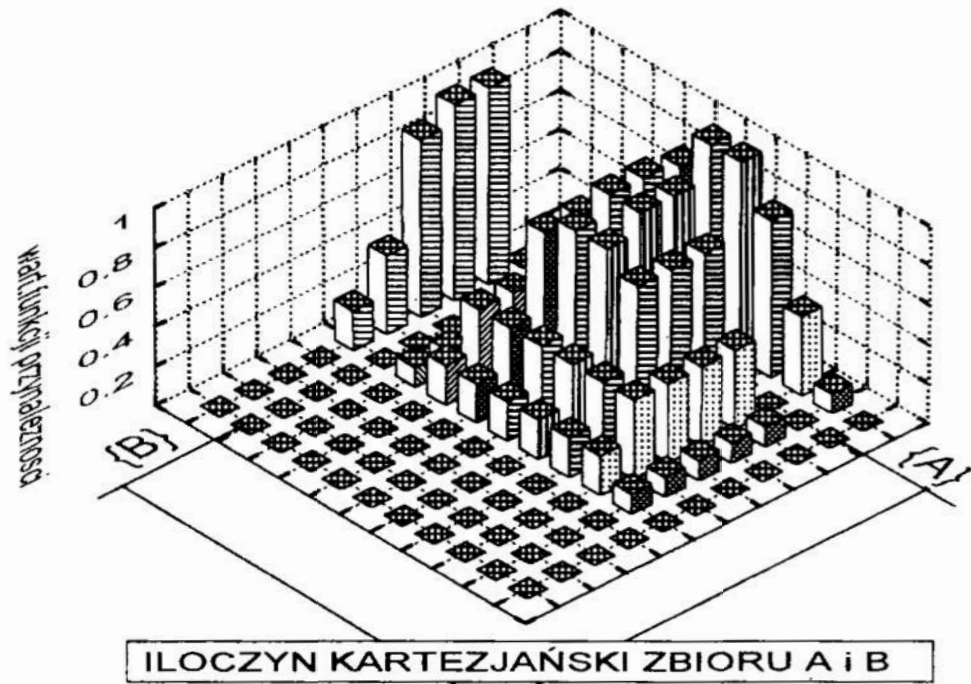
We wnioskowaniu przybliżonym wykorzystuje się liczby rozmyte L o określonych funkcjach przynależności np. liczby $\alpha - \beta$. Liczby takie są reprezentowane przez czwórkę parametrów (a, b, α , β). Pierwsze dwa parametry określają przedział, w którym funkcja przynależności μ_L posiada wartość 1, natomiast następne dwa określają szerokość lewego i prawego przedziału, w którym μ_L zmienia wartość odpowiednio od 0 do 1 i od 1 do 0 (rys. 4).

Funkcja przynależności jest zatem następująca:

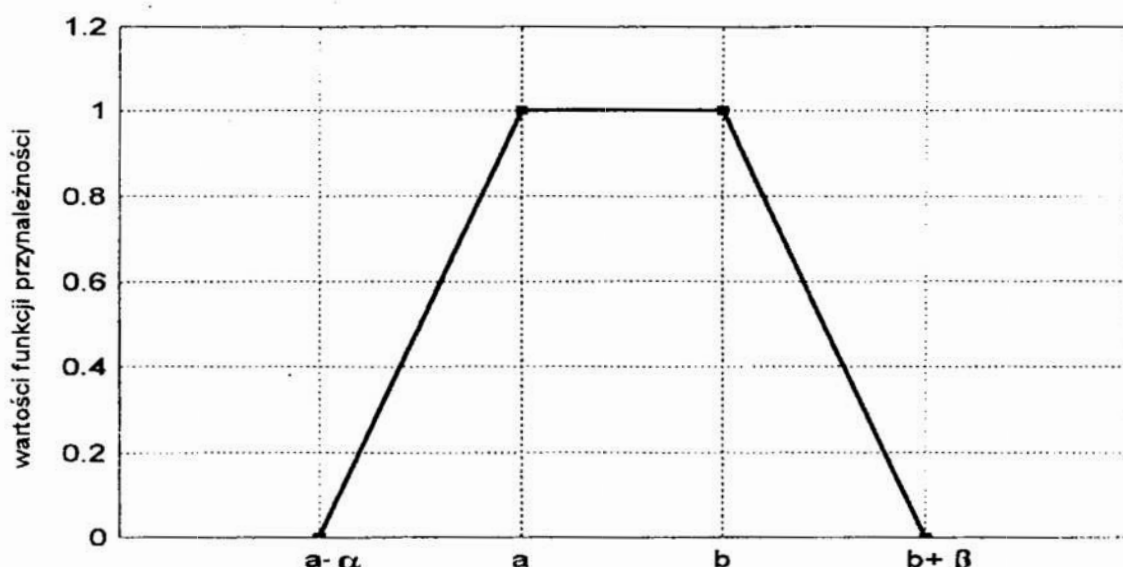
$$\mu_L(x) = \begin{cases} 0, & x < a - \alpha, \\ (1/\alpha)(x - a + \alpha), & x \in [a - \alpha, a), \\ 1, & x \in [a, b], \\ (1/\beta)(b + \beta - x), & x \in (b, b + \beta], \\ 0, & x > b + \beta. \end{cases}$$



Rys.2. Ilustracja zastosowania operatorów rozmytych
 Fig.2. Presentation of fuzzy operators application



Rys.3. Wartości funkcji przynależności dla iloczynu kartezjańskiego
 zbiorów A i B
 Fig.3. Values of membership function for the product of Cartesian sets
 A and B



Rys.4. Wykres funkcji przynależności liczby rozmytej α - β
 Fig.4. Diagram of membership function for a fuzzy number α - β

3. Przykład zastosowania teorii zbiorów rozmytych do oceny jakości ściernicy

W procesie wnioskowania rozmytego występują cztery typy operacji: szacowanie stopni niepewności przesłanek połączonych operatorami koniunkcji, ocena stopni niepewności konkluzji na podstawie analizy przesłanek i reguł wnioskowania, ocena niepewności tych konkluzji, do których prowadzi więcej niż jedna ścieżka wnioskowania, minimalizacja przedziału niepewności tych samych zdarzeń (cech) określanych z różnych źródeł danych (teoria, kolejne eksperymenty).

Rodzina funkcji stosowanych jako operatory przekształceń, przydatnych do tego celu są normy trójkątne (T-normy) (przecięcie) i co-normy trójkątne (T-co-normy, albo S-normy) [1, 6, 7, 8] (suma), przy czym dla określonego operatora (np. operatora negacji) wybór określonego przekształcenia dla koniunkcji (T) ustala jednoznacznie przekształcenie dla dysjunkcji (S) i odwrotnie:

$$s(x, y) = 1 - t(1-x, 1-y), \quad t(x, y) = 1 - s(1-x, 1-y).$$

Oprócz standardowych (nieparametrycznych) operatorów agregacji zbiorów rozmytych stosowane są również parametryczne T-normy i T-co-normy wprowadzone przez Yagera, Schweizera, Hamachera, Franka, Dombi, Webera i Duboisa [1, 6].

W dalszej części pracy - w rozpatrywanym przykładzie - jako nieparametryczne operatory przecięcia ($A \cap B$) (koniunkcji) przyjęto: średnią harmoniczną $2xy/(x+y)$, średnią geometryczną $(xy)^{1/2}$, średnią arytmetyczną $(x+y)/2$, dualną średnią geometryczną $1-((1-x)(1-y))^{1/2}$ oraz dualną średnią harmoniczną $(x+y-2xy)/(2-x-y)$.

Operator relacji rozmytej "i"(AND) został zdefiniowany następująco:

$$\mu_{\text{AND}}(x, y) = \gamma \min(x, y) + (1-\gamma)(x+y)/2 \quad \gamma \in [0, 1],$$

natomiast operator relacji rozmytej "albo" (OR) jak poniżej:

$$\mu_{\text{OR}}(x, y) = \gamma \max(x, y) + (1-\gamma)(x+y)/2 \quad \gamma \in [0, 1].$$

Parametr γ określa wagę jaką przypisuje się do relacji "i" oraz "albo". Dla $\gamma=1$ operator relacji rozmytej AND staje się operatorem MIN, a operator OR staje się operatorem MAX. Dla $\gamma=0$ obydwaj operatory stają się średnią arytmetyczną. Przydatnym operatorem okazał się operator GAMMA, który dla n argumentów ma postać:

$$\mu_{\text{COMP}}(x_1, \dots, x_n) = (\prod x_i)^{(1-\gamma)} (1 - \prod(1-x_i))^\gamma \quad \gamma \in [0, 1].$$

Operator ten dla $\gamma=1$ staje się T-co-normą a dla $\gamma=0$ jest T-normą. Kolejnymi przydatnymi operatorami, będącymi kombinacją odpowiadających sobie par T-norm i T-co-norm, z uwzględnieniem wagi γ są: operator MIN-MAX i operator ALGEBR-PROD-SUM definiowane następująco:

$$\mu_{\text{MIN-MAX}}(x_1, \dots, x_n) = \gamma \min(x_1, \dots, x_n) + (1-\gamma) \max(x_1, \dots, x_n) \quad \gamma \in [0, 1],$$

$$\mu_{\text{ALGEBR-PROD-SUM}}(x_1, \dots, x_n) = \gamma (\prod x_i) + (1-\gamma) (1 - \prod(1-x_i)) \quad \gamma \in [0, 1].$$

Dla przykładowego zobrazowania zastosowań rozpatrzono uproszczony problem doboru właściwości ściemicy do realizacji określonej operacji szlifowania. Jako kryteria wyboru przyjęto: maksymalną graniczną wydajność szlifowania w danych warunkach (MWS), trwałość ściemicy dla ustalonej wydajności szlifowania (TS), koszt ściemicy (KS) i względną jakość warstwy wierzchniej (JWW).

Dla cech MWS (rys.5), TS i JWW, których wartości im są większe tym korzystniej, przyjęto funkcję przynależności typu S w postaci:

$$\mu_s(x, a, \delta) = \begin{cases} 0, & x < a-\delta, \\ 2((x-a+\delta)/2\delta)^2, & x \in [a-\delta, a), \\ (1-2((a-x+\delta)/2\delta)^2), & x \in (a, a+\delta], \\ 1, & x > a+\delta. \end{cases}$$

a dla cechy KS funkcję przynależności typu Z: $\mu_z(x, a, \delta) = 1 - \mu_s(x, a, \delta)$, co oznacza, że im wyższy jest koszt narzędzia tym, mniejsza jest wartość funkcji przynależności.

Wartości funkcji przynależności odpowiadające wszystkim rozpatrywanym cechom dla porównywanych ściemicy oraz wartości operatorów (z zaznaczeniem ich maksimum) zestawiono w tablicy I.

W rozpatrywanym przypadku maksymalne wartości operatorów (poza jednym) wskazują na ściemicy nr 4. Analizy powyższe są jedynie wprowadzeniem do tworzenia algorytmów automatycznego wnioskowania z wykorzystaniem teorii zbiorów rozmytych. Mają one nad innymi metodami wnioskowania

przewagę wynikającą ze zgodności sposobu wnioskowania z rodzajem informacji, jakie stają się podstawą wniosków o badanych procesach.

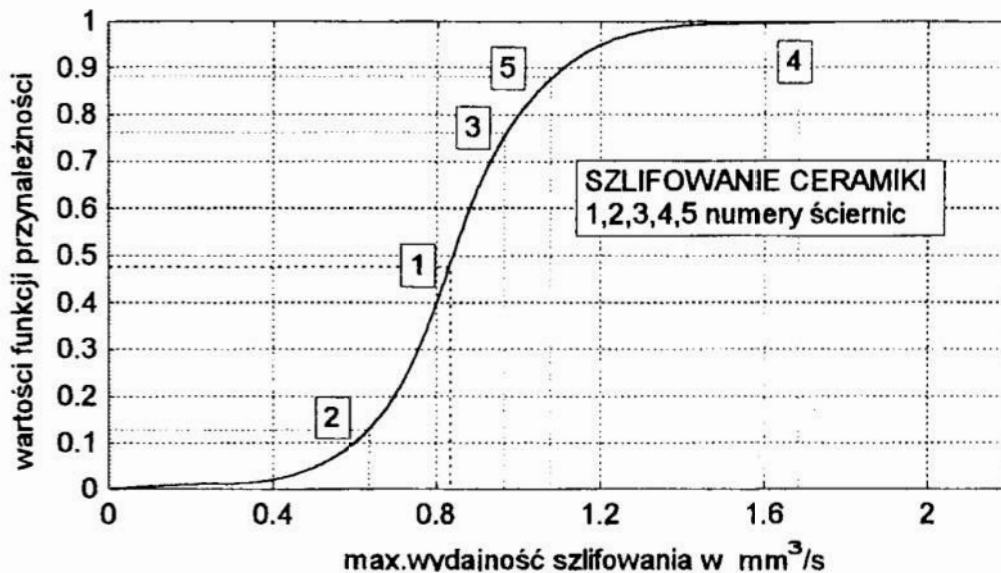
Tablica I. Wartości funkcji przynależności i wartości operatorów służących do oceny przydatności ściernicy
 Table I. Values of membership function and operators values serving for the grinding wheel suitability evaluation

CECHA	ŚCIERNICA				
	1	2	3	4	5
MAX. WYDAJNOŚĆ SZLIF.	0.47	0.13	0.76	1	0.88
TRWAŁOŚĆ ŚCIERNICY	0.75	0.92	0.53	0.71	0.66
KOSZT ŚCIERNICY	0.62	0.74	0.59	0.38	0.44
JAKOŚĆ WARSTWY WIERZCH.	0.58	0.25	0.52	0.8	0.76
OPERATORY (I ICH WARTOŚCI MAKSYMALNE)					
MINIMUM WART. (max(MIN))	0.47	0.13	0.52	0.38	0.44
ŚREDNIA ARYTMET. (max(WS))	0.61	0.51	0.6	0.72	0.69
ŚREDNIA GEOMETRYCZNA	0.6	0.39	0.59	0.68	0.66
DUAL ŚREDNIA HARMONICZNA	0.62	0.66	0.61	1	0.73
OPERATOR GAMMA=0.6	0.43	0.22	0.43	0.54	0.52
OPERATOR GAMMA=0.5	0.35	0.15	0.35	0.46	0.44
OPERATOR GAMMA=0.4	0.29	0.1	0.28	0.4	0.37
ALGEBR-PROD-SUM GAMMA=0.7	0.38	0.31	0.38	0.45	0.43
ALGEBR-PROD-SUM GAMMA=0.4	0.64	0.6	0.64	0.69	0.67
OPERATOR MIN-MAX GAMMA=0.6	0.58	0.45	0.62	0.63	0.62

Inną zaletą zastosowań wnioskowania z wykorzystaniem teorii zbiorów rozmytych jest to, że jest ono zbliżone do naturalnego wnioskowania człowieka [5]. Może zatem stanowić podstawy automatycznych systemów wnioskowania, które będą mogły wyręczać operatora lub wspomagać podejmowanie przez niego optymalnych decyzji.

4. Podsumowanie

Automatyczne wnioskowanie jest warunkiem dużej elastyczności systemów wytwarzania. Cechy tej nie posiadają układy, których działanie opisano w sposób zdeterminowany i, gdy działanie to nie jest przez określony czas zmieniane. Procesy wytwarzania są ze swej natury procesami, z których wydobywamy informacje niepełne i nieściśle. W pracy podkreślono, iż w takiej sytuacji wykorzystanie teorii zbiorów rozmytych pozwala lepiej opisywać procesy podejmowania decyzji. Zakres aplikacji tej teorii szybko się poszerza. Zamieszczone przykłady wskazują na możliwość nowych zastosowań.



Rys.5. Wartości funkcji przynależności dla cechy MWS dla pięciu ściernic przeznaczonych do szlifowania ceramiki

Fig.5. Values of membership function of the MWS feature for five grinding wheels destined for grinding of ceramics

8. Literatura

1. *Bolc L., Brodziewicz W., Wójcik M.*: Podstawy przetwarzania informacji niepewnej i niepełnej. Wydawnictwo Naukowe PWN, Warszawa 1991.
2. *Kacalak W.*: Mechanizm kumulacji skutków zakłóceń losowych oraz zmian warunków obróbki w aspekcie optymalizacji procesów skrawania. Zeszyty Naukowe Politechniki Poznańskiej. Politechnika Poznańska nr 30/84 Poznań, 1984, s. 145-149.
3. *Kacalak W.*: Wyznaczenie trwałości narzędzi ściernych z uwzględnieniem losowego charakteru procesu szlifowania z nałożonymi warunkami ograniczającymi. Zeszyty Naukowe Politechniki Poznańskiej. Politechnika Poznańska nr 30/84 Poznań, 1984, s. 151-156.
4. *Zadeh L.A.*: Fuzzy Sets. Information and Control 8/1965, s.338-353.
5. *Zadeh L.A.*: Making Computers Think Like People. IEEE Spectrum. Vol. 21, 8/1984, s.26-32.
6. *Mayer A., Mechler B., Schlinwein A., Wolke R.*: Fuzzy Logic. Addison-Wesley, 1993.
7. *Tilli T.*: Fuzzy-Logic: Grundlagen, Anwendungen, Hard- und Software. Monachium, Franzis, 1991.
8. *Trautzi G.*: Werkzeuge für Fuzzy-Logic-Entwicklungen. KI 4/1991, s.69-72.

INTRODUCTION TO FUZZY INFERENCE IN APPLICATION TO FEATURE ANALYSIS OF THE GRINDING PROCESSES

Summary

Automatic inference is a condition for big elasticity of manufacturing systems. There are systems, with the operation described in the determined manner, which do not possess this feature. Manufacturing processes are in their nature the processes providing incomplete and inaccurate information. Making use of fuzzy sets enables one for better description of the processes.

METODYKA PRZETWARZANIA INFORMACJI NIEPEWNEJ I TWORZENIA ALGORYTMÓW DO STEROWANIA PROCESAMI SZLIFOWANIA Z WYKORZYSTANIEM TEORII ZBIORÓW ROZMYTYCH

Ważnym celem wielu zastosowań nowych narzędzi matematyki i informatyki jest automatyzacja wnioskowania i sterowania w warunkach dysponowania informacją niepewną i nieściłą. Informacja, jaką wydobywamy z realizacji procesów wytwarzania ma właśnie takie cechy. Połączenie metod analizy informacji niepewnej z metodami typowymi dla teorii zbiorów rozmytych, reprezentujących informację nieprecyzyjną, pełniej opisuje wiele problemów, takich jak np. propagacja i kumulowanie niepewności w procesach szlifowania.

1. Przetwarzanie informacji niepewnej przez operatora

Dobłą drogą do ulepszania systemów automatycznego wnioskowania jest analiza postępowania człowieka w sytuacjach, gdy odbiera on informacje z otoczenia, które zwykle są niepewne i nieściłe. Trzeba jednak przy tym zauważyć, iż nie wszystkie cechy operatora warto naśladować [1]. Pierwszą zauważalną cechą operatora jest ograniczenie jego percepcji do kilku (najczęściej 5...9) stopni różnicowania odbieranych sygnałów. Innym ograniczeniem jest ilość danych, jakie mogą być jednorazowo zapamiętane. W postępowaniu operatora korzystną cechą jest umiejętność adaptacji do odbioru informacji niepewnej, niekorzystną natomiast skłonność do "konserwatyizmu" i ulegania presji oczekiwania na informację spodziewaną.

Ciekawą właściwością operatora jest obniżanie poziomu konserwatyizmu (opóźnienia w dokonywaniu przeszacowań prawdopodobieństwa zdarzeń wraz z napływaniem nowych informacji), gdy ilość danych jest mała, gdy nagradzana jest dokładność a nie wydajność, a także w przypadkach zaskakujących wartości pierwszych danych. Dla tematu niniejszej pracy istotne jest określenie czy i w jakim stopniu subiektywne prawdopodobieństwa podlegają takim samym prawom jak prawdopodobieństwa obiektywne. Wyniki badań [2] wskazują, iż u operatora występuje skłonność do zawyżania prawdopodobieństwa iloczynu zdarzeń niezależnych A i B (sytuacji, gdy zajdzie zarówno zdarzenie A jak i zdarzenie B) oraz skłonność do zaniżania prawdopodobieństwa sumy zdarzeń A i B (sytuacji, gdy zajdzie przynajmniej jedno z tych zdarzeń). W przypadku możliwego wystąpienia jednego ze zdarzeń niezależnych A i B, gdy zdarzenie A występowało z częstością większą od oczekiwanej, operator ulega skłonności do zawyżania prawdopodobieństwa oczekiwanego zdarzenia B.

* Wyższa Szkoła Inżynierska w Koszalinie, ul. Raławicka 15/17, 75-620 Koszalin

Typową, korzystną cechą operatora jest świadome i podświadome korzystanie z umiejętności odkrywania nowych zależności. Występują przy tym jednak błędy, wynikające z pomijania wpływu wielkości próbki, skłonności do przypisywania większych prawdopodobieństw informacjom łatwiej zapamiętywanym i skłonności do przedwczesnego odrzucania możliwych współzależności. Inną cechą operatora jest często zbyt mała wnikliwość w ocenie, czy możliwa jest współzależność cech, których wartości są w danym procesie skorelowane.

2. Rozmyte zbiory cech procesów szlifowania

W procesach szlifowania liczne zjawiska o małym często znaczeniu w obróbce skrawaniem nabierają znaczenia decydującego o przebiegu i efektach procesu. Wynika to z tego, iż stosunek parametrów charakteryzujących rozproszenie wartości danej cechy do parametrów będących miarami położenia osiąga duże wartości - zbliżone lub nawet większe od 1. Przyczyną tego są takie zjawiska jak [3, 4, 5, 6]: drgania elementów układu technologicznego, nieciągłość procesu tworzenia mikrowiórów, ciepłne i mechaniczne sublokalne odkształcenia narzędzi i materiału obrabianego, losowe zmiany topografii aktywnej powierzchni narzędzia oraz zmiana warunków mikroskrawania wzdłuż drogi pracy każdego z ostrzy stykających się z obrabianym materiałem.

We wszystkich sposobach obróbki zagłębienie naroża ostrza w materiał jest zmienne wzdłuż toru skrawania. Jest to skutkiem wielu czynników, takich jak [3]: zmienność zagłębienia nominalnego (zależna od kinematycznych cech metody), nierówność powierzchni przedmiotu w strefie obróbki, podatność lokalna materiału obrabianego i ziaren ściemych, lokalne mikroprzemieszczenia narzędzia i ziaren ściemych oraz mikronieciągłość procesu tworzenia wiórów. To wszystko, wraz z niedoskonałością metod wydobywania i przetwarzania informacji, powoduje iż zbiory cech charakteryzujących procesy szlifowania są zbiorami rozmytymi.

Zbiór rozmyty jest charakteryzowany przez funkcję $\mu_A(x)$, nazywaną funkcją przynależności, przekształcającą elementy przestrzeni X (wartości cech) w liczby z zakresu $0...1$ (znormalizowany zbiór rozmyty), odpowiadające prawdopodobieństwu przypisania elementów przestrzeni X do właściwości (cechy) odpowiadającej zbiorowi A .

Ważnym elementem teorii zbiorów rozmytych, wykorzystywanym w niniejszej pracy, jest zmienna lingwistyczna [1, 9,10,11,12], definiowana przez podanie pięciu parametrów: $\{N, T(N), X, GN, ZW\}$, gdzie N jest nazwą zmiennej. $T(N)$ jest zbiorem rozmytych wartości zmiennej N a X jest przestrzenią obiektów, w której określane są zbiory $T(N)$. GN jest zbiorem reguł syntaktycznych do generowania nazw wartości zmiennej N , a ZW jest zbiorem reguł semantycznych nadających znaczenie wartościom zmiennej N .

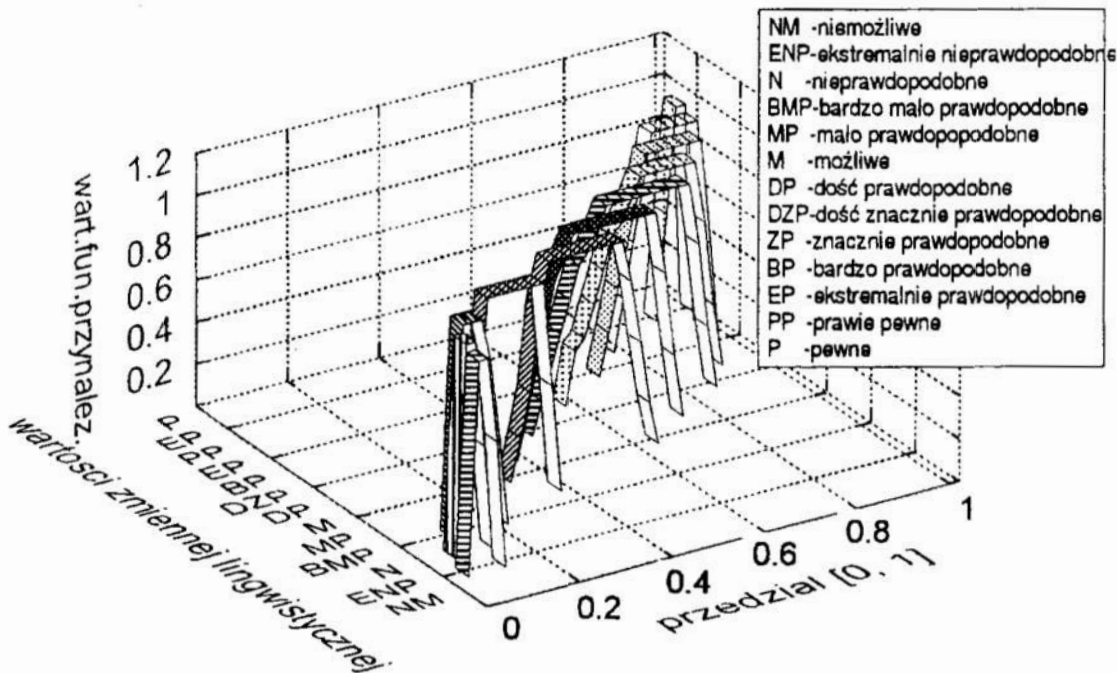
W przykładowym zbiorze stopni niepewności o dużej rozdzielczości można wyróżnić 13 elementów, będących rozmytymi zmiennymi lingwistycznymi, zapisywanymi w postaci: { wiarygodność, T(wiarygodność), [0, 1], GN, ZW }.

Wartości zmiennych T są następujące:

T(wiarygodność) =

{ niemożliwe, ekstremalnie-nieprawdopodobne, nieprawdopodobne, bardzo-mało-prawdopodobne, mało-prawdopodobne, możliwe, dość-prawdopodobne, dość-znacznie-prawdopodobne, znacznie-prawdopodobne, bardzo-prawdopodobne, ekstremalnie-prawdopodobne, prawie-pewne, pewne }.

Semantyka (ZW) każdego elementu zmiennej wiarygodność jest liczbą rozmytą określoną na przedziale [0, 1], której parametry (a, b, α , β) mogą być np. takie jak na rysunku 1, przy czym przedziały a- α odpowiadają lewym zboczom funkcji przynależności, przedziały a-b wartościom funkcji przynależności równym jedności, a przedziały b+ β odpowiadają prawym zboczom funkcji przynależności.



Rys. 1. Rozkład funkcji przynależności dla zmiennych zbioru "wiarygodność"
 Fig. 1. Distribution of membership function for variables of the "credibility"

3. System wykorzystywania informacji rozmytych

W systemie automatycznego wnioskowania w warunkach przetwarzania informacji niepewnej i nieściślej można wyróżnić trzy poziomy: reprezentacji, wnioskowania i sterowania (decyzji) (rys.2). Na poziomie reprezentacji w odniesieniu do procesów szlifowania należy wymagać:

- ♦ opisu niepewności wyników procesu w całym zakresie możliwych parametrów i przyjętym zakresie warunków obróbki,
- ♦ analizy wrażliwości wyników procesu na zmiany parametrów szlifowania,
- ♦ oceny wpływu czasu obróbki (głównie zmian stanu narzędzi w okresie trwałości) na cechy procesu
- ♦ wyczerpującego opisu stanów wątpliwych, sprzeczności eksperymentu i teorii,
- ♦ analizy mechanizmów kumulacji skutków zakłóceń dla ważniejszych zjawisk.

Na poziomie wnioskowania wskazane jest:

- ♦ unikanie przedwczesnych założeń o niezależności zdarzeń,
- ♦ dążenie do uzupełniania zbioru hipotez, który zazwyczaj nie jest wyczerpujący,
- ♦ staranność w ocenie przyjmowanych modeli mechanizmów propagacji i kumulacji niepewności z dążeniem do przejrzystej semantyki.

W trzecim poziomie - sterowania - ważnym jest:

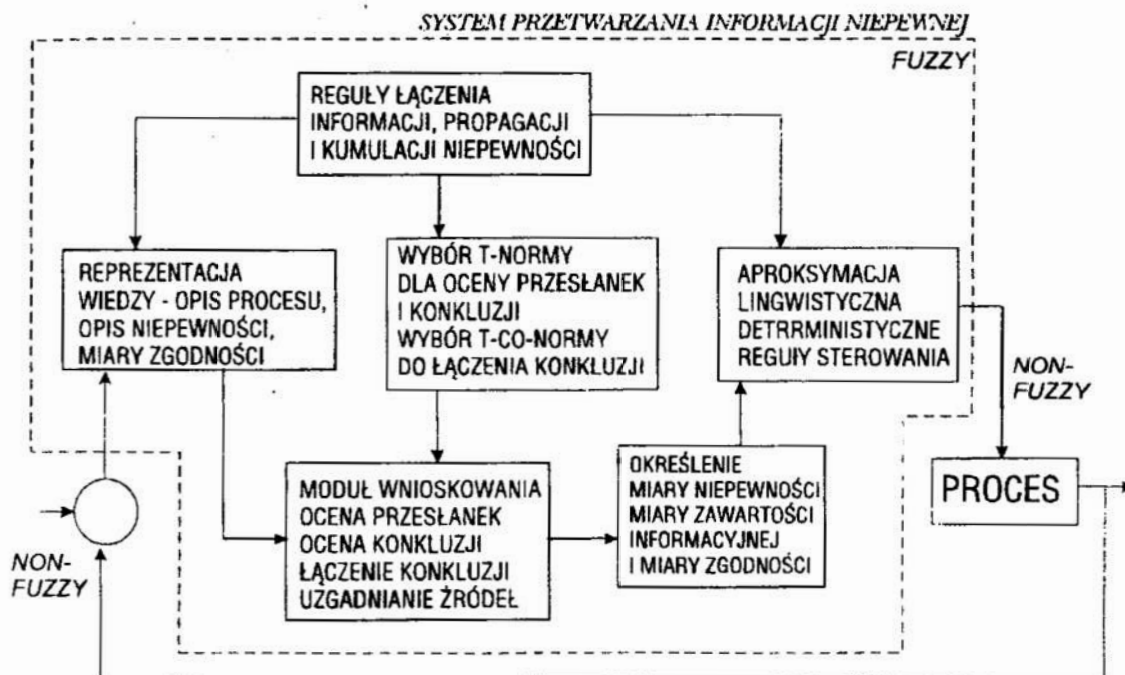
- ♦ wyraźne odróżnienie sytuacji sprzecznych, wynikających z charakteru procesu od sytuacji wynikających z braku kompletnego opisu procesu,
- ♦ przeprowadzenie analizy drogi propagacji niepewności i ich znaczenia dla konkluzji.

Na poziomie wnioskowania występują cztery typy operacji: szacowanie stopni niepewności przesłanek połączonych operatorami koniunkcji, ocena stopni niepewności konkluzji na podstawie analizy przesłanek i reguł wnioskowania, ocena niepewności tych konkluzji, do których prowadzi więcej niż jedna ścieżka wnioskowania, minimalizacja przedziału niepewności tych samych zdarzeń (cech) określanych z różnych źródeł danych (teoria, kolejne eksperymenty).

Na tym etapie wykorzystuje się operatory koniunkcji i dysjunkcji określające stopień pewności przesłanek składających się z wielu części. Rodziną funkcji stosowanych do przekształceń, przydatnych do tego celu są normy trójkątne (T-normy) (jako operatory przecięcia zbiorów rozmytych) i co-normy trójkątne (T-co-normy, albo S-normy) [1, 9, 10, 11, 12] (jako operatory sumy zbiorów rozmytych), przy czym wybór określonego przekształcenia dla koniunkcji (T) ustala jednoznacznie przekształcenie dla dysjunkcji (S) i odwrotnie:

$$s(x, y) = 1 - t(1-x, 1-y), \quad t(x, y) = 1 - s(1-x, 1-y).$$

Oprócz standardowych (nieparametrycznych) operatorów agregacji zbiorów rozmytych stosowane są również parametryczne T-normy i T-co-normy wprowadzone przez Yagera, Schweizera, Hamachera, Franka, Dombi, Webera i Duboisa [1, 9].



Rys.2. Struktura układu sterowania w systemie przetwarzania informacji niepewnej

Fig.2. Structure of the control system in the system of processing of uncertain informatuin

5. Przykład algorytmu

Budowa algorytmu automatycznego wnioskowania zostanie pokazana na przykładach podejmowania decyzji o odnowieniu czynnej powierzchni ściernicy (przykład P1) oraz problemu doboru prędkości posuwu w procesie szlifowania elementów ceramicznych (przykład P2). Do budowy algorytmu automatycznego wnioskowania z wykorzystaniem teorii zbiorów rozmytych można wykorzystać poniższe zestawienie kolejnych kroków:

określenie alternatyw (w przykładzie P1 występują dwie alternatywy - akceptacja właściwości określonej ściernicy lub brak akceptacji, w przykładzie P2 trzy alternatywy - zwiększyć prędkość posuwu, pozostawić bez zmian, zmniejszyć prędkość posuwu),

określenie zmiennych lingwistycznych, a w tym:

określenie nazw zmiennej (atrybutów)

(w przykładzie P1: względny przyrost mocy szlifowania (PMS), względna jakość warstwy wierzchniej (JWW) oraz dokładność szlifowania (DS), decyzja o odnowieniu - DOO;

w przykładzie P2: przyrost mocy szlifowania (PMS), częstość wykruszeń krawędzi lub pęknięć szlifowanego przedmiotu (CW), oraz dokładność szlifowania (DS), decyzja o zmianie posuwu - DZP,

- **określenie zbioru rozmytych wartości zmiennych lingwistycznych, zwanych termami,**
(w przykładzie P1 i P2:
PMS (UJEMNY, BLISKI ZERA, UMIARKOWANY, DUŻY, BARDZO DUŻY),
JWW(NIEDOSTATECZNA, DOSTATECZNA, DOŚĆ-DOBRA, DOBRA, BARDZO-DOBRA),
DS (NISKA, DOŚĆ-DOBRA, WYSOKA),
CW (NISKA, UMIARKOWANA, DOŚĆ-DUŻA, DUŻA, BARDZO-DUŻA))
DOO(NIE, TAK),
DZP (ZMNIJSZENIE, BEZ-ZMIANY, ZWIĘKSZENIE),
- ♦ **określenie przestrzeni, w których tworzone są rozmyte zbiory, będące wartościami zmiennych lingwistycznych**
(w rozpatrywanych przykładach:
PMS $\in\{-0,5, 0,5\}$, JWW $\in\{0, 1\}$,
KS $\in\{0, 100\}$ (μm), CW $\in\{0, 0,1\}$
DOO $\{0, 1\}$,
DZP $\{-1, 1\}$,
- ♦ wyznaczenie zbioru wielkości wejściowych i wyznaczenie odpowiadających im wartości zmiennych lingwistycznych,
- ♦ zbudowanie grafu relacji (sieci wnioskowania),
- ♦ dobór klasy operatorów i typu operatora dla danej klasy (np. klasy operatorów agregacji: T-normy, ŚREDNIE, T-co-normy i dla nich odpowiednich operatorów parametrycznych lub nieparametrycznych),
- ♦ określenie relacji rozmytych (dla rozpatrywanego przykładu, tak jak w poniższym fragmencie),

if	(PMS	is	UMIARKOWANY)
and	(JWW	is	BARDZO-DOBRA)
and	(DS	is	WYSOKA)
then	(DOO	is	NIE)
with	WSPÓŁCZYNNIK-PEWNOŚCI = 0.90		
.	.	.	.
.	.	.	.
.	.	.	.

- ♦ określenie wyników; pierwotnym wynikiem działania systemu jest funkcja przynależności dla zmiennej lingwistycznej będącej określoną decyzją (zmienną wyjściową), wtórnym wynikiem, który pomija część informacji ale za to, przestaje być rozmytym, jest skalarna wielkość sterująca,
- ♦ testowanie i korekta systemu polegająca na modyfikacji funkcji przynależności, przyjętych operatorów, współczynników pewności.

Przejście od zmiennych rozmytych do nie rozmytej wielkości wyjściowej może nastąpić z wykorzystaniem jednej z kilku metod, z których najbardziej znanymi są: metoda środka ciężkości oraz metoda maksimum. Pierwsza z nich polega na znalezieniu wartości odciętych (x), będącej współrzędną środka ciężkości powierzchni ograniczonej linią przedstawiającą wykres funkcji przynależności zmiennej decyzyjnej.

Metoda ta może być stosowana, gdy wynikowy zbiór rozmyty (funkcja przynależności) nie posiada lokalnych maksimów. W przypadku dwóch alternatyw zmiennej wyjściowej metoda środka ciężkości nie daje dostatecznie wyraźnej decyzji. W takim przypadku lepszą jest metoda maksimum, której wynikiem jest wartość dla której funkcja przynależności osiąga maksimum globalne.

Dla zbudowania grafu wnioskowania i doboru operatorów bardzo przydatnym jest zbudowanie tablicy zależności decyzji od atrybutów, tak jak w poniższych przykładach: dla problemu podejmowania decyzji o odnowieniu powierzchni czynnej ściernicy (tabl. II) i problemu doboru prędkości posuwu w procesie szlifowania elementów ceramicznych (tabl. III).

Tablica II. Fragment zestawienia decyzji o odnowieniu powierzchni ściernicy w zależności od wartości zmiennych lingwistycznych charakteryzujących proces szlifowania

Table II. Juxtaposition fragment of decision of renewal of the wheel surface dependent on the values of linguistic variables characterizing the grinding process

NR	PRZYRÓST MOCY SZLIFOWANIA	JAKOŚĆ WARSTWY WIERZCHNIEJ	DOKŁADNOŚĆ SZLIFOWANIA	DECYZJA O ODNOWIENIU POW. CZYNNEJ ŚCIERNICY	WSPÓŁ. PEWNOŚCI
1	UJEMNY	B. DOBRA	WYSOKA	NIE	1
2	BLISKI ZERA	DOBRA	WYSOKA	NIE	1
3	UMIARKOWANY	DOŚĆ DOBRA	WYSOKA	NIE	0.9
4	DUŻY	DOSTATECZNA	NISKA	TAK	0.8
5	BARDZO DUŻY	NIEDOSTAT.	NISKA	TAK	1
6	UJEMNY	DOBRA	DOŚĆ DOBRA	NIE	0.9
7	BLISKI ZERA	DOBRA	DOŚĆ DOBRA	NIE	0.85
8	UMIARKOWANY	DOBRA	DOŚĆ DOBRA	NIE	0.8
9	DUŻY	DOSTATECZNA	NISKA	TAK	0.7
10	BARDZO DUŻY	DOSTATECZNA	NISKA	TAK	1
11	UJEMNY	DOŚĆ DOBRA	WYSOKA	NIE	0.95
12	BLISKI ZERA	DOŚĆ DOBRA	WYSOKA	NIE	0.9
13	UMIARKOWANY	DOŚĆ DOBRA	WYSOKA	NIE	0.85
14	DUŻY	DOŚĆ DOBRA	NISKA	TAK	0.6
15	BARDZO DUŻY	DOSTATECZNA	DOŚĆ DOBRA	TAK	0.7

Tablica III. Fragment zestawienia decyzji o zmianie prędkości posuwu (ruchu obrotowego stołu) w zależności od wartości zmiennych lingwistycznych charakteryzujących proces szlifowania elementów ceramicznych

Table III. Juxtaposition fragment of decision of feed change (for the table rotary motion) dependent on the values of linguistic variables characterizing the grinding process of ceramic elements

NR	PRZYROST MOCY SZLIFOWANIA	CZĘSTOŚĆ WYSTĘPOWANIA WYKRUSZEŃ KRAWĘDZI LUB PEKNIĘĆ PRZEDMIOTU	DOKŁADNOŚĆ SZLIFOWANIA	DECYZJA ZMIANIE PRĘDKOŚCI POSUWU	WSPÓŁCZYNNIK PEWNOŚCI
1	UJEMNY	NISKA	WYSOKA	ZWIĘKSZENIE	0.8
2	BLISKI ZERA	NISKA	WYSOKA	ZWIĘKSZENIE	0.6
3	UMIARKOWANY	NISKA	WYSOKA	BEZ-ZMIANY	0.8
4	DUŻY	UMIARKOWANA	NISKA	ZMNIEJSZENIE	0.85
5	BARDZO DUŻY	UMIARKOWANA	NISKA	ZMNIEJSZENIE	0.95
6	UJEMNY	UMIARKOWANA	DOŚĆ DOBRA	BEZ-ZMIANY	0.9
7	BLISKI ZERA	UMIARKOWANA	DOŚĆ DOBRA	BEZ-ZMIANY	0.8
8	UMIARKOWANY	UMIARKOWANA	DOŚĆ DOBRA	BEZ-ZMIANY	0.7
9	DUŻY	DOŚĆ DUŻA	NISKA	ZMNIEJSZENIE	1
10	BARDZO DUŻY	DOŚĆ DUŻA	NISKA	ZMNIEJSZENIE	1
11	UJEMNY	DOŚĆ DUŻA	WYSOKA	ZMNIEJSZENIE	0.85
12	BLISKI ZERA	DOŚĆ DUŻA	WYSOKA	ZMNIEJSZENIE	0.9
13	UMIARKOWANY	DOŚĆ DUŻA	WYSOKA	ZMNIEJSZENIE	0.95
14	DUŻY	DUŻA	NISKA	ZMNIEJSZENIE	1
15	BARDZO DUŻY	DUŻA	DOŚĆ DOBRA	ZMNIEJSZENIE	1
16	UJEMNY	DUŻA	DOŚĆ DOBRA	ZMNIEJSZENIE	0.9
17	BLISKI ZERA	DUŻA	DOŚĆ DOBRA	ZMNIEJSZENIE	0.95
18	UMIARKOWANY	DUŻA	DOŚĆ DOBRA	ZMNIEJSZENIE	1
19	DUŻY	BARDZO DUŻA	NISKA	ZMNIEJSZENIE	1
20	BARDZO DUŻY	BARDZO DUŻA	NISKA	ZMNIEJSZENIE	1

Interesujące możliwości zastosowań mogą wynikać z połączenia korzystnych właściwości sieci neuronowych i systemów relacji rozmytych [13...20]. Obecnie w wielu ośrodkach naukowych prowadzone są intensywne badania nad różnorodnymi formami integracji tych systemów.

6. Obiektowe rozwiązanie problemu w języku C++

Klasa zbiorów opisuje obiekty określonego typu i zawiera wirtualną klasę typu elementów, podstawowe funkcje do określania rodzaju elementów i przeprowadzania na nich operacji.

Dostęp do składników klasy jest następujący [21, 22]: składniki zadeklarowane jako publiczne są dostępne we wszystkich częściach programu, składniki zadeklarowane jako prywatne są dostępne tylko w ciele składników ich klasy i funkcji zaprzyjaźnionych, składniki zadeklarowane jako chronione są w tej klasie dostępne tak jak składniki prywatne, z tą różnicą, że będą one dziedziczone i znajdują się w klasie potomnej.

Składnikami publicznymi są między innymi: konstruktory i destruktory obiektów, funkcje określania cech zbiorów, wprowadzania i wyprowadzania danych. Składnikami prywatnymi są typy funkcji przynależności zaś składnikami chronionymi są dane do prezentacji określonej zmiennej. Oddzielną klasę utworzono dla zmiennych lingwistycznych a kolejną dla operatorów.

Program umożliwia edycję i określanie wartości zmiennych lingwistycznych, relacji i operatorów, zapisywanie zestawów danych oraz graficzną prezentację wyników w postaci funkcji przynależności dla zmiennych lingwistycznych, będących decyzjami sterowania (postępowania).

7. Podsumowanie

Automatyczne wnioskowanie jest warunkiem dużej elastyczności systemów wytwarzania. Cechy tej nie posiadają układy, których działanie opisano w sposób zdeterminowany i, gdy działanie to nie jest przez określony czas zmieniane. Procesy wytwarzania są ze swej natury procesami, w których złożony proces kumulacji i propagacji zakłóceń jest przyczyną nieściśłości ich opisu. W pracy podkreślono, iż w takiej sytuacji wykorzystanie teorii zbiorów rozmytych pozwala właściwiej charakteryzować problemy oraz tworzyć algorytmy i systemy automatycznego podejmowania decyzji. Przedstawiona metodyka oraz podane przykłady wskazują na możliwości wykorzystania teorii zbiorów rozmytych.

8. Literatura

1. Bolc L., Brodziewicz W., Wójcik M.: Podstawy przetwarzania informacji niepewnej i niepełnej. Wydawnictwo Naukowe PWN, Warszawa 1991.
2. Kahneman D., Tversky A.: On the psychology of prediction. *Psychological Review* 1973, vol.80, s.123-141.
3. Kacalak W.: Mechanizm kumulacji skutków zakłóceń losowych oraz zmian warunków obróbki w aspekcie optymalizacji procesów skrawania. *Zeszyty Naukowe Politechniki Poznańskiej. Politechnika Poznańska nr 30/84 Poznań, 1984 r. s. 145-149.*
4. Kacalak W.: Wyznaczenie trwałości narzędzi ściernych z uwzględnieniem losowego charakteru procesu szlifowania z nałożonymi warunkami ograniczającymi. *Zeszyty Naukowe Politechniki Poznańskiej. Politechnika Poznańska nr 30/84 Poznań, 1984 r. s. 151-156.*
5. Kacalak W.: Własności eksploatacyjne narzędzi ściernych i problemy ich oceny. *Materiały VII Szkoły Naukowej Obróbki Ściernej. Wyższa Szkoła Inżynierska w Koszalinie. 1985 r., s. 113-132.*

6. Kacalak W.: Technologiczne i ekonomiczne korzyści optymalizacji szlifowania z uwzględnieniem probabilistycznego charakteru procesu. Materiały PAN, Sekcja PT KBM. IX Szkoła Naukowa Obróbki Ściernej. Kraków, 1986 r., s. 113-120.
7. Zadeh L.A.: Fuzzy Sets. Information and Control 8/1965, s.338-353.
8. Zadeh L.A.: Making Computers Think Like People. IEEE Spectrum. Vol. 21, 8/1984, s.26-32.
9. Mayer A., Mechler B., Schlinwein A., Wolke R.: Fuzzy Logic. Addison-Wesley, 1993.
10. Palm R., Hellendorn H.: Fuzzy-Control - Grundlagen und Entwicklungsmethoden. KI 4/1991, s.18-23.
11. Tilli T.: Fuzzy-Logic: Grundlagen, Anwendungen, Hard- und Software. Monachium, Franzis, 1991.
12. Trautzi G.: Werkzeuge für Fuzzy-Logic-Entwicklungen. KI 4/1991, s.69-72.
13. Kosko B.: Fuzzy Entropy and Conditioning. Information Sciences, vol. 40, s.165-174, 1986..
14. Yasuhiko Dote Y., James C. Hung J.C.: Autonomous Fuzzy Grasping Force Controller with Neural Networks for Manipulator Hand. First Workshop on Neural Networks. Auburn University, 1990.
15. Li-Min Fu Li-Min: Back-Propagation in Neural Networks with Fuzzy Conjunction Units. IJCNN, 1990, San Diego, vol.I, s. 613-618.
16. Toshiaki Iwata T., Machida K., Toda Y.: Fuzzy Control Using Neural Network Techniques. IJCNN, 1990, San Diego, vol.III, s. 365-370.
17. Seong-Gon Kong S., Kosko B.: Comparison of Fuzzy and Neural Truck Backer-Upper Control Systems. IJCNN, 1990, San Diego, vol.III, s. 349-358.
18. Lee C.: Fuzzy Logic in Control Systems: Fuzzy Logic Controller -Part I, II. SMC, vol. 20, 2/1990 s.404-418, 419-435.
19. Takagi H.: Fusion Technology of Fuzzy Theory and Neural Networks: Survey and Future Directions. Fuzzy Logic and Neural Networks. Izzuka 90, 1990.
20. Xingui H.: Fuzzy Computational Reasoning And Neural Networks Tools for Artificial Intelligence. Herndon VA, 1990, s. 706-711.
21. Borland C++ for Windows. Zbiór materiałów do Borland C++ 4.0.
22. Lipman S.B.: Podstawy języka C++. WNT, Warszawa 1993.

METHOD OF PROCESSING OF UNCERTAIN INFORMATION AND FORMATION OF ALGORITHMS TO CONTROL THE GRINDING PROCESSES USING THE FUZZY SETS THEORY

Summary

An important purpose of numerous application of new mathematics and informatics tools is the automation of the inference and control under the conditions of disposing of uncertain and inaccurate information. Application of typical method to the fuzzy sets theory in higher detail describes many problems, such as e.g. propagation and cumulation of the uncertainty in grinding processes. The presented method of forming the algorithms and delivered examples may provide for new applications of these effective inference method.

ALGORYTM ROZMYTEJ KOMPENSACJI NIEREGULARNYCH ZAKŁÓCEŃ W UKŁADACH TECHNOLOGICZNYCH

Procesy technologiczne zawsze podlegają różnorodnym zakłóceniom. Poziom zakłóceń można oceniać na podstawie stopnia zmienności wielkości wyjściowych. W pewnych przypadkach, opisanych w pracy, gdy poziom zakłóceń jest znaczny, opis procesu opiera się na informacjach niepewnych i nieściśle. Kompensacja rozmyta zakłóceń w takich warunkach jest dobrym i stosunkowo prostym rozwiązaniem problemu podwyższania dokładności procesów technologicznych. W pracy przedstawiono przykładowy algorytm postępowania oraz istotniejsze fragmenty programu w języku C++.

1. Zakłócenia w układach i procesach technologicznych

Ogólnie opis procesu można wyrazić następująco:

$$y_j = f_j(x_{i1}, w_{i2}, z_{i3}),$$

gdzie x_{i1} to parametry procesu, czyli wielkości mierzalne i sterowalne, w_{i2} to warunki procesu, czyli wielkości mierzalne lecz niesterowalne oraz z_{i3} to zakłócenia, czyli wielkości niemierzalne i niesterowalne, których liczby ani oddziaływań nie znamy. Miernikiem stopnia losowości rozpatrywanych procesów może być średnia wartość stosunku odchylenia standardowego wielkości wyjściowej do wartości oczekiwanej. W wielu procesach np. w procesach dokładnego szlifowania stosunek ten może osiągać wartości zbliżające się do 1 [2, 3].

Analizując rozkład określonej cechy jako wynik kumulacji oddziaływań wielu przyczyn rozproszenia jej wartości, w procesie uznanym za stacjonarny, można wyróżnić następujące przypadki modelowe [2, 3]:

- ♦ rozproszenie wartości obserwowanej cechy jest wynikiem sumowania się wielu oddziaływań losowych, przy czym rozróżnić należy sytuacje, w których jeden lub kilka czynników posiada wpływ dominujący, od przypadku, gdy żaden z czynników nie dominuje nad pozostałymi,
- ♦ rozproszenie wartości badanej cechy jest wynikiem geometrycznego sumowania się oddziaływań losowych,
- ♦ rozkład wartości badanej cechy zależy od iloczynu dużej liczby zmiennych, czyli jest skutkiem działania mechanizmu multiplikatywnego,

* Wyższa Szkoła Inżynierska w Koszalinie, ul. Raławicka 15/17, 75-620 Koszalin

- ♦ rozkład wynikowy jest złożoną funkcją oddziaływań, np. rozkładem sum iloczynów wektorowych, funkcją wykładniczą, funkcją potęgową,
- ♦ rozkład wartości obserwowanej cechy jest rozkładem wartości ekstremalnych skutków zakłóceń.

W warunkach niezależności czynników wywołujących odchyłki dokładności, jeżeli żaden z czynników nie dominuje nad pozostałymi, to w przypadku addytywnego mechanizmu kumulacji nawet bez znajomości liczby składowych losowych i ich rozkładów, można zakładać, iż rozkład sumy skutków jest zbliżony do normalnego. W procesach szlifowania addytywny mechanizm kumulacji skutków zakłóceń nie występuje zbyt często, zaś skłonność do posługiwania się takim modelem jest nadmierna.

Geometryczne sumowanie się zakłóceń jest charakterystycznym mechanizmem kumulacji przemieszczeń elementarnych narzędzia i przedmiotu o losowym kierunku i wartości. Typowym modelem rozkładu odpowiadającym multiplikatywnemu mechanizmowi kumulacji skutków jest rozkład logarytmo-normalny. Ten model kumulacji skutków zakłóceń jest wystarczająco dobry do opisu rozproszenia takich cech jak: wymiary ziaren ściernych, odchyłki wymiarów i kształtu, trwałość narzędzi dla niektórych kryteriów trwałości.

W procesach szlifowania występuje zazwyczaj bardzo złożony mechanizm kumulacji skutków zarówno losowych, harmonicznym i monotonicznym przyczyn niedokładności. Składowa deterministyczna jest tym większa im większe są skutki odkształceń cieplnych i mechanicznych obrabianego przedmiotu oraz niedokładności obrabiarki. Składowa losowa w największym stopniu zależy od zużycia ściernicy oraz losowych przyczyn niedokładności kinematycznej układu technologicznego, przy czym rozkład składowej losowej jest zbliżony do rozkładu logarytmo-normalnego.

W praktyce eksperymentalnej stwierdzenie "proces może być uznany za stacjonarny" jest związane z przyjęciem określonej, dopuszczalnej niedokładności oceny. Rzeczywiste procesy nie są procesami stacjonarnymi. Błędy wynikające z przyjęcia założenia o stacjonarności procesów zależą również od momentu rozpoczęcia oraz okresu zbierania danych. Rozkład sumaryczny różni się od rozkładów chwilowych tym bardziej, im dłuższy jest okres zbierania danych oraz im mniej stabilne są składowe systematyczne. Składowe te mogą powodować wzrost wariancji rozkładu sumarycznego, czyli momentu rzędu drugiego, zmianę współczynnika asymetrii - czyli momentu rzędu trzeciego oraz zmianę współczynnika spłaszczenia - momentu rzędu czwartego. Niestabilność składowych losowych powoduje natomiast zmianę postaci rozkładu.

Kompensacja zakłóceń w precyzyjnych procesach technologicznych jest dość częstym postulatem użytkowników i ważnym zadaniem producentów układów technologicznych. Jest ona dość łatwa w układach o stabilnych charakterystykach, zwłaszcza gdy dotyczy wielkości łatwo sterowalnych. Wprowadzenie kompensacji staje się trudnym zadaniem, gdy charakterystyka układu jest zmienna (np. sztywność wynikowa układu technologicznego w procesie szlifowania

wałków, czy długich śrub zmienia się wraz z przemieszczaniem ściernicy wzdłuż przedmiotu), oraz gdy wielkość kompensowanego odchylenia staje się porównywalna z wpływem nieznanego zakłócenia lub jest zbliżona do elementarnej wartości korekty.

Przykładem takiego problemu może być zadanie kompensacji względnych, promieniowych odchyień położenia w procesie precyzyjnego szlifowania śrub lub wałków o znacznej długości, gdy przyrosty tego odchylenia są bardzo małe a sumaryczne maksymalne odchylenie nie przekracza 2-3 μm . W takim przypadku nie są przydatne typowe układy do realizacji przemieszczeń. Właściwym rozwiązaniem może być kompensacja realizowana poprzez wprowadzenie do systemu sprężystego lub cieplnego kompensatora. W opisanej sytuacji opis układu staje się jednak niedostatecznie ścisły i trudno jest wyprowadzić analityczne rozwiązanie systemu regulacji, zwłaszcza gdy problem jest niestacjonarny. Trzeba przy tym też podkreślić, iż w takim wypadku wyznaczanie analitycznego opisu jest bardzo kosztowne i długotrwałe.

Zastosowanie opisanej poniżej kompensacji rozmytej jest korzystniejszą alternatywą, stosunkowo prostą w realizacji, której działanie w bardziej rozbudowanych algorytmach (z wykorzystaniem właściwości sieci neuronowych) może podlegać adaptacji, będącej skutkiem uczenia się systemu.

2. Algorytm kompensacji rozmytej

Kompensacja rozmyta polega na wykorzystaniu systemu przetwarzania informacji niepewnej i nieścisłej w sposób opisany poniżej.

- ▶ Pierwszym zadaniem jest określenie wielkości (zmiennej wyjściowej) wywołującej efekt kompensacji. Dla przykładu przyjęto, iż zmienną taką będzie siła kompensacji w odkształcalnym specjalnym układzie, powodująca zmniejszanie odchyłki położenia przedmiotu i narzędzia.
- ▶ Kolejnym zadaniem jest określenie zmiennych wejściowych (wielkości mierzonych), które wpływać będą na wynik działania kompensującego. Przyjęto w przykładzie rozwiązania, iż cechami tymi będą: chwilowe wartości odchyłki położenia i jej pierwszej pochodnej.
- ▶ Następnym zadaniem staje się określenie relacji rozmytych między zmiennymi wejściowymi a zmienną wyjściową.

Zarówno zmienne wejściowe, jak i zmienna wyjściowa są zmiennymi lingwistycznymi, definiowanymi przez podanie między innymi nazwy zmiennej lingwistycznej oraz termów, czyli nazw rozmytych zbiorów z $[0,1]$, lub inaczej "nazw" wartości zmiennej lingwistycznej a także przez podanie przestrzeni obiektów, w której określane są te zbiory.

W opracowanym w systemie Borland C++ [6, 7] programie KOMP-ROZ zdefiniowano klasy: Zmienna, Term i RelacjaRozmyta, a następnie utworzono

obiekty dla tych klas, odpowiadające poszczególnym zmiennym, przypisanym im termom oraz relacjom:

```
Zmienna ZmLingwistSila("Sila F"); // definicja obiektu ZmLingwistSila klasy Zmienna
// definicje obiektów SilaUjemna, SilaBliskaZeru i SilaDodatnia klasy Term
Term SilaUjemna (&ZmLingwistSila, "F<0", TYP_Z, aS1, deltaS1);
Term SilaBliskaZeru (&ZmLingwistSila, "F=0", TYP_PI, aS0, deltaS0);
Term SilaDodatnia (&ZmLingwistSila, "F>0", TYP_S, aS2, deltaS2);

Zmienna ZmLingwistDeltaY("DeltaY"); // definicja obiektu ZmLingwistDeltaY klasy Zmienna
// definicje obiektów DeltaYUjemna, DeltaYZerowa i DeltaYDodatnia klasy Term
Term DeltaYUjemna (&ZmLingwistDeltaY, "DeltaY<0", TYP_Z, aDY1, deltaDY1);
Term DelatYZerowa (&ZmLingwistDeltaY, "DeltaY=0", TYP_PI, aDY0, deltaDY0);
Term DeltaYDodatnia(&ZmLingwistDeltaY, "DeltaY>0", TYP_S, aDY2, deltaDY2);

Zmienna ZmLingwistPredkY("Predk.DeltaY"); // def. obiektu ZmLingwistPredkY klasy Zmienna
// definicje obiektów PredkDeltaYUjemna, PredkDeltaYZerowa i PredkDeltaYDodatnia
// klasy Term
Term PredkDeltaYUjemna (&ZmLingwistPredkY, "vDeltaY<0", TYP_Z, aPDY1, deltaPDY1);
Term PredkDeltaYZerowa (&ZmLingwistPredkY, "vDeltaY=0", TYP_PI, aPDY0, deltaPDY0);
Term PredkDelatYDodatnia (&ZmLingwistPredkY, "vDeltaY>0", TYP_S, aPDY2, deltaPDY2);
```

Konstruktor klasy Term przyjmuje następujące argumenty: wskazanie do zmiennej lingwistycznej, której składnikiem jest dany term, nazwę termu, typ funkcji przynależności (TYP_Z, TYP_PI, TYP_S) oraz jej dwa parametry: a (aS1, ..., aPDY2) i delta (deltaS1, ..., deltaPDY2), co pokazano niżej w definicji klasy Term.

```
class Term // klasa wartości zmiennych lingwistycznych zwanych termami
{
private:
    Zmienna *WskDoZmLingwist; // wskazanie do zmiennej lingwistycznej
    char NazwaTermu[30]; // nazwa termu
    int TypFu; // typ funkcji przynależności
    float a, delta; // charakterystyczne wartości termu
    float MaxWartFunPrzypZRelacji; // max.wartość.funkcji.przynależności jako wynik relacji
    Term *NastepnyTerm; // wskazanie do następnego termu
public:
    Term(Zmienna *p, char *Nazwa, int TypFuPn, float Wa, float Wdelta); // konstruktor klasy Term
    void InicjalizacjaListyTermow(Term *t); // inicjacja wskazań do termów
    Term *NastTerm(void); // funkcja do określenia wskazania do następnego termu
    char *Nazwa(void); // funkcja do przekazywania wskazania do nazwy termu
    float WartFunPrzynal(void); // funkcji do określania wartości funkcji przynależności termu
    float WartFunPrzynal(float Wart); // funkcja określania wartości funkcji przynależności
    // z uwzględnieniem zastosowania operatorów
    float Min(void) { return(a-delta); } // określanie minimum wartości termu
    float Max(void) { return(a+delta); } // określanie maksimum wartości termu
    // określenie max.wart.funkcji przynależn.jako wyniku relacji rozmytych
    void WyznaczMaxRelacji(float max; // inicjalizacja (zerowanie) max.wart.wyników relacji rozmytych
    void InicjalizacjaMaxRelacji(void);
};
```

Konstruktor klasy Term jest następujący:

```

// konstruktor klasy Term
Term::Term(Zmienna *p, char *Nazwa, int TypFuP, float Wa, float Wdelta)
{
    WskDoZmLingwist=p; // wskazanie do zmiennej lingwistycznej
    strcpy(NazwaTermu, Nazwa); // skopiowanie nazwy termu
    TypFu=TypFuP; // typ funkcji przynależności
    a=Wa, delta=Wdelta; // parametry charakterystyczne funkcji przynależności
    WskDoZmLingwist->InicjalizacjaListyTermow(this); //aktualizacja listy termów
// wskaźnik this jest dodatkowym ukrytym argumentem funkcji składowych klasy wskazującym
// na obiekt, który wywołuje daną funkcję
    NastepnyTerm=0; // zakończenie listy
    MaxWartFunPrzypZRelacji=0; // początkowa wart.max.funkcji.przyn.w wyniku relacji rozmytej
}

```

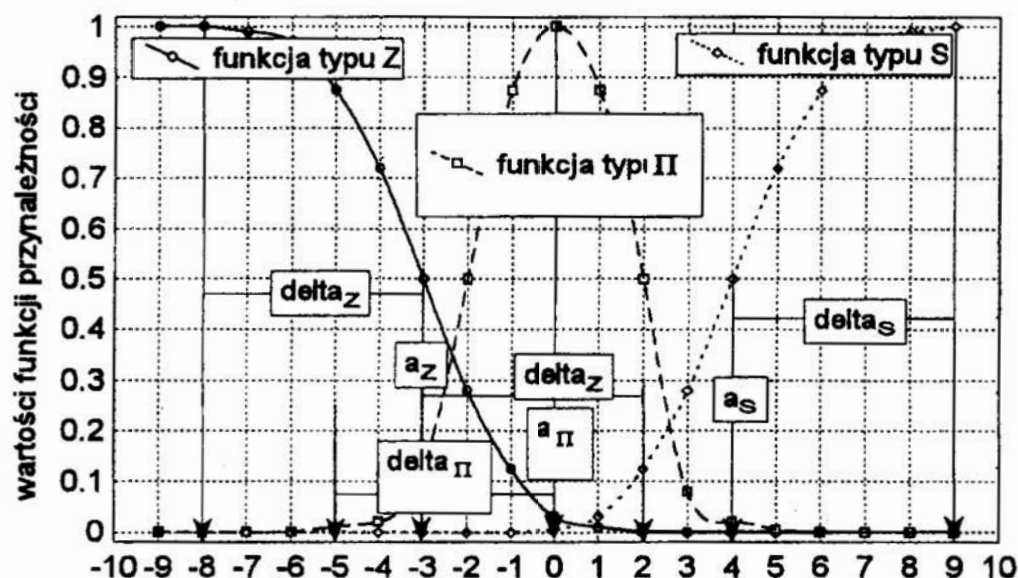
Zastosowanymi funkcjami przynależności [1, 4, 5] były (rys.1):

- ♦ funkcja typu S, określona poniżej

$$\mu_S(x, a, \delta) = \begin{cases} 0, & x < a-\delta, \\ 2((x-a+\delta)/2\delta)^2, & x \in [a-\delta, a), \\ (1-2((a-x+\delta)/2\delta))^2, & x \in (a, a+\delta], \\ 1, & x > a+\delta, \end{cases}$$

- ♦ funkcja typu Z: $\mu_Z(x, a, \delta) = 1 - \mu_S(x, a, \delta)$,
- ♦ funkcja typu II zdefiniowana następująco:

$$\mu_{II}(x, a, \delta) = \begin{cases} \mu_S(x, a - \delta/2, \delta/2) & x < a \\ \mu_Z(x, a + \delta/2, \delta/2) & x \geq a. \end{cases}$$



Rys. 1. Funkcje przynależności S, Z i II

Fig.1. Membership function S, Z and II

Działanie sieci z jedną warstwą ukrytą i neuronami z liniowymi funkcjami opisują następujące równania:

$$X = W_H \cdot I \quad (7)$$

$$y = W_O \cdot X \quad (8)$$

gdzie: X - wektor sygnałów wyjściowych warstwy ukrytej,
 y - sygnał wyjściowy sieci (w tym przypadku skalar),
 W_H - wektor wag pomiędzy wejściem a warstwą ukrytą,
 W_O - wektor wag pomiędzy warstwą ukrytą a wyjściem,

Przeprowadzone eksperymenty wskazują, że podczas projektowania sieci należy wziąć pod uwagę następujące wytyczne:

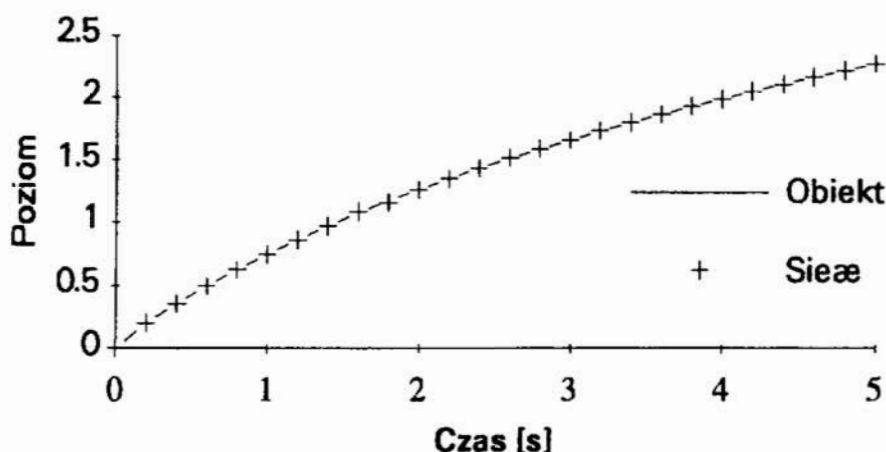
- liczba wejść nie mniejsza niż rząd obiektu;
- liczba warstw ukrytych - 1;
- liczba neuronów w warstwie ukrytej nie mniejsza niż liczba wejść;
- współczynnik uczenia (learning rate) od 0.005 do 0.5;
- współczynnik stałego kierunku modyfikacji (momentum term) od 0.01 do 0.5;
- sygnał sterujący uczący - sygnał losowy o równomiernym rozkładzie prawdopodobieństwa.

Jako algorytmu uczenia używa się Wstecznej Propagacji Błędów.

Dla neuronowego modelu obiektu zbiornika ze swobodnym odpływem opisanego równaniem (6) przyjęto następujące parametry: liczba wejść - 3, liczba neuronów w warstwie ukrytej - 3, learning rate - 0.2, momentum term - 0.1.

Symulator zaprojektowanej sieci neuronowej został napisany w języku Borland Pascal. Dla tego obiektu proces modelowania prowadzono wykorzystując losowy sygnał sterujący z zakresu od 0 (zawór dopływowy zamknięty) do 1 (zawór w pełni otwarty) i zakończono po 100.000 kroków.

Na rys. 5 zaprezentowano odpowiedź modelu neuronowego zbiornika ze swobodnym odpływem na tle odpowiedzi obiektu wyliczonej z równania (6) dla wymuszenia w postaci skoku jednostkowego (pełne otwarcie zaworu). W tym przypadku wartość błędu liczonego według wzoru (5) wynosi 0.00008.



Definicję klasy zmiennych lingwistycznych i następnie klasy relacji rozmytych podano poniżej.

```

class Zmienna // klasa zmiennych lingwistycznych
{
protected:
    char NazwaZmiennej[30]; // nazwa zmiennej lingwistycznej
    float AktWartZmiennej; // aktualna wartość zmiennej
    Term *PierwszyTerm; // wskazanie do pierwszego i ostatniego
    Term *OstatniTerm; // termu w łańcuchu zdefiniowanych termów

public:
    Zmienna(char *Nazwa); // konstruktor klasy Zmienna
    void NowyTerm(Term *t);
    void InicjalizacjaListyTermow(Term *t);
    Term *SzukajTermu(char *Nazwa); // poszukiwanie termu o określonej nazwie
    float Wartosc(void) { return(AktWartZmiennej); } // aktualna wartość zmiennej
    void NowaWart(float PodanaWartosc)
    { AktWartZmiennej = PodanaWartosc; } // nowa wartość zmiennej
    float ObliczPunktCiezk(void); // przejście do wyniku nierozmytego
    float fmax(float x); // maksymalna wartość funkcji przynależności z wszystkich
    // termów danej zmiennej rozmytej
    // inicjalizacja wszystkich termów zmiennej lingwistycznej
    void InicjalizacjaTermow(void);
};

class RelacjaRozmyta // klasa relacji rozmytych
{
protected:
    Term *TermWyjscia; // wskazanie do termu zmiennej wyjściowej
    Term *TermWejscia1; // wskazanie do termu zmiennej wejściowej_1
    Term *TermWejscia2; // wskazanie do termu zmiennej wejściowej_2
    OP_TYP Operator; // typ operatora
    float WartFunPrzynWyniku; // wartość funkcji przynależności jako wyniku relacji
    RelacjaRozmyta *NastepnaRelacja; // wskazanie do następnej relacji

public:
    RelacjaRozmyta(Zmienna *Wyjscie, char *NazwaWyjscia,
        Zmienna *ParWE1, char *NazwaWejscia1,
        OP_TYP TypOperatora,
        Zmienna *ParWE2, char *NazwaWejscia2);

    float WynikRelacjiRozmytej(void);
    RelacjaRozmyta *NastRel(void); // wskazanie do następnej relacji
    void InicjujRelacje(RelacjaRozmyta *Nastepna)
    { NastepnaRelacja=Nastepna; } // utworzenie listy relacji rozmytych
    float Wartosc(void) { return(WartFunPrzynWyniku); } // przekazanie wartości funkcji
    // przynależności jako w wyniku relacji
};

```

Konstruktor klasy RelacjaRozmyta jest następujący:

```
RelacjaRozmyta::RelacjaRozmyta(Zmienna *Wyjscie, char *NazwaWyjscia,  
                                Zmienna *ParWE1, char *NazwaWejscia1,  
                                OP_TYP TypOperatora,  
                                Zmienna *ParWE2, char *NazwaWejscia2)  
{  
    Operator = TypOperatora; // przekazany typ operatora  
    // określenie wskaźników do termów wejścia (1i2) oraz wyjścia  
    TermWejscia1 = ParWE1->SzukajTermu(NazwaWejscia1);  
    TermWejscia2 = ParWE2->SzukajTermu(NazwaWejscia2);  
    TermWyjscia = Wyjscie->SzukajTermu(NazwaWyjscia);  
    WartFunPrzynWyniku = 0; // początkowa wart.funkcji przynależności wyniku  
    if(PoprzedniaRelacja != NULL)  
        PoprzedniaRelacja->InicjujRelacje(this); // dołączenie nowej relacji  
    PoprzedniaRelacja = this; // rozbudowa listy  
    if(PierwszaRelacja==0) PierwszaRelacja=this;  
}
```

Dla rozpatrywanego problemu zdefiniowano dziewięć relacji opisujących wpływ wartości i znaku przemieszczenia oraz wartości i kierunku prędkości przemieszczenia na kierunek i wartość siły kompensacji.

```
RelacjaRozmyta Rel1(&ZmLingwistSila, "F<0", &ZmLingwistDeltaY, "DeltaY<0",  
                    AND, &ZmLingwistPredkY, "vDeltaY<0");  
// RelacjaRozmyta 2: wychylenie w prawo i ruch na prawo -> siła F>0  
RelacjaRozmyta Rel2(&ZmLingwistSila, "F>0", &ZmLingwistDeltaY, "DeltaY>0",  
                    AND, &ZmLingwistPredkY, "vDeltaY>0");  
// RelacjaRozmyta 3: wychylenie zerowe i bez ruchu -> F=0  
RelacjaRozmyta Rel3(&ZmLingwistSila, "F=0", &ZmLingwistDeltaY, "DeltaY=0",  
                    AND, &ZmLingwistPredkY, "vDeltaY=0");  
// RelacjaRozmyta 4: wychylenie w prawo i ruch na lewo -> siła F=0  
RelacjaRozmyta Rel4(&ZmLingwistSila, "F=0", &ZmLingwistDeltaY, "DeltaY>0",  
                    AND, &ZmLingwistPredkY, "vDeltaY<0");  
// RelacjaRozmyta 5: wychylenie w lewo i ruch na prawo -> siła F=0  
RelacjaRozmyta Rel5(&ZmLingwistSila, "F=0", &ZmLingwistDeltaY, "DeltaY<0",  
                    AND, &ZmLingwistPredkY, "vDeltaY>0");  
// RelacjaRozmyta 6: wychylenie zerowe i ruch na lewo -> F<0  
RelacjaRozmyta Rel6(&ZmLingwistSila, "F<0", &ZmLingwistDeltaY, "DeltaY=0",  
                    AND, &ZmLingwistPredkY, "vDeltaY<0");  
// RelacjaRozmyta 7: wychylenie zerowe i ruch na prawo -> F>0  
RelacjaRozmyta Rel7(&ZmLingwistSila, "F>0", &ZmLingwistDeltaY, "DeltaY=0",  
                    AND, &ZmLingwistPredkY, "vDeltaY>0");  
// RelacjaRozmyta 8: wychylenie na lewo i bez ruchu -> F<0  
RelacjaRozmyta Rel8(&ZmLingwistSila, "F<0", &ZmLingwistDeltaY, "DeltaY<0",  
                    AND, &ZmLingwistPredkY, "vDeltaY=0");  
// RelacjaRozmyta 9: wychylenie na prawo i bez ruchu -> F>0  
RelacjaRozmyta Rel9(&ZmLingwistSila, "F>0", &ZmLingwistDeltaY, "DeltaY>0",  
                    AND, &ZmLingwistPredkY, "vDeltaY=0");
```


Działanie programu KOMP-ROZ rozpoczyna się inicjalizacją utworzonych obiektów: zmiennych, ich termów oraz relacji. Konstruktor klasy Zmienna inicjuje początkowy stan obiektu:

```

Zmienna::Zmienna(char *Nazwa) // konstruktor klasy Zmienna
{
    strcpy(NazwaZmiennej, Nazwa); // skopiowanie nazwy zmiennej
    AktWartZmiennej=0; // początkowa wartość
    // początkowe wartości wskaźnik- brak termów na liście
    PierwszyTerm = OstatniTerm = NULL;
}

```

Następnie dla wszystkich zmiennych i odpowiadających im termów wywoływana jest funkcja NowyTerm, będąca składnikiem publicznym klasy Zmienna, która tworzy listę termów dla każdej zmiennej lingwistycznej.

```

void Zmienna::NowyTerm(Term *t)
{
    OstatniTerm->InicjalizacjaListyTermow(t); // wskazanie do poprzedniego termu
    OstatniTerm=t; // nowy term dołączony do listy
    if(PierwszyTerm==NULL) // jeżeli to już koniec listy to
        PierwszyTerm=t; // inicjalizacja pierwszego
}

```

Konstruktor klasy RelacjaRozmyta powoduje inicjowanie relacji rozmytych i przypisanie do nich odpowiednich wskaźników.

Obliczenia przebiegają następująco:

```

for(n=1; n<NMAX+1 ;n++) //do wykonania wszystkich kroków symulacji
{
    DeltaDY=PrzyrostyDY(); // przyrosty jako wynik zakłóceń
    DY+=DeltaDY; // nowa wartość odchyłki położenia przed kompensacją
    SumaDY+=DeltaDY; // odchyłka sumaryczna
    // obliczenia wartości zmiennej lingwistycznej SILA
    SILA = Obliczenia(&ZmLingwistSila, &ZmLingwistDeltaY, DY,
        &ZmLingwistPredkY, d_DY);
    // wyznaczenie nowej wartości odchyłki po kompensacji lub w przypadku
    // symulacji rozwiązanie równania różniczkowego opisującego obiekt
    RozRR2s_RK(SILA, &Czas, &DY, &d_DY, KrokKompensacji, Wart2Pochodnej);
    // wyznaczenie wart.drugiej pochodnej przemieszczenia
    d2_DY=Wart2Pochodnej(t, DY, d_DY);
    // zapamiętanie poprzednich wartości ( dla potrzeb prezentacji wyników)
    SumaDYOld=SumaDY;
    SILA_Old=SILA;
    DYOld = DY, d_DYOld = d_DY, d2_DYOld = d2_DY;
}

```

Funkcja Obliczenia otrzymuje jako parametry wskazania ParWE1, ParWE2 do zmiennych lingwistycznych wejścia oraz wskazanie Wyjscie do zmiennej wyjściowej a także wartości zmiennych wejścia vWE1 i vWE2. Poprzez sprawdzenie wszystkich relacji rozmytych poszukiwana jest relacja dająca największą wartość wyniku.

```
float Obliczenia(Zmienna *Wyjscie, Zmienna *ParWE1,
                float vWE1, Zmienna *ParWE2, float vWE2)
{
    RelacjaRozmyta *NastRel; // wskazanie do relacji rozmytej
    ParWE1->NowaWart(vWE1); // wartości zmiennych wejściowych
    ParWE2->NowaWart(vWE2);
    Wyjscie->InicjalizacjaTermow(); // inicjalizacja termów wyjścia
    // wykonanie wszystkich relacji rozmytych
    for(NastRel = PierwszaRelacja; NastRel!=NULL; NastRel = NastRel->NastRel())
        NastRel->WynikRelacjiRozmytej();
    }
    return(Wyjscie->ObliczPunktCiezk()); // obliczenie wyniku nierozmytego
}
```

Wynikiem relacji rozmytej, w przypadku operatora AND, jest minimum wartości funkcji przynależności obu składników relacji (w tym przypadku przemieszczenia DY i jego pochodnej d_DY). Wartością potrzebną dla dalszych operacji jest maksymalna wartość wyniku z wszystkich relacji rozmytych, bowiem interesuje nas ta relacja, która ma największy wpływ na wynik.

```
float RelacjaRozmyta::WynikRelacjiRozmytej(void)
{ float z1,z2; // wartości funkcji przynależności

    z1=TermWejscia1->WartFunPrzynal(); // przypisanie wartości wejścia
    z2=TermWejscia2->WartFunPrzynal();

    switch(Operator) // zależnie od typu relacji
    { case AND: // operator Minimum dla lingwistycznego AND
        WartFunPrzynWyniku=MIN(z1, z2); // (z1<z2)?(z1):(z2)
        break;
        case OR: // operator Maximum dla lingwistycznego OR
        WartFunPrzynWyniku=MAX(z1, z2); // (z1>z2)?(z1):(z2)
        break;
    } // ustalenie wartości maksymalnej funkcji przynależności wyniku
    TermWyjscia->WyznaczMaxRelacji(WartFunPrzynWyniku);
    return(WartFunPrzynWyniku);
}
```

Wynikiem kompensacji rozmytej jest wielkość nierozmyta, w tym przypadku wartość siły kompensacji przemieszczenia. Dla jej wyznaczenia wykonywana jest funkcja ObliczPunktCiezk będąca składnikiem publicznym klasy Zmienna.

```

// funkcja przetworzenia wnioskowania rozmytego w wynik nierozmyty
float Zmienna::ObliczPunktCiezk(void)
{
    const int LiKrokovCalk=100; // liczba kroków całkowania
    Term *Nast; // wskazanie do następnego termu na liście
    float minC, maxC; // granice całkowania
    float SumaDx=0, SumaYiDx=0, Yi; // zmienne w procesie całkowania
    float w, // zmienna pomocnicza
        Krok; // krok całkowania
    float x; // współrzędna x w procesie całkowania
    float fx, fxplusDx; // wartość funkcji przynależności dla x oraz dla x+dx

    minC = MAXFLOAT, maxC = -MAXFLOAT; // określanie granic całkowania wśród wszystkich termów
    for(Nast = PierwszyTerm; Nast != NULL; Nast = Nast->NastTerm())
    {
        if((w=Nast->Min())<minC) minC=w; // nowe minimum zakresu zmiennej
        if((w=Nast->Max())>maxC) maxC=w; // nowe maksimum zakresu zmiennej
    }
    Krok=(maxC-minC)/LiKrokovCalk; // długość kroku całkowania
    fx = fmax(minC); // max. wartość funkcji przynależności/ dla minC ze wszystkich termów
    for(x=minC+Krok; x<=maxC; x+=Krok)
    {
        fxplusDx = fmax(x); // max. wartość funkcji przynależności dla x ze wszystkich termów
        Yi = (fx + fxplusDx) / 2; // wartość średnia
        SumaDx += Yi, SumaYiDx += x * Yi;
        fx=fxplusDx;
    }
    if(SumaDx==0) {...; } // procedura informacji o braku poprawnych relacji
    SumaDx *= Krok; SumaYiDx *= Krok;
    return(SumaYiDx/SumaDx); // nie rozmyty wynik
}

```

W celu przejścia do wyniku nierozmytego wykorzystano średnią wartość funkcji przynależności, choć można też posługiwać się kryterium wartości maksymalnej. Do wyznaczenia średniej wartości funkcji przynależności wykorzystano funkcję `fmax`, która zwraca maksymalną wartość funkcji przynależności ze wszystkich termów zmiennej wyjściowej.

```

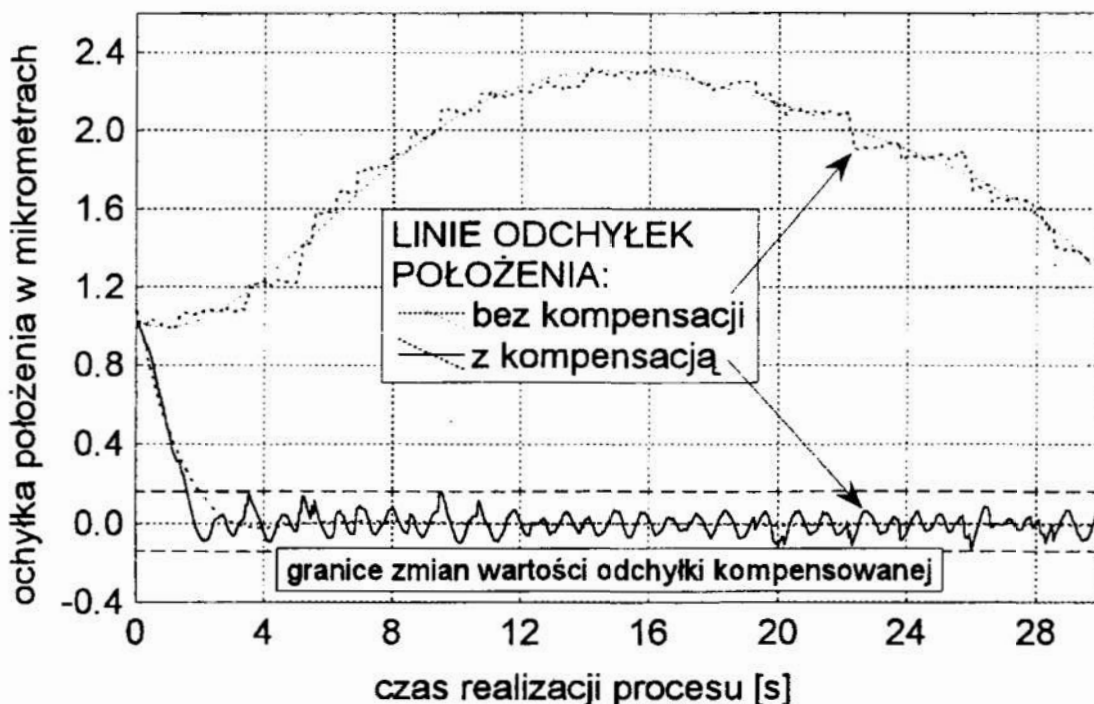
// obliczanie maksymalnej wartości funkcji przynależności
// ze wszystkich termów danej zmiennej
float Zmienna::fmax(float x)
{
    Term *Nast; // wskazanie do następnego termu na liście
    float max=0, wfp; // maksimum z przedziału [0, 1]
    for(Nast = PierwszyTerm; Nast != NULL; Nast = Nast->NastTerm())
        if((wfp = Nast->WartFunPrzynal(x)) > max)
            max = wfp;
    return(max);
}

```

Wartość funkcji przynależności obliczana jest z użyciem funkcji $WartFunPrzynal(x)$, przy czym wykorzystywane jest ograniczenie maksymalnej wartości funkcji przynależności, będące wynikiem relacji rozmytych.

```
float Term::WartFunPrzynal(float Wart) // oblicz.wartości funkcji przynależności
{ // przy ograniczeniu wart.maksymalnej
float Wynik, WartFu; // zmienne pomocnicze
// zerowa wartość funkcji przynależności poza przedziałem [min-max]
if((Wart<(a-delta)) || (Wart>(a+delta))) return(0);
if(Wart <= a)
{ WartFu=2 * pow((Wart-a+delta)/(2*delta), 2);
Wynik=(TypFu==TYP_S || TypFu==TYP_Z)? WartFu : 1-WartFu;
}
if(Wart > a)
{ WartFu=1-2*pow((-Wart+a+delta)/(2*delta),2);
Wynik=(TypFu==TYP_Z || TypFu==TYP_PI)? WartFu : 1-WartFu;
}
return (MIN(Wynik, MaxWartFunPrzynZRelacji)); // ograniczenie wart.fun.przynależności
}
```

Efekty zastosowania algorytmu kompensacji rozmytej przedstawiono w formie wykresu symulacji procesu kompensacji przemieszczenia względnego ściernicy i szlifowanego wałka o dużej długości (rys. 2).



Rys.2. Przykład symulacji kompensacji odkształceń mechanicznych szlifowanego wałka

Fig.2. Example of simulation of mechanical strains in the ground shaft

7. Podsumowanie

Automatyczne wnioskowanie jest współcześnie warunkiem dużej elastyczności systemów wytwarzania. Do niedawna elastycznymi nazywano takie systemy, które pozwalały na stosunkową łatwą zmianę procesu produkcyjnego, choć ich działanie opisane w sposób zdeterminowany uniemożliwiało adaptacją działania w warunkach silniejszych zakłóceń. Procesy wytwarzania są ze swej natury procesami podlegającymi zakłóceniom, wynikającym przede wszystkim z cieplnych i dynamicznych oddziaływań narzędzia i przedmiotu oraz zmian właściwości narzędzia podczas realizacji procesu. Analityczne rozwiązanie problemu kompensacji zakłóceń wymaga pełnego opisu systemu. W przypadku małych nieregularnych zakłóceń trudno jest wyznaczyć ściśle współzależności zakłóceń i ich skutków. Wskazane jest wówczas określenie relacji rozmytych i opracowanie systemu kompensacji rozmytej. Przedstawiony algorytm i przykład ilustrują zastosowanie tej metody.

8. Literatura

1. Bolc L., Brodziewicz W., Wójcik M.: Podstawy przetwarzania informacji niepewnej i niepełnej. Wydawnictwo Naukowe PWN, Warszawa 1991.
2. Kacalak W.: Mechanizm kumulacji skutków zakłóceń losowych oraz zmian warunków obróbki w aspekcie optymalizacji procesów skrawania. Zeszyty Naukowe Politechniki Poznańskiej. Politechnika Poznańska nr 30/84 Poznań, 1984 r. s. 145-149.
3. Kacalak W.: Wyznaczenie trwałości narzędzi ściernych z uwzględnieniem losowego charakteru procesu szlifowania z nałożonymi warunkami ograniczającymi. Zeszyty Naukowe Politechniki Poznańskiej. Politechnika Poznańska nr 30/84 Poznań, 1984 r. s. 151-156.
4. Mayer A., Mechler B., Schlinwein A., Wolke R.: Fuzzy Logic. Addison-Wesley, 1993.
5. Palm R., Hellendorn H.: Fuzzy-Control - Grundlagen und Entwicklungsmethoden. KI 4/1991, s.18-23.
6. Borland C++ for Windows. Zbiór materiałów do Borland C++ 4.0.
7. Lipman S.B.: Podstawy języka C++. WNT, Warszawa 1993.

ALGORITHM OF FUZZY COMPENSATION OF IRREGULAR DISTURBANCES IN TECHNOLOGICAL SYSTEMS

Summary

Analytical solution of the problem of disturbances compensation requires full description of the system properties. In case of small irregular disturbances in technological systems, it is difficult to determine strict relationships of the disturbances and their effects. It is advisable then to determine fuzzy relations and to work out a system of fuzzy compensation. The algorithm presented and the simulation example illustrate practical application of the method.

Józef GAWLIK

Krzysztof KARBOWSKI

Instytut Technologii Maszyn i Automatykacji Produkcji Politechniki Krakowskiej

Adam RUSZAJ

Instytut Obróbki Skrawaniem w Krakowie

KONCEPCJA STEROWANIA WYDAJNOŚCIĄ PROCESU OBRÓBKI ECM Z WYKORZYSTANIEM TEORII ZBIORÓW ROZMYTYCH

1. Wybrane pojęcia z teorii zbiorów rozmytych [2,5].

Logika rozmyta posługuje się pojęciami wieloznacznymi, nieostryimi jak np. "mała prędkość", "niskie napięcie", "dobra gładkość powierzchni" itp. Sformułowania tego typu są więc oceną a nie jednoznaczny, deterministycznym faktem.

Różnica pomiędzy logiką klasyczną a rozmytą tkwi w prawie wykluczonego środka. W klasycznej teorii zbiorów element należy (ocena "1") albo nie należy (ocena "0") do zbioru; zbiory tego typu mają ostre granice.

Elementy zbioru rozmytego należą do wielu rozłącznych zbiorów; granice zbiorów rozmytych zanikają łagodnie.

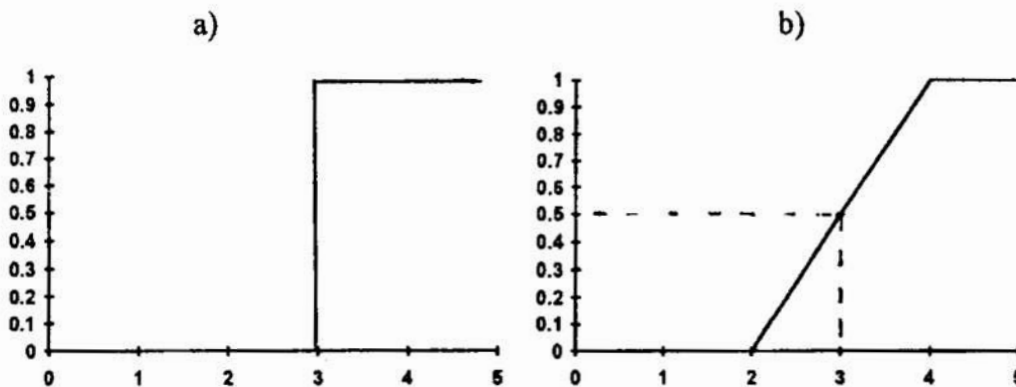
Logika rozmyta nakłada tylko jedno ograniczenie: suma stopni przynależności elementu do rozłącznych zbiorów musi wynosić jeden.

Stopień rozmycia zbioru nie jest miarą prawdopodobieństwa, lecz miarą zakresu zachodzenia lub występowania danych warunków.

Zbiorem rozmytym A nazywamy ogół par $\{x, \mu_A(x)\}$, przy czym x jest elementem zbioru, zaś $\mu_A(x)$ jest funkcją przynależności, która określa stopień przynależności elementu do zdarzenia (zbioru) A .

Funkcja przynależności $\mu_A(x)$ jest oszacowaniem z ciągłego przedziału $[0,1]$.

Graficzną interpretację funkcji przynależności w logice klasycznej (a) oraz w logice rozmytej (b) przedstawiono na rys. 1.

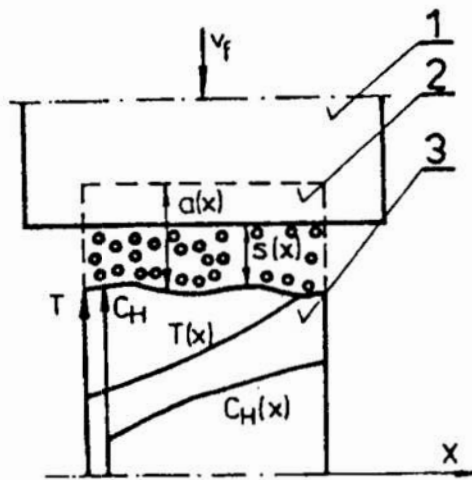


Rys. 1. Graficzna interpretacja funkcji przynależności.

Zgodnie z logiką klasyczną (a) element należy ($\mu_A(4)=1$), lub nie należy ($\mu_A(1)=0$) do zbioru A . W myśl zasad logiki rozmytej (b) stopień przynależności elementu do zbioru A zawiera się w przedziale $[0,1]$ (np. $\mu_A(3)=0.5$).

2. Czynniki wpływające na nieokreśloność procesu obróbki elektrochemicznej (ECM).

W obróbce elektrochemicznej naddatek jest usuwany w wyniku roztwarzania elektrochemicznego [6,7], któremu towarzyszą procesy transportu masy, ładunku elektrycznego oraz ciepła. Ciepło, produkty roztwarzania anodowego, wodór wydzielający się na katodzie są usuwane z obszaru obróbki przez przepływający elektrolit. Na skutek tego w szczelinie występują nierównomierne rozkłady temperatury, koncentracji produktów roztwarzania, koncentracji wodoru, właściwej przewodności elektrolitu, obrabialności elektrochemicznej oraz spadku potencjału w warstwach przyelektrodowych. Wyżej wymienione rozkłady są trudne do dokładnego wyznaczenia i właśnie ten fakt uniemożliwia jednoznaczne określenie grubości szczeliny międzyelektrodowej. Grubość szczeliny jest podstawowym wskaźnikiem technologicznym procesu determinującym jego wydajność, dokładność obróbki i jakość powierzchni obrabianej. Grubość szczeliny w najprostszym przypadku obróbki płaskiej powierzchni płaską elektrodą (rys.2), można wstępnie obliczyć z zależności (1).



Rys.2. Schemat drążenia płaską elektrodą. 1 - elektroda, 2 - przedmiot przed obróbką, 3 - przedmiot w trakcie obróbki, $s(x)$, $C_H(x)$, $T(x)$, $a(x)$ - rozkłady, odpowiednio: grubości szczeliny międzyelektrodowej, koncentracji objętościowej wodoru, temperatury i grubości usuwanego naddatku.

$$s = \frac{\eta k_v \kappa (U - \Delta U_p)}{v_f}; \quad j = \frac{v_f}{\eta k_v} \quad (1)$$

$$s = \frac{\kappa (U - \Delta U_p)}{j}$$

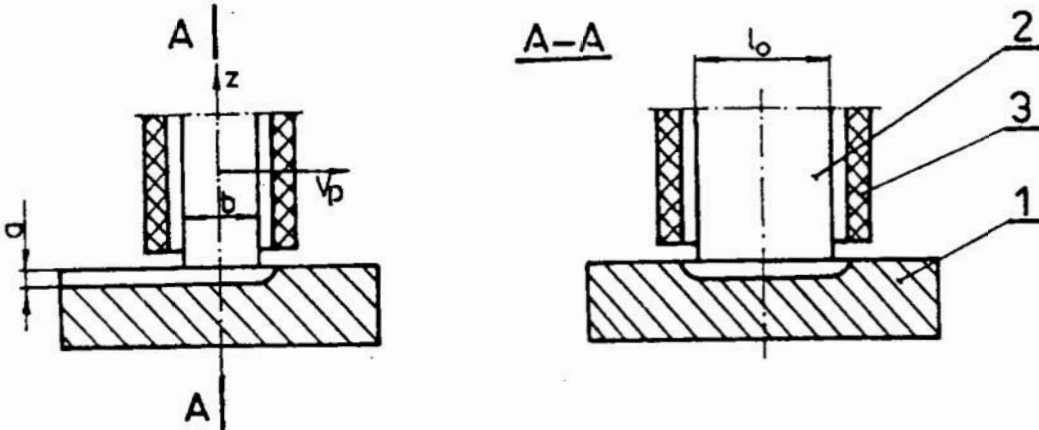
gdzie:

s - grubość szczeliny międzyelektrodowej, η - wydajność prądowa procesu roztwarzania, k_v - równoważnik elektrochemiczny materiału obrabianego, ηk_v - współczynnik obrabialności elektrochemicznej, U - napięcie międzyelektrodowe, ΔU_p - spadek potencjału w warstwach przyelektrodowych, v_f - prędkość przesuwu elektrody, j - gęstość prądu.

Z zależności (1) wynika, że grubość szczeliny zależy przede wszystkim od gęstości prądu, napięcia międzyelektrodowego oraz jego spadku w warstwach przyelektrodowych. Z kolei gęstość prądu zależy głównie od prędkości przesuwu elektrody (wydajności procesu roztwarzania), współczynnika obrabialności elektrochemicznej oraz od wielkości powierzchni

obrabaniej. Z uwagi na to, że nawet w przypadku obróbki płaską elektrodą nie ma możliwości oszacowania z wystarczającą dokładnością: κ , ΔU_p oraz ηk_v zależność (1) nie może stanowić podstawy do deterministycznego sterowania grubością szczeliny. Z tego względu, w praktyce steruje się prędkością przesuwu elektrody. Jest to sterowanie w układzie otwartym, nie uwzględniające sprzężenia zwrotnego pomiędzy stanem fizycznym szczeliny międzyelektrodowej a wartością v_f .

Aby zmniejszyć nieokreśloność procesu wynikającą z nierównomierności rozkładów η k_v , ΔU_p oraz κ stosuje się obróbkę nieprofilowaną elektrodą (rys.3), przemieszczającą się ponad powierzchnią obrabianą wzdłuż odpowiednio zaprojektowanego toru.



Rys.3. Schemat obróbki powierzchni płaskiej elektrodą prostopadłościenną przemieszczającą się ponad powierzchnią obrabianą; 1 - przedmiot obrabiany, 2 - elektroda, 3 - dysza doprowadzająca elektrolit.

W najprostszym przypadku elektrody płaskiej (rys.3) wydajność obróbki, grubość nadatku oraz gęstość prądu można oszacować z zależności:

$$V_w = l_0 \cdot a \cdot v_p \quad (2)$$

gdzie:

V_w - wydajność obróbki, l_0 - szerokość elektrody roboczej, a - grubość nadatku usuwanego w jednym przejściu elektrody, v_p - prędkość przesuwu elektrody.

Na podstawie praw Ohma i Faradaya grubość nadatku oraz średnią gęstość prądu można wyznaczyć z zależności:

$$a = \sqrt{2 \eta k_v (U - \Delta U_p) \frac{b}{v_p} + s_0^2} - s_0$$

$$j_{sr} = \frac{2 \kappa (U - \Delta U_p)}{a + 2s_0} \quad (3)$$

$$a = \frac{2 \kappa (U - \Delta U_p)}{j_{sr}} - 2s_0$$

$$V_w = l_0 v_p \left(\frac{2 \kappa (U - \Delta U_p)}{j_{sr}} - s_0 \right)$$

gdzie:

s_0 - grubość początkowa szczeliny międzyelektrodowej, b - długość elektrody roboczej, j_{sr} - średnia gęstość prądu w obszarze obróbki.

Z uwagi na to, że powierzchnia elektrody jest wielokrotnie mniejsza od powierzchni obrabianej można przyjąć, że κ , $\eta\kappa_v$ oraz ΔU_p są stałe w obszarze obróbki, co umożliwia w przypadku płaskiej elektrody oszacowanie grubości szczeliny elektrody z dokładnością większą niż przy drążeniu klasycznym. Jednak w praktyce, gdy obróbka powierzchni kształtowych jest realizowana w wielu przejściach i zwykle elektrodą o bardziej złożonym kształcie a grubość naddatku do usunięcia jest zmienna wzdłuż toru elektrody trudno jest oszacować a priori warunki roztwarzania wynikające z wartości s_0 , ΔU_p oraz $\eta\kappa_v$. Z tego względu również w tym przypadku zależności (2) i (3) nie mogą być podstawą do deterministycznego sterowania procesem (grubością usuwanego naddatku). W istniejących obecnie rozwiązaniach technicznych procesem obróbki ECM steruje się pośrednio przez sterowanie prędkością przesuwu elektrody wzdłuż odpowiednio zaprojektowanej trajektorii. Przy czym podobnie jak w przypadku drążenia, sterowanie jest realizowane w układzie otwartym ponieważ brak jest sprzężenia pomiędzy wartością v_p a stanem fizycznym szczeliny. Ponadto obróbka nieprofilowaną elektrodą odbywa się tylko w obszarze jej oddziaływania na powierzchnię obrabianą. Tym samym wydajność obróbki jest znacznie mniejsza niż w przypadku drążenia elektrochemicznego. Fakt ten wymusza warunek sterowania procesem przy kryterium:

$$V_w \rightarrow \max \quad (4)$$

Z dotychczasowych rozważań wynika, że sterowanie procesem obróbki elektrochemicznej w oparciu o algorytm deterministyczny jest, ze względu na znaczny stopień nieokreśloności procesu nieefektywne. Dlatego też podjęto prace zmierzające do sprawdzenia możliwości sterowania procesem ECM z wykorzystaniem zasad logiki rozmytej. Przedstawione powyżej modele deterministyczne wykorzystano do sformułowania ogólnych reguł decyzyjnych.

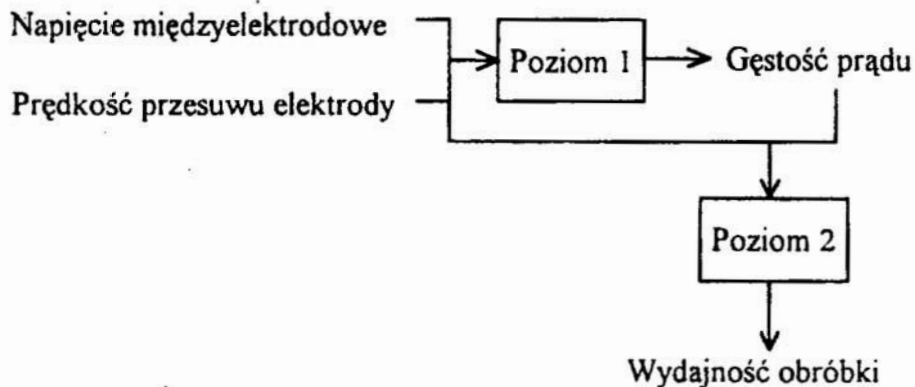
3. Przykład sterowania rozmytego wydajnością procesu obróbki ECM nieprofilowaną elektrodą.

Jak już wykazano (por. pkt 2), wydajność obróbki ECM zależy głównie od napięcia międzyelektrodowego, gęstości prądu i prędkości przesuwu elektrody. Gęstość prądu zależy przede wszystkim od napięcia i prędkości przesuwu elektrody.

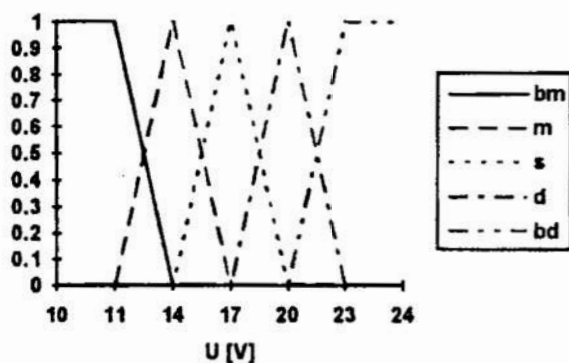
Po analizie przykładów zastosowania logiki rozmytej w sterowaniu procesami i układami technicznymi [1,3,4] do sterowania procesem obróbki elektrochemicznej nieprofilowaną elektrodą zaproponowano dwustopniowy sterownik rozmyty (rys.4). Sterownik zaprogramowano używając systemu Mathcad. Wielkościami wejściowymi sterownika są: napięcie międzyelektrodowe "U" oraz prędkość przesuwu elektrody " v_f ". Wielkościami wyjściowymi są: gęstość prądu "j" i wydajność obróbki " V_w " (rys. 1).

Do opisu wielkości wejściowych i wyjściowych sterownika zastosowano pięć klas zbiorów: "bardzo mały" (*bm*), "mały" (*m*), "średni" (*s*), "duży" (*d*), "bardzo duży" (*bd*).

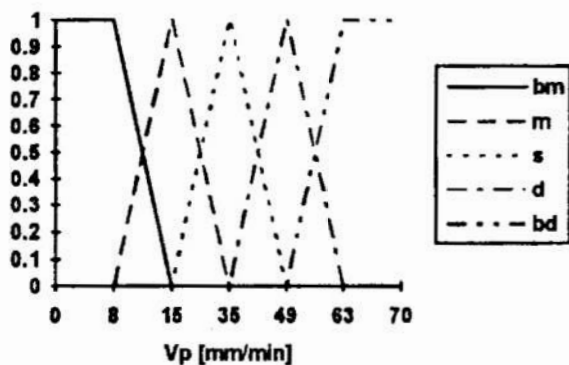
Przynależność poszczególnych elementów do zbiorów rozmytych opisano za pomocą funkcji liniowych (rys.5 - rys.8).



Rys.4: Schemat sterownika rozmytego do sterowania wydajnością obróbki ECM nieprofilowaną elektrodą.

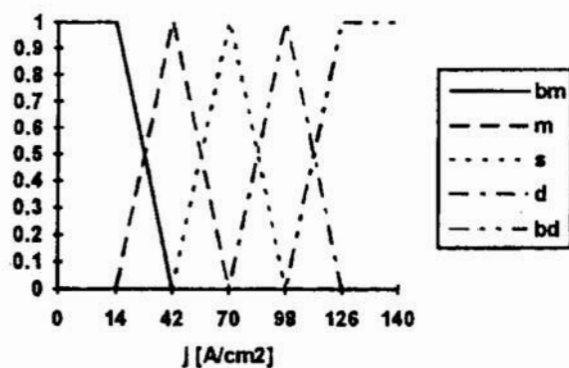


Rys.5: Funkcje przynależności dla wielkości "Napięcie międzyelektrodowe U".

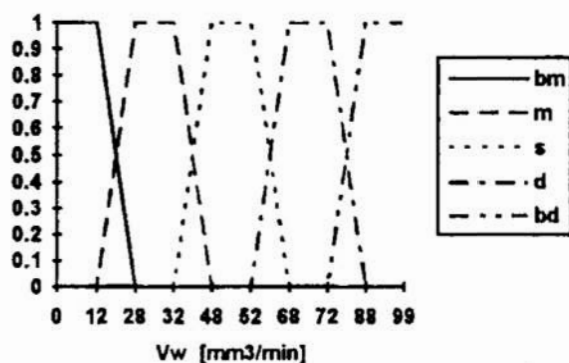


Rys.6: Funkcje przynależności dla wielkości "Prędkość przesuwu elektrody".

W sterowniku zastosowano 50 reguł postaci: "jeżeli napięcie międzyelektrodowe jest małe i prędkość przesuwu elektrody jest duża to gęstość prądu jest średnia". Kompletny zestaw reguł podano w tabl.1 i tabl.2.



Rys.7: Funkcje przynależności dla wielkości "Gęstość prądu".



Rys.5: Funkcje przynależności dla wielkości "Wydajność obróbki".

Tablica 1: Reguły opisujące zależność gęstości prądu od napięcia międzyelektrodowego i prędkości przesuwu elektrody.

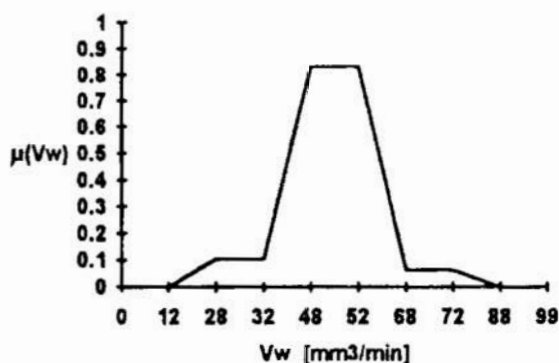
		Napięcie międzyelektrodowe				
		bm	m	s	d	bd
Prędkość przesuwu elektrody	bm	bm	m	s	s	s
	m	m	s	s	d	d
	s	m	s	d	d	bd
	d	m	s	d	bd	bd
	bd	m	s	d	bd	bd

Tablica 2: Reguły opisujące zależność wydajności obróbki od gęstości prądu i prędkości przesuwu elektrody.

		Gęstość prądu				
		bm	m	s	d	bd
Prędkość przesuwu elektrody	bm	bm	m	m	s	s
	m	bm	m	m	s	bd
	s	bm	m	m	d	bd
	d	bm	m	s	s	bd
	bd	bm	bm	m	s	d

Dla konkretnych (nierozmytych) wielkości wejściowych sterownik oblicza ich stopień przynależności do rozłącznych zbiorów rozmytych, a następnie odpowiednio do ustalonych wcześniej stopni przynależności oraz podanych reguł modyfikuje zbiory rozmyte (funkcje przynależności) wielkości wyjściowych.

Na rys.9 pokazano zmodyfikowaną funkcję przynależności elementów "Wydajność obróbki" dla wielkości wejściowych: napięcie międzyelektrodowe $U=15$ V, prędkość przesuwu elektrody $V=50$ mm/min.



Rys.6: Odpowiedź sterownika.

Ponieważ odpowiedź sterownika w postaci rozmytej jest nieprzydatna do celów praktycznych należy obliczyć ściśle określoną wartość będącą ostateczną odpowiedzią sterownika. Dokonano tego obliczając środek masy figury zamkniętej krzywą przynależności i osią odciętych korzystając ze wzorów:

$$m = \int_0^{99} \int_0^{\mu(V_w)} d\mu dV_w \quad (5)$$

$$V_{ster} = \frac{1}{m} \int_0^{99} \int_0^{\mu(V_w)} w d\mu dV_w$$

W przykładzie pokazanym na rys.9 odpowiedzią sterownika (wydajnością obróbki) jest wielkość $V_{ster}=49$ mm³/min.

Literatura.

1. Breiden O.: Praktische Entwicklung eines Fuzzy-Logik-Systems. Sonderdruck aus "elektronik industrie" Nr 7/92.
2. Breiden O.: Einführung in Theorie und Funktion der Fuzzy-Logik. Elektronik Entwicklung Nr 7-8/1992
3. Du R.X., Elbestawi M.A., Li S.: Tool Condition Monitoring in Turning Using Fuzzy Set Theory. Int.J.Mach.Tools Manufact, Vol.32, No.6, pp.781-796, 1992.
4. Elektroerosion: la logique floue est la, faisons donc le point. MACH'PRO - 24 fevrier 1993
5. Kacprzyk J: Zbiory rozmyte w analizie systemowej. PWN W-wa 1986.
6. Kozak J., Davidov A.D.: Vysokoskorostnoe elektrochimičeskoe formoobrazovanie. Izd. Nauka, Moskva 1990.
7. Ruszaj A.: Procesy obróbki elektrochemicznej i elektroerozyjnej w różnych odmianach kinematycznych. Zeszyty Naukowe IOS Nr 76 1989.

IDENTYFIKACJA BADANEGO PROCESU LINIOWYM RÓWNANIEM REGRESJI W OPARCIU O RACHUNEK MACIERZOWY

Streszczenie

Do identyfikacji procesu skrawania lub innych pokrewnych procesów stosuje się uniwersalny program regresji liniowej oparty na rachunku macierzowym. Program działa w trybie interakcyjnym. Funkcja regresji może być zadeklarowana jako wykładnicza, logarytmiczna albo wielomianowa. Liczbę zmiennych wejściowych można dobrać z przedziału <1,5>, liczba doświadczeń może wynosić maksimum 220. Hiperpłaszczyzna regresji opisana równaniem macierzowym $(X^T X)^{-1} X^T Y = b$ stanowi rozwiązanie układu równań normalnych $(X^T X) b = X^T Y$, którego stopień jest uzależniony od liczby zmiennych wejściowych. Istotność uzyskanej funkcji regresji jest oceniana za pomocą współczynnika korelacji wielowymiarowej R i testu F Snedecora. Istotność kolejnych współczynników funkcji regresji szacuje się za pomocą testu t Studenta. Poprawność działania programu sprawdzono porównując opracowania statystyczne wyników badań skrawaniowych i obróbki ściernej uzyskanych przez innych badaczy tych procesów stosujących inne metody obliczeniowe.

1. Wstęp

Większość zależności opisujących procesy w technologii maszyn posiada postać funkcji wykładniczych, logarytmicznych albo wielomianowych. Zależności te, na przykład siła skrawania, prędkość okresowa, itp, wyrażają się w funkcji zmiennych wejściowych, którymi są parametry obróbki. Postać wykładniczą funkcji przedstawiono w zależności (1)

$$Y = \exp(b_0) X_1^{b_1} X_2^{b_2} \dots X_M^{b_M} \quad (1)$$

postać logarytmiczną funkcji w zależności (2)

$$\ln(Y) = b_0 + b_1 X_1 + b_2 X_2 + \dots + b_M X_M \quad (2)$$

postać liniową funkcji w zależności (3)

$$Y = b_0 + b_1 X_1 + b_2 X_2 + \dots + b_M X_M \quad (3)$$

Ponieważ postać wykładniczą funkcji (1) można sprowadzić do postaci liniowej przez logarytmowanie, można przyjąć, że zależności opisujące proces obróbki jako funkcje parametrów obróbki opisują funkcje liniowe o postaci (3). Postać logarytmiczną (2) stosuje się do opisu procesu obróbki, jeśli występuje znaczna różnica ilościowa pomiędzy zmiennymi wejściowymi X_j ($j=1, \dots, M$). W tym przypadku transformuje

się zmienne wejściowe X_j do postaci kodowej [2]. Taka transformacja pozwala zwiększyć dokładność obliczenia stałych w zależności (2) na komputerze PC. Postać wielomianowa zależności (3) występuje najczęściej podczas skalowania przyrządów pomiarowych. Zmienne X_1, X_2, \dots, X_M nazywamy zmiennymi wejściowymi, zmienną Y nazywamy zmienną wyjściową. Wartości liczbowe stałych ustala się w wyniku badań doświadczalnych realizowanych wg dobranego planu badań [2]. Do obliczenia stałych wykorzystuje się zasadę najmniejszej sumy kwadratów względem hiperpłaszczyzny (3).

2. Zapis wyników badań

2.1 Zapis wyników badań według symboliki sumacyjnej

Badania wykonywane zgodnie z zasadami na przykład planowanego eksperymentu [2] umożliwiają ustalić stałe w funkcji regresji (3), która identyfikuje badany proces. Stałe w funkcji regresji (3) traktuje się jako niewiadome, natomiast zmienne wejściowe X_j ($j=1, \dots, M$) oraz zmienna wyjściowa Y stanowią zbiór znanych wielkości w próbie, ustalonych bądź mierzonych. Wzór (3) użyty do zapisu wyników jednego pomiaru ma więc postać:

$$Y_i = b_0 + b_1 X_{i1} + b_2 X_{i2} + \dots + b_j X_{ij} + \dots + b_M X_{iM} \quad (4)$$

gdzie indeksy oznaczają

- "i" - oznacza numer kolejnego doświadczenia, ($i=1, \dots, N$),
- "j" - oznacza kolejną zmienną niezależną, ($j=1, \dots, M$),

Kolejne doświadczenia zapisuje się wg (4) zmieniając numer indeksu "i" od $1 \div N$, gdzie N jest liczbą badań doświadczalnych.

2.2 Zapis wyników badań według symboliki macierzowej

Wyniki pomiarów w postaci rzeczywistej lub transformowanych do postaci kodowej zapisuje się do dwóch tablic. W pierwszej tablicy wektorowej oznaczonej Y zapisuje się ciąg mierzonych wyników (zmienna wyjściowa). W drugiej tablicy prostokątnej oznaczonej X o wymiarze $(M \times N)$ zapisuje się w pierwszej kolumnie wektor jedynek stanowiący tzw kolumnę elementów fikcyjnych, w pozostałych kolumnach macierzy X zapisuje się zmienne wejściowe. Kolumna elementów fikcyjnych jest niezbędna do zrealizowania zapisu macierzowego wyników badań w postaci zapisu (4) przy pomocy reguł rachunku macierzowego [1].

Zapis wyników badań wg symboliki macierzowej ma postać:

$$Y = X \cdot b \quad (5)$$

W zapisie (5) zawarte są następujące dane

$$Y = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_M \\ \vdots \\ Y_{NX} \end{pmatrix}; \quad X = \begin{pmatrix} 1. & X_{1,1} & X_{1,2} & \dots & X_{1,M} & \dots & X_{1,MX} \\ 1. & X_{2,1} & X_{2,2} & \dots & X_{2,M} & \dots & X_{2,MX} \\ \vdots & \vdots & \vdots & & \vdots & & \vdots \\ 1. & X_{M,1} & X_{M,2} & \dots & X_{M,M} & \dots & X_{M,MX} \\ \vdots & \vdots & \vdots & & \vdots & & \vdots \\ 1. & X_{NX,1} & X_{NX,2} & \dots & X_{NX,M} & \dots & X_{NX,MX} \end{pmatrix}; \quad b = \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_M \\ \vdots \\ b_{NX} \end{pmatrix}$$

... .. (6)

gdzie:

- NX - liczba wierszy w macierzy prostokątnej X;
- MX - liczba kolumn w macierzy prostokątnej X;
- N - liczba pomiarów doświadczalnych;
- M - liczba zmiennych wejściowych;
- Y_i - wartości zmiennej wyjściowej w kolejnych eksperymentach;
- X_{ij} - wartości zmiennych wejściowych ustalonych w eksperymencie, indeks "i" oznacza kolejny numer eksperymentu, indeks "j" oznacza kolejny numer zmiennej wejściowej,
- 1. - są to elementy fikcyjne w pierwszej kolumnie macierzy X niezbędne do zrealizowania zapisu (4), podczas mnożenia macierzowego.

3. Estymacja według procedury najmniejszych kwadratów

Równanie liniowe wielowymiarowej funkcji regresji ma postać:

$$\hat{Y} = b_0 + b_1 X_1 + b_2 X_2 + \dots + b_M X_M \quad (7)$$

gdzie wektor

$$b^T = [b_0, b_1, b_2, \dots, b_M] \quad (8)$$

jest estymatorem wektora $\beta^T = [\beta_1, \beta_2, \beta_3, \dots, \beta_M]$ rzeczywistych parametrów. Wyjście obiektu Y określa związek

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_M X_M + \varepsilon \quad (9)$$

Przyrost ε oznacza rozrzut wartości poza funkcją regresji \hat{Y} . Stałe parametry wektora (8) ustala się wg procedury najmniejszych kwadratów. Po zróżniczkowaniu uzyskuje się tzw układ równań normalnych. Stopień macierzy układu równań związany jest z ilością zmiennych wejściowych. Dla jednej

zmiennej wejściowej otrzymuje się 2-gi stopień macierzy układu równań normalnych, dla "M" zmiennych wejściowych stopień macierzy "M+1" układu równań normalnych. Zapis układu równań normalnych wg symboliki sumacyjnej dla jednej zmiennej niezależnej ma postać

$$\begin{aligned}
 b_1 \sum_{i=1}^N X_{i1}^2 + b_0 \sum_{i=1}^N X_{i1} &= \sum_{i=1}^N X_{i1} Y_i \\
 b_1 \sum_{i=1}^N X_{i1} + b_0 \sum_{i=1}^N 1 &= \sum_{i=1}^N Y_i
 \end{aligned}
 \tag{10}$$

Jest to układ równań liniowych z macierzą kwadratową stopnia drugiego. Według symboliki macierzowej układ posiada zapis

$$\mathbf{X}^T \mathbf{X} \mathbf{b} = \mathbf{X}^T \mathbf{Y}
 \tag{11}$$

Układ równań normalnych dla trzech zmiennych wejściowych wg symboliki sumacyjnej ma zapis

$$\begin{aligned}
 b_1 \sum_{i=1}^N X_{i1}^2 + b_2 \sum_{i=1}^N X_{i1} X_{i2} + b_3 \sum_{i=1}^N X_{i1} X_{i3} + b_0 \sum_{i=1}^N X_{i1} &= \sum_{i=1}^N X_{i1} Y_i \\
 b_1 \sum_{i=1}^N X_{i1} X_{i2} + b_2 \sum_{i=1}^N X_{i2}^2 + b_3 \sum_{i=1}^N X_{i2} X_{i3} + b_0 \sum_{i=1}^N X_{i2} &= \sum_{i=1}^N X_{i2} Y_i \\
 b_1 \sum_{i=1}^N X_{i1} X_{i3} + b_2 \sum_{i=1}^N X_{i2} X_{i3} + b_3 \sum_{i=1}^N X_{i3}^2 + b_0 \sum_{i=1}^N X_{i3} &= \sum_{i=1}^N X_{i3} Y_i \\
 b_1 \sum_{i=1}^N X_{i1} + b_2 \sum_{i=1}^N X_{i2} + b_3 \sum_{i=1}^N X_{i3} + b_0 \sum_{i=1}^N 1 &= \sum_{i=1}^N Y_i
 \end{aligned}
 \tag{12}$$

Jest to układ równań liniowych z macierzą kwadratową stopnia czwartego. Według symboliki macierzowej układ ma zapis

$$\mathbf{X}^T \mathbf{X} \mathbf{b} = \mathbf{X}^T \mathbf{Y}
 \tag{13}$$

Zatem zapisy macierzowe układu równań normalnych (11), (13) są formalnie takie same i różnią się tylko stopniem macierzy kwadratowej. Parametry wektora \mathbf{b} (8) równania estymacji (7) otrzymuje się z (11) przez transformację macierzową mnożenia lewostronnego [2]:

$$\mathbf{b} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}
 \tag{14}$$

W zapisie sumacyjnym wektor parametrów \mathbf{b} (8) uzyskuje się

przez rozwiązywanie układu równań normalnych, które różnią się jakościowo na przykład (10) lub (12) w zależności od ilości zmiennych wejściowych M .

4. Transformacje macierzowe stosowane przy obliczaniu wektora b równania estymacji Y

Przy używaniu zapisu macierzowego do obliczenia wektora b (14) niezbędne są podprogramy do transponowania macierzy (X^T), mnożenia macierzy ($X^T X$) i odwracania macierzy $(X^T X)^{-1}$. Są to znane transformacje realizowane przez podprogramy dostępne w bibliotekach SSP (Scientific Subroutine Package) [5] w języku Fortran w komputerach PC. Odwracanie macierzy układu równań normalnych stwarza jednak problem w związku ze zmieniającym się stopniem macierzy ($X^T X$) uzależnionym od ilości zmiennych wejściowych w równaniu (4). Podprogram odwracający macierz oblicza wartość wyznacznika odwracanej macierzy $(X^T X)^{-1}$. W przypadku gdy wyznacznik ma wartość zerową przerywa się tok obliczeń, gdyż macierz jest osobliwa (dzielenie przez zero). Praktycznie oznacza to występowanie w macierzy kwadratowej dwóch identycznych wierszy lub kolumn, względnie wiersza lub kolumny z zerami. Ostatni wariant występuje w omawianym przypadku. Wynika to z następujących założeń. Do zapisu wyników badań przyjęto macierz prostokątną (X) o wymiarach $(N \times M)$. Z macierzy tej powstaje macierz kwadratowa ($X^T X$) o wymiarze (M, M) , której podmacierze są również kwadratowe o wymiarze (M, M) , gdzie stopień M odpowiada ilości zmiennych wejściowych. Niewykorzystane wiersze macierzy ($X^T X$) o wymiarze (M, M) są zerowe. Odwrócenie takiej macierzy daje wyznacznik zerowy tj. macierz osobliwą. Tę trudność można usunąć w sposób następujący. W zależności od stopnia M układu równań normalnych tworzy się (deklaruje) macierze wektorowe $(M \times M)$ elementowe (dla $M \in \langle 2, 5 \rangle$) nadając im odpowiednie nazwy. Dwuwymiarową macierz kwadratową (M, M) elementową przetwarza się na macierz wektorową $(M \times M)$ elementową uprzednio zdefiniowaną. Tę macierz odwraca się przez podprogram MINV z biblioteki SSP. Po odwróceniu macierz wektorową $(M \times M)$ elementową transformuje się ponownie na dwuwymiarową macierz kwadratową stopnia (M) gdyż stopień M jest niezbędny do dalszych przekształceń macierzowych przy obliczaniu wektora parametrów b . Jeśli przy odwracaniu macierzy wektorowej stopnia M wyznacznik ma wartość zerową, oznacza to, że macierz $(X^T X)$ jest osobliwa. W tym przypadku w sposób wymuszony przerywa się tok dalszych obliczeń i równocześnie wprowadza się na ekran i na zbiór (wcześniej deklarowany) komunikat o osobliwości macierzy układu równań normalnych i o przerwaniu programowym toku obliczeń.

5. Macierz współczynników korelacji

Dla zmiennych wejściowych X_j ($j=1, \dots, m$) oraz zmiennej wyjściowej Y oblicza się macierz współczynników korelacji

$$R_{X_1, X_2, \dots, X_5} = \begin{array}{c|c} \begin{array}{c} R_{X_1, X_1}, R_{X_1, X_2}, \dots, R_{X_1, X_5} \\ R_{X_2, X_1}, R_{X_2, X_2}, \dots, R_{X_2, X_5} \\ \dots \\ R_{X_5, X_1}, R_{X_5, X_2}, \dots, R_{X_5, X_5} \end{array} & \begin{array}{c} R_{X_1, Y} \\ R_{X_2, Y} \\ \dots \\ R_{X_5, Y} \end{array} \\ \hline R_{Y, X_1}, R_{Y, X_2}, \dots, R_{Y, X_5} & R_{Y, Y} \end{array} \dots (15)$$

gdzie

$$R_{X_j, X_j} = \frac{\text{cov}(X_j, X_j)}{\sqrt{\text{var}(X_j)} \sqrt{\text{var}(X_j)}} = 1$$

$$R_{X_j, X_k} = \frac{\text{cov}(X_j, X_k)}{\sqrt{\text{var}(X_j)} \sqrt{\text{var}(X_k)}} = \frac{E[X_j - E(X_j)][X_k - E(X_k)]}{\sqrt{\text{var}(X_j)} \sqrt{\text{var}(X_k)}}$$

$$R_{Y, X_j} = \frac{\text{cov}(Y, X_j)}{\sqrt{\text{var}(Y)} \sqrt{\text{var}(X_j)}} = \frac{E[Y - E(Y)][X_j - E(X_j)]}{\sqrt{\text{var}(Y)} \sqrt{\text{var}(X_j)}}$$

$$R_{Y, Y} = \frac{\text{cov}(Y, Y)}{\sqrt{\text{var}(Y)} \sqrt{\text{var}(Y)}} = 1$$

Elementy macierzy korelacji spełniają warunek

$$R_{X_j, X_k} \in \langle -1, +1 \rangle$$

Udowodniono, że jeśli zmienne losowe są niezależne to ich kowariancja jest równa zero [1], a zatem i odpowiedni współczynnik korelacji jest równy zero. Wynika z tego, że analiza macierzy korelacji pozwala zatem wnioskować o zależności między zmienną wynikową Y a zmiennymi wejściowymi X_j ($j=1, \dots, m$), jak również o zależności między samymi zmiennymi wejściowymi X_j , co daje odpowiedź na pytanie odnośnie celowości tworzenia modeli wyższego rzędu niż

model liniowy. Wartość $R_{x_j, x_k} = 0$ oznacza, że zmienne są niezależne (nieskorelowane). Z wartości $R_{x_j, x_k} = 1$ wynika, że mamy do czynienia z zależnością funkcyjną między tymi zmiennymi.

6. Obliczenie współczynnika korelacji wielowymiarowej R

W praktycznych rozważaniach ważną sprawą jest siła związku pomiędzy równaniem estymacji (7)) a równaniem regresji (9). W celu oszacowania siły związku oblicza się współczynnik korelacji wielowymiarowej R, który jest wielkością unormowaną przyjmującą wartość $R \in \langle 0, 1 \rangle$. W zapisie macierzowym współczynnik R oblicza się wg wzoru [2]

$$R = \sqrt{\frac{\mathbf{b}^T \mathbf{X}^T \mathbf{Y} - N \bar{Y}^2}{\mathbf{Y}^T \mathbf{Y} - N \bar{Y}^2}} \quad (16)$$

gdzie

\mathbf{Y} - wektor zmiennej wyjściowej,

N - liczba pomiarów,

\bar{Y} - wartość średnia wektora zmiennej wyjściowej

$$\bar{Y} = \frac{\sum_{i=1}^N Y_i}{N} \quad (17)$$

Wartość współczynnika R jest miarą zależności pomiędzy zmienną wyjściową Y a zmiennymi wejściowymi $X_j (j=1, \dots, n)$. Wartość $R=0$ oznacza, że zmienne są niezależne (nieskorelowane). Z wartości $R=1$ wynika, że mamy do czynienia z zależnością funkcyjną.

7. Obliczenie testu F Snedecora

Do ustalenia poziomu istotności obliczonego wektora \mathbf{b} estymowanego równania regresji (7) można posłużyć się testem F Snedecora badającego stosunek oszacowania wariancji (σ^2_Y) estymowanego równania regresji Y , do oszacowania wariancji resztowej ($\sigma^2_{Y-\hat{Y}}$). Zapis macierzowy testu F Snedecora ma postać [2]

$$F = \frac{N - M - 1}{M} \frac{R^2}{1 - R^2} \quad (18)$$

gdzie

N - liczba doświadczeń,

M - liczba zmiennych wejściowych.

R - współczynnik korelacji wielowymiarowej,

Zmienna losowa F ma rozkład F Snedecora o $r_1 = (N-M-1)$ oraz $r_2 = M+1$ stopniach swobody.

8. Obliczenie testu t Studenta

Współczynniki równania estymacji b_j ($j=0,1,\dots,M$) można oszacować posługując się testem t Studenta. Współczynniki b_j są traktowane jako zmienne losowe, które mają wartość oczekiwaną b_j oraz odchylenie standardowe $\sigma_{b_j} = \sigma \sqrt{c_{j,j}}$. Stałe $c_{j,j}$ są to mnożniki Gaussa, obliczane wg macierzy kowariancyjnej, która ma postać $(X^T X)^{-1}$ [2]. Nieznane odchylenie standardowe σ szacuje się przy pomocy wariancji resztowej $\hat{\sigma}_{Y-\hat{Y}}^2$

$$\hat{\sigma}_{Y-\hat{Y}}^2 = \frac{1}{N-M-1} \sum_{i=1}^{i=N} (y_i - \hat{y})^2 = \frac{1}{N-M-1} (Y^T Y - b X^T Y) \quad (19)$$

gdzie

M - liczba zmiennych wejściowych,

N - liczba pomiarów.

Zmienna losowa t_j ($j=0,1,\dots,M$), opisana równaniem

$$t_j = \frac{|b_j|}{\hat{\sigma}_{Y-\hat{Y}} \sqrt{c_{j,j}}} \quad (20)$$

ma rozkład t Studenta o $r=(N-M-1)$ stopniach swobody.

9. Weryfikacja hipotez statystycznych

Funkcję regresji w całości weryfikuje się testem F Snedecora, zaś kolejne współczynniki funkcji regresji b_j ($j=1,\dots,M$) weryfikuje się testem t Studenta. O przydatności otrzymanego estymowanego równania regresji w całości lub niektórych jego współczynników decyduje prowadzący badania eksperymentalne.

10. Przykłady i możliwe zastosowania

W tabelicy 1 zamieszczono wyniki obliczeń dla danych

Tablica 1. Ustalenie równania trwałości ostrza jako funkcji szybkości, posuwu i głębokości skrawania [3]
 Establishing the relationship between the machining parameters (cutting speed, feedrate and depth of cut) and the tool life

IDENTYFIKACJA PROCESU ROWNANIEM REGRESJI WIELOKROTNEJ			
DANE DO TABLIC			
18 3			
X1	X2	X3	Y
40.000	.200	1.000	59.200
40.000	.350	1.000	52.950
75.000	.350	3.000	29.100
75.000	.500	3.000	23.200
MACIERZ WSPOLCZYNNIKOW KORELACJI (B(K),B(J))			
ZMIENNA 1			
1.00000	.00000	.00000	-.38873
ZMIENNA 2			
.00000	1.00000	.00000	-.34671
ZMIENNA 3			
.00000	.00000	1.00000	-.78645
ZMIENNA 4			
-.38873	-.34671	-.78645	1.00000
WSPOLCZYNNIKI FUNKCJI REGRESJI: B(1),B(2),B(3),B(4)			
4.8612	-.2839	-.2111	-.3980
STALA FUNKCJI REGRESJI W POSTACI WYKLADNICZEJ			
EXP(4.8612) =	129.1845		
ODCHYLENIE STANDARDOWE WSPOLCZYNNIKOW REGRESJI B(1)-B(4)			
.26844	.06480	.05400	.04490
WSPOLCZYNNIK KORELACJI WIELOWYMIAROWEJ R			
R =	.9433		
TEST F SNEDECORA			
F =	37.6895		
TEST T STUDENTA DLA WSPOLCZYNNIKOW FUNKCJI REGRESJI			
	4.38198		
	3.90826		
	8.86523		

uzyskanych podczas badań trwałości ostrza [4] ($X_1, X_2, X_3 \rightarrow v, p, g$, - parametry obróbki, $Y \rightarrow T$ -trwałość), zaś w tablicy 2 wyniki badań procesu obróbki ścierniej [3] ($X_1, X_2, X_3, X_4 \rightarrow$ ziarnistość, struktura, twardość, siła szlifowania, $Y \rightarrow$ wydajność obróbki). Opracowany program generuje funkcje wykładnicze o postaci (1) albo funkcje logarytmiczne o postaci (2) albo liniowe o postaci (3). Postać funkcji ustalana jest konwersacyjnie na początku realizacji programu.

Tablica 2. Ustalenie równania wydajności ściernicy jako funkcji ziarnistości, twardości, struktury i siły szlifowania
 Establishing the relationship between the grinding conditions (wheel grit size, wheel hardness, wheel structure grinding force) and the volume removal rate in grinding

IDENTYFIKACJA PROCESU ROWNANIEM REGRESJI WIELOKROTNEJ					
DANE DO TABLIC					
42 4					
X1	X2	X3	X4	Y	
.275	.520	.375	2.000	.460	
.275	.520	.375	3.000	.580	
.	
.116	.500	.345	3.000	.890	
.116	.500	.345	4.000	.480	
MACIERZ WSPOLCZYNNIKOW KORELACJI (B(K),B(J))					
ZMIENNA 1					
1.00000	-.00278	-.01109	-.13416	-.20566	
ZMIENNA 2					
-.00278	1.00000	-.00134	.08504	.32418	
ZMIENNA 3					
-.01109	-.00134	1.00000	.13250	.44017	
ZMIENNA 4					
-.13416	.08504	.13250	1.00000	.43463	
ZMIENNA 5					
-.20566	.32418	.44017	.43463	1.00000	
WSPOLCZYNNIKI FUNKCJI REGRESJI: B(1), B(2), B(3), B(4), B(5)					
6.6227	-.5096	6.6746	3.7395	.9206	
STALA FUNKCJI REGRESJI W POSTACI WYKLADNICZEJ					
EXP(6.6227) =			751.9958		
ODCHYLENIE STANDARDOWE WSPOLCZYNNIKOW REGRESJI B(1)-B(5)					
2.45845	.40456	2.76880	1.16940	.34170	
WSPOLCZYNNIK KORELACJI WIELOWYMIAROWEJ R					
R = .668990					
TEST F SNEDECORA					
F = 7.4935					
TEST T STUDENTA DLA WSPOLCZYNNIKOW FUNKCJI REGRESJI					
1.259636					
2.410657					
3.197794					
2.694069					

Znaki minus przy funkcji regresji o postaci wykładniczej oznaczają postać ilorazową funkcji (1). Ponieważ współczynniki funkcji regresji mają znaki ujemne oznacza to, że zmienne ($X_1, X_2, X_3 \Leftrightarrow v, p, g$) są odwrotnie proporcjonalne do zmiennej wyjściowej ($Y \Leftrightarrow T$). Opracowany program pozwala sprawdzać poprawność zależności podawanych w lite-

raturze. Na przykład [3] str. 110+112, funkcja iloczynowa dla danych tam przytoczonych powinna mieć następujące wartości stałych

$$T = \frac{129,184}{V^{0,2839} p^{0,2111} g^{0,3980}} \quad [\text{min}] \quad (21)$$

Funkcji tej odpowiada współczynnik korelacji $R = 0.94$ oraz test F Snedecora $F=27.69$.

11. Podsumowanie

Opracowany program jest wygodnym środkiem identyfikacji badanego procesu. Macierz korelacji pozwala ustalić siłę związku pomiędzy zmiennymi w procesie. Daje to podstawę do wnioskowania odnośnie celowości budowania modelu stopnia drugiego, w którym występują kwadraty i iloczyny mieszane zmiennych wejściowych X_j ($j=1, \dots, m$). Program można również wykorzystać do ustalenia funkcji skalowania przyrządów pomiarowych. Postać funkcji skalowania jako wykładniczej (1), logarytmicznej (2) lub wielomianowej (3) ustala się w trybie interakcyjnym.

Application of multiple linear regression and matrix calculus for modelling of machining processes

Summary

Modelling of the investigated process has been carried out as a hyperplane of regression which is described by matrix equation $\mathbf{b} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}$. This program is operating in an interactive mode. Number of input variable varies from 1 to 5. The maximum number of the investigated points is 220. Correlation matrix, coefficients of the multiple regression equations, t and F parameters of statistical tests are calculated by this program.

Literatura

1. Kryslcki W., Włodarski L.: Analiza matematyczna w zadaniach, Cz.I, Warszawa, PWN, 1974.
2. Mańczak K.: Technika planowania eksperymentu, Warszawa, WNT, 1976.
3. Marciniak M., Zawora J., Lang H.V., Pacanowski K: Rola struktury ściernic w szlifowaniu na sucho, Podstawy i technika obróbki ścierniej, Zbiór prac pod redakcją T.Karpińskiego, Wydawnictwo WSI Koszalin 1993.
4. Pajak E., Wleczorkowski K.: Podstawy optymalizacji operacji technologicznych, Warszawa - Poznań, PWN, 1982.
5. IBM Corporation : System/360 Scientific Subroutine Package (360 A-CM-03X) Version III, Technical Publications Department, 1133 Westchester Avenue, White Plains, New York 10604, 1969.

POLIOPTYMALIZACJA UŁATWIA FORMALNE UJĘCIE PROBLEMÓW PROJEKTOWYCH W SYSTEMACH CAD

Tezą pracy jest, że polioptymalizacja może być sposobem formalizacji zadań heurystycznych w projektowaniu, walnie przyczyniając się do rozwoju systemów CAD. W pracy sformułowano zadanie polioptymalizacji, w szczególności w projektowaniu technicznym. Wskazano na główne walory ujęcia polioptymalnego, tj.: (1) łatwa ocena wpływu ograniczeń, (2) łatwa ocena wpływu dyskretyzacji zmiennych decyzyjnych, (3) matematyczne formułowanie zasad projektowania, (4) jednoczesna optymalizacja struktury i parametrów.

1. Wstęp i teza pracy

Szeroko rozumiane projektowanie (i konstruowanie - tutaj są to synonimy) maszyn, obiektów przemysłowych i komunikacyjnych, planowanie procesów (np. wytwórczych), działań (np. militarnych) i organizacji (np. masowej obsługi) jest ważnym działaniem, ponieważ ma decydujący wpływ na życie gospodarcze. Istotną rolę odgrywają działania optymalizacyjne w projektowaniu. Skuteczność optymalizacji zależy nie tylko od **dokładności** działań optymalizacyjnych, lecz być może w głównej mierze od **sformułowania zadania** oraz umiejętnej **interpretacji** otrzymanych wyników, a to są działania niealgorytmiczne.

Drugą przesłanką niniejszej pracy jest, że jedną z głównych barier rozwoju komputerowych systemów wspomagania projektowania i konstruowania (CAD) jest brak algorytmów dla działań **twórczych**, które są heurystyczne [4]. W tej sytuacji komputer powinien przejąć rolę wydajnego pomocnika, opracowującego dla projektanta potrzebne informacje i prezentującego je w najdogodniejszej dla człowieka postaci.

W pracy starano się wykazać, że polioptymalizacja może być bardzo pomocna do budowy niektórych modułów systemów CAD, dostarczając narzędzi i technik zasypujących przepaść między metodami algorytmicznymi i heurystycznymi.

2. Podstawy teoretyczne polioptymalizacji

W projektowaniu typową jest sytuacja, kiedy należy uwzględnić wiele kryteriów, zwanych tu *kryteriami oceny*, niezależnie od *ograniczeń nierównościowych*, nałożonych na właściwości obiektu.

Kryteria powinny być ujmowane *probabilistycznie*, to znaczy stochastyczny powinien być opis preferencji, ponieważ zależą one od losowych warunków eksploatacji. Także sam optymalizowany obiekt powinien być opisany losowo, gdyż losowy jest proces jego wytwarzania i losowe są zakłócenia zewnętrzne (np. zużycie paliwa przez silnik zależy od przypadkowych parametrów stanu otoczenia). Zauważmy także, że niektóre kryteria mają charakter *rozmyty* (np. wygoda czy bezpieczeństwo) i taki też powinien być ich opis.

W dalszym ciągu ograniczymy się do dyskusji przypadków deterministycznych.

Wygodnie jest wyróżnić dwa przypadki optymalizacji (i polioptymalizacji): przypadek zmiennych decyzyjnych *dyskretnych* i przypadek zmiennych decyzyjnych *ciągłych*.

W pierwszym przypadku optymalizacja (lub polioptymalizacja) jest przeprowadzana na *skończonym* (czyli przeliczalnym) zbiorze wariantów o liczności zwykle nie przekraczającej kilkanaście do kilkudziesięciu. Przykładami mogą być: wybór koncepcji konstrukcji, wybór zespołu handlowego z katalogu lub wybór oferty (np. oferty licencyjnej).

Załóżmy, że w przypadku zmiennych dyskretnych dane są:

1) Zbiór wariantów (np. nazw) $A = \{a_v : v = 1, \dots, V\}$

Jest to zbiór skończony, indeks v jest zmienną decyzyjną,

2) Zbiór kryteriów oceny $K = \{k_i : i = 1, \dots, M\}$,

przy czym pewne kryteria mają osiągnąć wartości maksymalne: $k_i \rightarrow \max!$ ($i=1, \dots, n$), zaś pozostałe kryteria wartości minimalne: $k_j \rightarrow \min!$, ($j=n+1, \dots, M$),

3) Macierz ocen $[k_{iv}]$, $i = 1, \dots, M$, $v = 1, \dots, V$

Oceny k_{iv} mogą być liczbami, rozkładami prawdopodobieństwa (np. gęstościami $f(k_{iv})$) lub funkcjami przynależności $\mu_v(k_{iv})$ w przypadku oceny rozmytej, charakteryzują one osiągi obiektu,

4) Ograniczenia $g_j(k_{1v}, \dots, k_{Mv}) > 0 : j = 1, \dots, J$, gdzie k_{iv} są ocenami.

W przypadku drugim (zmiennych ciągłych) optymalizacja przeprowadzana jest na zbiorze ograniczonym nieprzeliczalnym. Jest to tzw. *optymalizacja parametryczna* konstrukcji lub procesu. Dane są:

1) Zbiór zmiennych decyzyjnych: $X = \{x_1, \dots, x_p\}$

2) Zbiór kryteriów oceny K (jak powyżej),

3) Ograniczenia (jak powyżej) w postaci:

$$g_j(x_1, \dots, x_p, k_1, \dots, k_M) > 0 : j = 1, \dots, J,$$

4) Model matematyczny projektowanego obiektu:

$$k_i = k_i(x_1, \dots, x_p) : i = 1, \dots, M.$$

Zadanie optymalizacji brzmi: **należy znaleźć taki wariant $a_0 \in A$ (zwany dalej optymalnym), który spełni możliwie dobrze kryteria a jednocześnie spełni bezwzględnie ograniczenia :**

$$a_0 = a_v \in A : \left\{ \bigcap_{i \in [1, M]} k_i(a_v) \rightarrow \min! \vee \max! \wedge \bigcap_{j \in [1, J]} g_j(k_{1v}, \dots, k_{Mv}) \geq 0 \right\}$$

W przypadku drugim wariant a_v jest określony jako wektor zmiennych decyzyjnych: $a_v = [x_1, \dots, x_p]$

Jeśli liczba kryteriów $M > 2$, powyższe zadanie jest zadaniem *polioptymalizacji*. Definicje rozwiązań (wariantów) polioptymalnych można oprzeć na pojęciu relacji porządku częściowego określonej przez stożki dominowania [1] lub na pojęciu dominacji wariantów w przestrzeni kryteriów w ujęciu teoriomnogościowym [2].

W ogólnym przypadku jest wiele wariantów polioptymalnych, zatem rozwiązanie zadania polioptymalizacji jest niejednoznaczne. Wymaga więc od użytkownika arbitralnej decyzji co do wyboru jednego konkretnego wariantu.

Problemy te rozważano w pracy [2], gdzie zaproponowano teorię procesu wyboru a także zbiór metod wspomagających wybór. W oparciu o tę pracę, dla przypadku dyskretnego, w ramach Programu RP.I.06 stworzono system do komputerowego wspomaganie procesu wyboru, pracujący w trybie dialogowym z projektantem.

3. Charakterystyka zadań polioptymalizacji w projektowaniu i konstruowaniu

Wygodnie jest wyróżnić trzy przypadki optymalizacji i polioptymalizacji [2, s.71]:

a) optymalizacja na nielicznym zbiorze wariantów, przy dyskretnych zmiennych (patrz p.4), np. zbiór koncepcji konstrukcji lub katalogowy zbiór podzespołów handlowych, jedna lub kilka dyskretnych zmiennych decyzyjnych, np. nazwa. Jest to przypadek typowy dla konstruowania wstępnego.

b) optymalizacja na licznym przeliczalnym zbiorze wariantów, przy wielu dyskretnych zmiennych decyzyjnych, np. rozstawienie maszyn w hali produkcyjnej lub sterowanie dyspozycyjne ruchem wagonów na stacji rozrządowej. Przypadek typowy dla projektowania tzw. procesów dyskretnych, np. procesów produkcyjnych, transportowych i innych na poziomie ich organizacji, zarządzania itp.

c) optymalizacja na nieprzeliczalnym ('ciągłym') zbiorze parametrycznym, przy ciągłych zmiennych decyzyjnych, np. optymalizacja zbiornika. Przypadek ten jest typowy dla konstruowania szczegółowego.

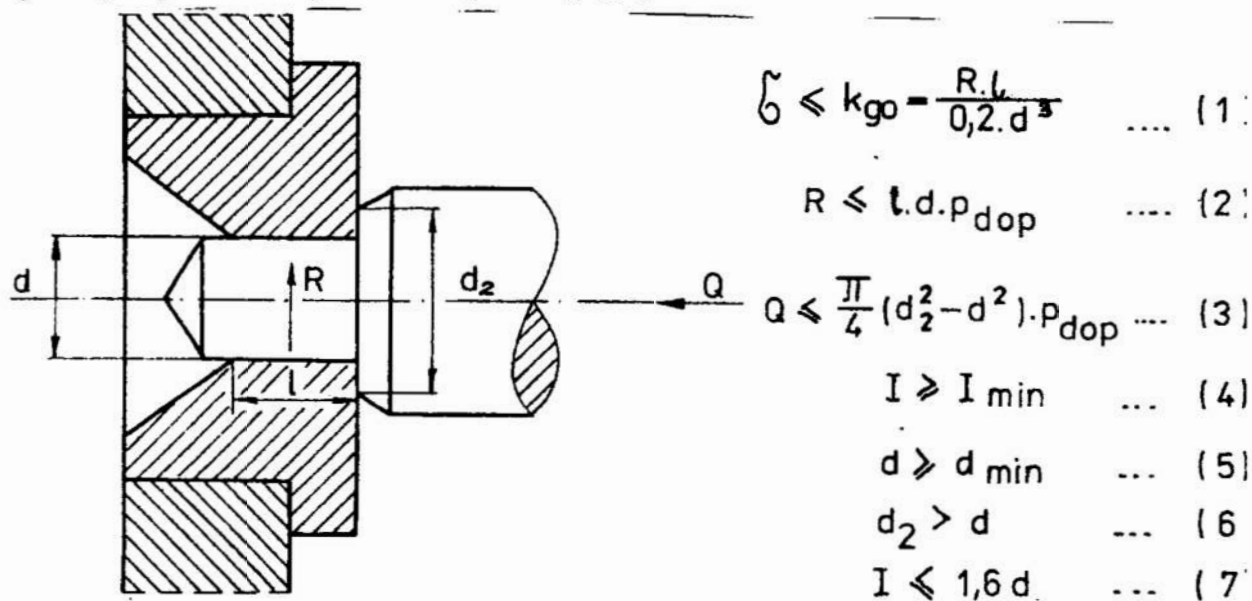
Na podstawie wielu przebadanych przypadków można pokusić się o podanie kilku osobliwości zadań polioptymalizacji w projektowaniu (i konstruowaniu).

Dla przypadku pierwszego typowym jest występowanie licznego zbioru kryteriów oceny, wśród nich wielu probabilistycznych (a także niektórych rozmytych). Polioptymalizacja może być pomocna do wyeliminowania wariantów zdominowanych, przy czym w przypadku kryteriów deterministycznych można to zrobić 'ręcznie'.

W przypadku drugim i trzecim zbiór kryteriów nie jest liczny (dwa do cztery, prawie nigdy powyżej ośmiu). Wiele spośród nich powinno być ujęte probabilistycznie, niektóre opisane w sposób rozmyty, np. właściwości ergonomiczne. Zbiór zmiennych decyzyjnych liczny, często kilkadziesiąt, przy czym są to z reguły zmienne deterministyczne i ostre (nie rozmyte).

4. Korzyści płynące z ujęcia polioptymalnego [3]

Optymalizacja umożliwia znalezienie jednego (o ile ono istnieje) wariantu, najlepszego ze względu na jedno kryterium. Natomiast polioptymalizacja, prócz tego że prowadzi do znalezienia zbioru wariantów niezdominowanych (polioptymalnych), umożliwia projektantowi zapoznanie się z kontekstem zadania polioptymalizacji, tzn. pokazuje, jaki wpływ na rozwiązanie optymalne mają ograniczenia, pozostałe kryteria, dyskretyzacja określonych zmiennych decyzyjnych itd.



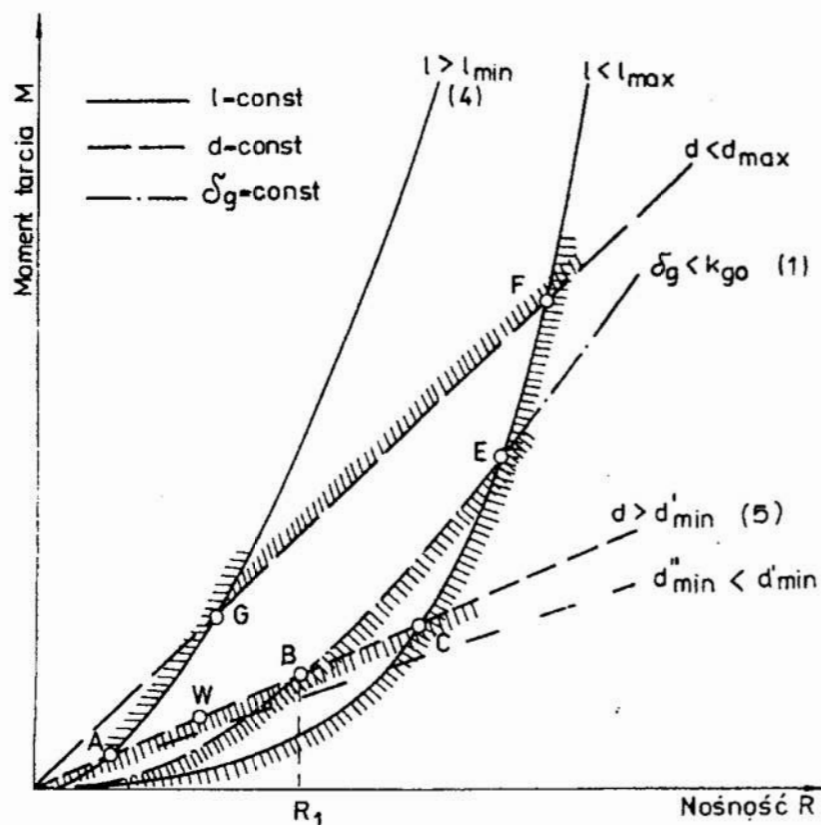
Rys. 1. Małogabarytowe łożysko ślizgowe jako przykład polioptymalizacji

Poniżej omówimy cztery charakterystyczne rodzaje korzyści jakie może przynieść ujęcie zadania optymalizacji w kategoriach polioptymalizacji. Zilustrowane zostaną przykładem ciernego małowabarytowego łożyska ślizgowego (rys.1), omówionym w pracy [5].

4.1. Analiza ograniczeń konstrukcyjnych

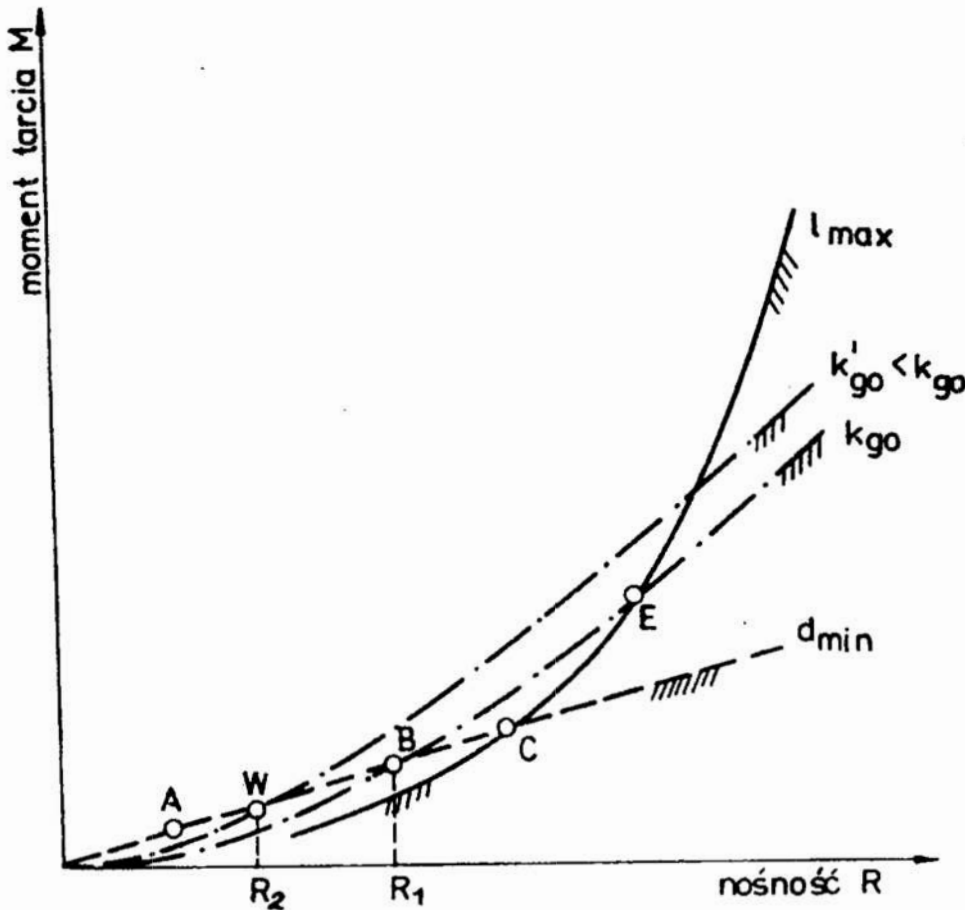
Trudno tego dowieść, ale doświadczenie uczy, że zbiór wariantów polioptymalnych leży najczęściej w całości na ograniczeniach (dwa przeciwne przypadki podano w [2]). Oznacza to, że odpowiednio zmieniając dane ograniczenie ('rozluźniając' je) można uzyskać nowy, lepszy zbiór.

Na przykład, takim ograniczeniem dla łożyska ślizgowego [5] może być minimalna średnica czopa d (rys.2): zmniejszając ją z wartości d' do wartości d'' otrzymujemy lepszy zbiór rozwiązań polioptymalnych P . Przy graficznym sposobie przedstawiania zbioru rozwiązań polioptymalnych widać więc explicite jaki jest techniczny i ekonomiczny skutek 'zacieśniania' ograniczeń (często zresztą nieuzasadnionego). Ten aspekt polioptymalizacji jest o tyle ważny, że często ograniczenia ustalane są arbitralnie, na podstawie intuicyjnego sądu projektanta, i dlatego mogą być zmienione.



Rys. 2. Przykład ilustrujący zasadę, że 'rozluźnianie' ograniczenia na którym leży zbiór wariantów polioptymalnych (np. $d > d'_{\min}$) powoduje poprawę tego zbioru

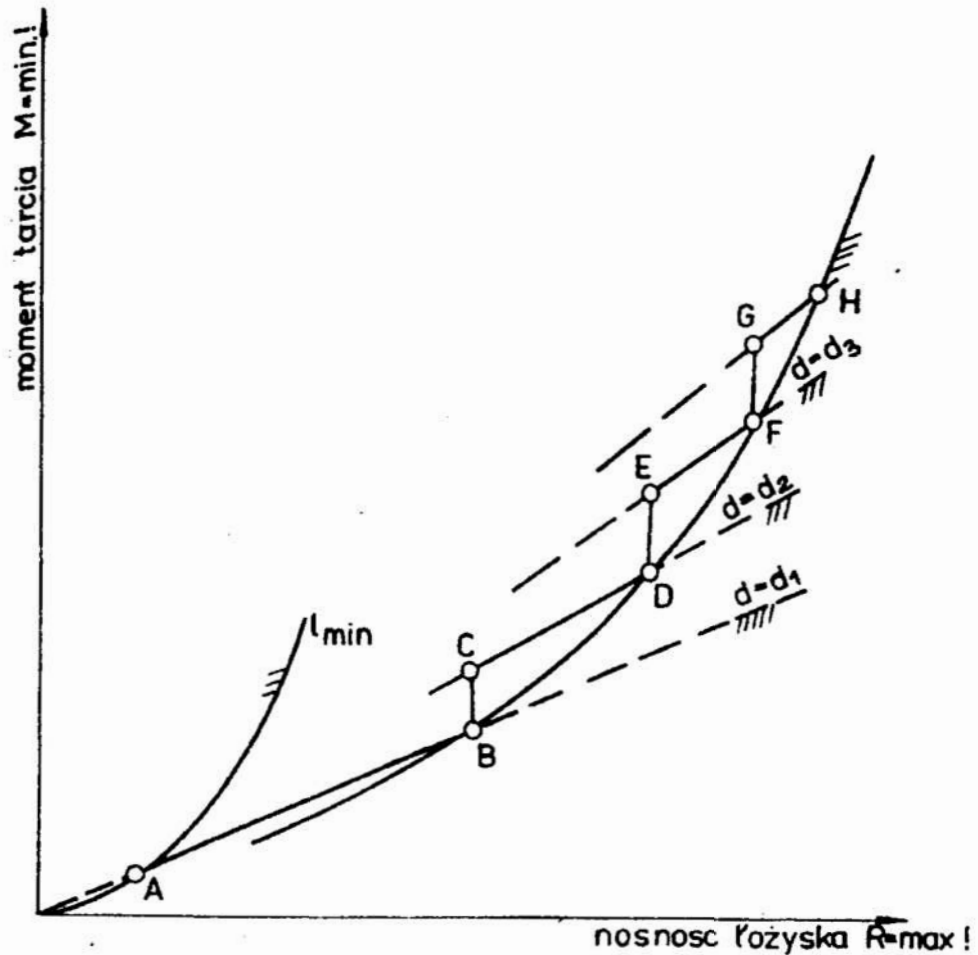
Jest powszechnie znane, że pewne ograniczenia są nieefektywne, jednak z góry nie wiadomo które. Odpowiedź może przynieść ujęcie polioptymalizacyjne. Na tym samym przykładzie łożyska można pokazać, że ta 'nieefektywność' ograniczenia może zależeć od wartości innych zmiennych: np. ograniczenie wytrzymałościowe $s < k$ jest efektywne tylko gdy wymagana nośność $R > R_1$. Oznacza to, że jeśli konstruktor zdecyduje się na wariant W (rys.2) to można zastosować gorszy materiał o wytrzymałości k' mniejszej niż pierwotnie zakładano k (rys.3).



Rys. 3. Dla wariantu W o nośności $R_2 < R_1$ ograniczenie $\sigma_g < k_{go}$ jest nieefektywne; efektywnym staje się ograniczenie $d > d_{min}$

4.2. Koszty dyskretyzacji zmiennych decyzyjnych

W ogólności, zbiór wariantów polioptymalnych pogarsza się, gdy jedna zmienna decyzyjna może przyjmować tylko niektóre wartości, np. staje się zmienną dyskretną, jak w przypadku normalizacji średnic otworów. Ujęcie polioptymalizacyjne pozwala ocenić stopień pogorszenia oraz wskazać te nieliczne warianty, których jakość nie uległa pogorszeniu wskutek dyskretyzacji (np. punkty B, D, F, H na rys.4).



Rys. 4. Zbiór wariantów polioptymalnych A-B-D-F-H gdy średnica d jest zmienną ciągłą; linia A-B-C-D-E-F-G-H gdy średnica d jest zmienną dyskretną

4.3. Formułowanie niektórych szczegółowych zasad konstrukcyjnych

Jak wspomniano w p. 4.1 warianty polioptymalne często leżą na ograniczeniach. Z definicji wariantu polioptymalnego wynika, że jest nieracjonalne stosowanie rozwiązań nie należących do zbioru wariantów polioptymalnych, zatem dla konkretnej klasy zadań konstrukcyjnych można sformułować konkretne zalecenie, np. takie: jeśli

$$M \leq M_E \text{ to } \frac{l}{d} \leq \sqrt{0.2 \frac{k_{go}}{P_{dop}}} \text{ gdzie } M_E \text{ jest momentem tarcia odpowiadającym}$$

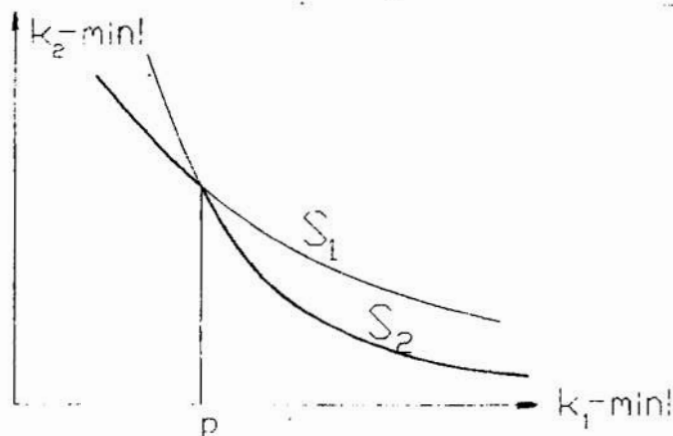
punktowi E (rys.2 i 3). Wzór ten jest przekształceniem ograniczeń (1) i (2) podanych na rys.1 oraz uwzględnia, że $M = 0,5R\mu d$.

Zasady takie z reguły dotychczas formułowane były intuicyjnie (heurystycznie) przez doświadczonych konstruktorów, bez formalnego uzasadnienia ani bez podania zakresu ich ważności i często podawane były w postaci rozmytej. Obecnie, wykorzystując ujęcie polioptymalizacyjne można je sformułować w postaci matematycznej i uzasadnić.

4.4. Równoczesna optymalizacja struktury i parametrów

Zdefiniujmy tu pojęcie struktury: powiemy, że dwa obiekty mają różną strukturę jeśli ich modele matematyczne (służące do optymalizacji) są różne, tzn. różnią się

postacią równań a nie tylko parametrami równań. Wówczas, przy klasycznym podejściu, optymalizacja musi być przeprowadzana niezależnie dla każdego obiektu. W praktyce najczęściej projektant intuicyjnie wybiera strukturę obiektu i na nim dokonuje optymalizacji parametrycznej. Nie trzeba dowodzić, że ta pierwsza intuicyjna decyzja dotycząca struktury jest o wiele ważniejsza i że takie działanie jest nieracjonalne.



Rys. 5. Przykładowe zbiory rozwiązań polioptymalnych: dla struktury S_1 i dla struktury S_2 : widać, że dla małych wartości k_1 optymalna jest struktura S_1 , zaś dla $k_1 > p$ optymalna jest struktura S_2

Tymczasem ujęcie polioptymalizacyjne umożliwia jednoczesną optymalizację parametrów i struktury obiektu, tzn. umożliwia w jednym zadaniu określenie zarówno struktury jak i parametrów obiektu. Osiąga się to przez wyznaczenie zbiorów wariantów niezdominowanych dla każdej struktury niezależnie i porównanie ich między sobą. To porównanie jest natychmiastowe, jeśli odpowiednie zbiory wariantów polioptymalnych przedstawione są w tej samej przestrzeni celów (Rys. 5)

Literatura

- [1] Osiński J., Wróbel J.: Teoria konstrukcji maszyn, PWN Warszawa 1982.
- [2] Tarnowski W.: Model procesu wyboru w projektowaniu technicznym. Zeszyty Naukowe Politechniki Śląskiej, seria Automatyka, z.72, Gliwice 1984.
- [3] Tarnowski W.: Polyoptimization as a Decision - Aiding Tool in Design. Proceedings of International Conference on Engineering Design ICED '88, Budapest, vol. 1, pp.315-322, ed. Heurista, Zurich 1988.
- [4] Tarnowski W.: Metody koncipowania - heurystyczne metody poszukiwania rozwiązań projektowych. Skrypt Uczelniany Politechniki Śląskiej, nr 1277, Gliwice 1986.
- [5] Tarnowski W.: Polioptymalizacja i optymalizacja hierarchiczna w projektowaniu technicznym - przykład. Materiały XXIV Symposium PTMTiS "Modelowanie w mechanice", ss. 159-174, Beskid Śląski 1985.

HOW THE POLYOPTIMIZATION BRIDGES THE GAP BETWEEN ALGORITHM AND HEURISTIC PROBLEM IN THE CAD SYSTEMS

SUMMARY

The polyoptimization problem in engineering design is formulated. It is argued that in the preliminary design the polyoptimization is of a secondary importance, while the multiattribute scalar optimization plays a primary role. Then the specific advantages of polyoptimization in the detail design are exhibited, namely: (i) the possibility of an assessment of specific requirements and specifications in design, (ii) the possibility of an estimation of costs of decision variables discretization. It is emphasised that the polyoptimization may play a generalization role in design, for example: 1) the simultaneous optimization of the structure and its parameters in one task, 2) the formulation of new design rules, as well as possibility of verification of already known heuristic principles.

POSZUKIWANIE USZKODZEŃ W URZĄDZENIACH MECHANICZNO-ELEKTRONICZNYCH

W pracy został przedstawiony system ekspertowy do lokalizacji uszkodzonych elementów mechanicznych i elektronicznych. Wiedza specjalisty jest reprezentowana w postaci reguł, a jej przetwarzanie jest realizowane metodą wnioskowania wstępującego oraz zstępującego. System został praktycznie przetestowany w zastosowaniu do stacji radiolokacyjnej.

1. Wprowadzenie

Pojawienie się na rynku szybkich komputerów o dużej pojemności pamięci operacyjnej umożliwiło upowszechnienie technik sztucznej inteligencji, zwanych systemami ekspertowymi. Odzwierciedlają one sposób wnioskowania specjalisty w pewnej wąskiej dziedzinie, wykorzystującego swoją wiedzę doświadczalną do rozwiązania określonego zadania. Jedną z dziedzin, w których systemy ekspertowe osiągnęły duże praktyczne zastosowanie jest diagnostyka. Najbardziej znane diagnostyczne systemy ekspertowe w medycynie to: MYCIN [1], ONOCIN [2] oraz w nauce i technice to: SACON [3], DART [4], REACTOR [5]. Są one stosowane jako konsultant: lekarza, inżyniera i technika zatrudnionego przy naprawie sprzętu.

Celem naszej pracy jest rozwiązanie wybranych problemów diagnozowania urządzeń mechaniczno-elektronicznych z zastosowaniem metod sztucznej inteligencji oraz wykazanie, że zastosowanie metod reprezentacji i przetwarzania wiedzy, obok tradycyjnych metod numerycznych i diagnostycznych, stwarza nowe możliwości w zakresie automatyzacji diagnozowania tych urządzeń.

W procesie usuwania uszkodzeń można wyróżnić 4 etapy:

1. Stwierdzenie uszkodzenia.
2. Lokalizacja uszkodzonego podzespołu .
3. Identyfikacja uszkodzonego elementu.
4. Usunięcie uszkodzenia, regulacja i sprawdzenie zdolności systemu do pracy.

Opracowany system dotyczy w głównej mierze dwóch śród-kowych etapów: drugiego i trzeciego nazywanych poszukiwaniem uszkodzeń lub diagnozowaniem urządzeń.

Etapy skrajne w systemie są realizowane: pierwszy poprzez wywiad przeprowadzony z użytkownikiem, ustalając objawy uszkodzenia oraz czwarty poprzez podanie: porady jakie elementy, bądź element należy wymienić lub jaką procedurę regulacyjną zastosować ustalając warunki pracy.

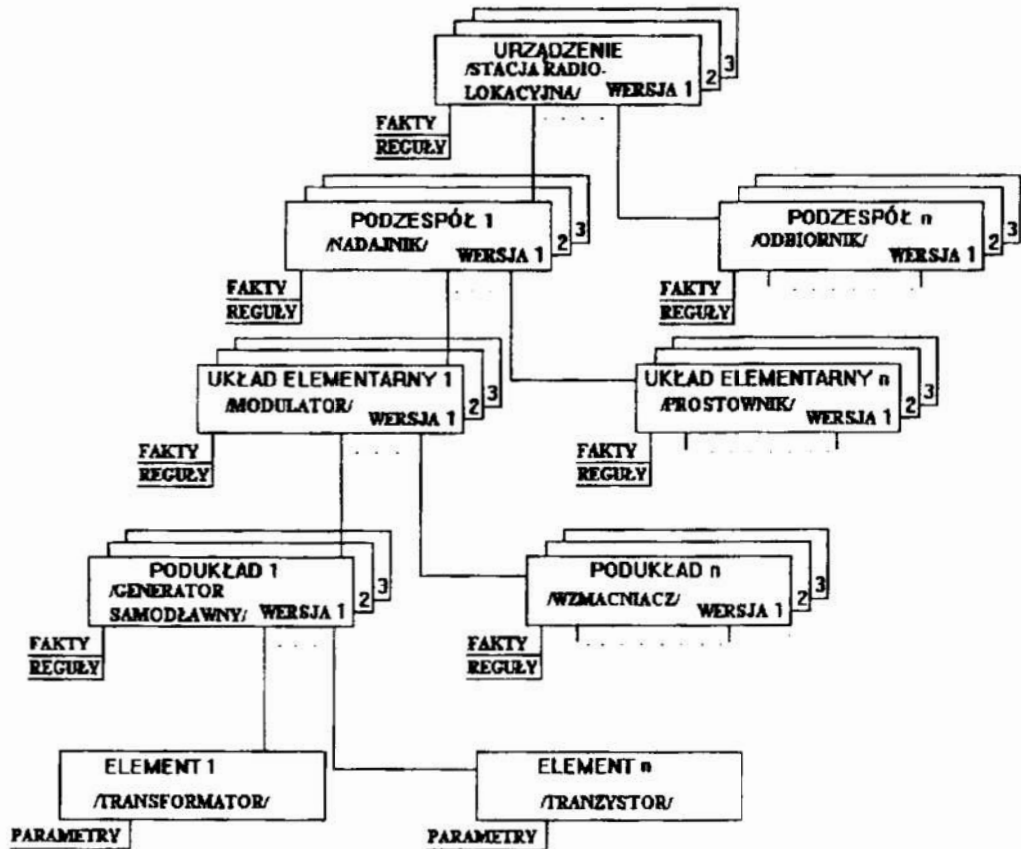
Podstawowymi elementami prezentowanego systemu ekspertowego są baza wiedzy i mechanizm wnioskowania.

* Wyższa Szkoła Inżynierska w Koszalinie

** Wyższa Szkoła Oficerska Wojsk Obrony Przeciwlotniczej w Koszalinie

2. Baza wiedzy

Baza wiedzy zawiera wiedzę specjalistów o urządzeniu mechaniczno-elektronicznym: o jego podstawowych sygnałach i przebiegach generowanych przez poszczególne podzespoły, wiedzę inżynierską o funkcjach podzespołów i elementów oraz wiedzę serwisową techników naprawiających sprzęt.



Rys. 1. Ramowa reprezentacja wiedzy do diagnostyki stacji radiolokacyjnej

Baza wiedzy akceptowana przez komputer jest nazwana reprezentacją wiedzy. W opracowanym systemie zastosowano formę mieszaną reprezentacji wiedzy, tj. w postaci ram [6], faktów i reguł.

Ramy pokazane na rys.1 są węzłami sieci semantycznej, tworzącej opis urządzenia np. stacji radiolokacyjnej i są one zapisywane w postaci predykatów. System ramowy umożliwił podzielenie wiedzy zgromadzonej w bazie wiedzy. Węzłami są: urządzenie, podzespoły, układy, podukłady i elementy. Każdy z węzłów ma określone własności; w postaci reguł, faktów, porad czy pytań. Ponadto urządzenie traktowane jest jako jedna z wersji urządzeń rozbudowujących system ramowy w górę. Tak zbudowana sieć semantyczna jest hierarchiczną strukturą drzewa, w której ramy są liśćmi, a gałęzie określają relację dziedziczenia; czyli na przykład do ram takich jak element będzie się stosowała wiedza o podukładzie, do którego element należy.

Ramy zapisane jako predykaty, reguły i fakty są szczególnymi przypadkami klauzul Horna w logice predykatów pierwszego rzędu. Przedstawienie wiedzy w postaci klauzul jest warunkiem wystarczającym do przeprowadzenia dedukcyjnego wnioskowania.

Klauzula składa się z rozłączonych atomów lub negacji atomów ewentualnie poprzedzonych kwantyfikatorem połączonych spójnikami logicznymi $\wedge, \vee, \Rightarrow, \sim$.

Ogólnie posiada ona postać:

$$B_1, B_2, \dots, B_m \leftarrow A_1, A_2, \dots, A_n \quad m, n \geq 0 \quad (1)$$

gdzie: $B_1, B_2, \dots, B_m, A_1, A_2, \dots, A_n$ są atomami. Atomy A_1, A_2, \dots, A_n są przesłankami klauzuli, a B_1, B_2, \dots, B_m alternatywnymi konkluzjami klauzuli.

Jeżeli klauzula zawiera zmienne x_1, x_2, \dots, x_k to jest ona interpretowana:

$$\forall_{x_i} (A_1 \wedge A_2 \wedge \dots \wedge A_n) \Rightarrow (B_1 \vee B_2 \vee \dots \vee B_m) \quad i=1, \dots, k \quad (2)$$

Atom jest wyrażeniem o postaci $P(t_1, t_2, \dots, t_n)$, gdzie P jest symbolem predykatu, a t_1, t_2, \dots, t_n termami.

Term może być stałą (symbolem), zmienną, lub wyrażeniem $f(t_1, t_2, \dots, t_n)$, gdzie f jest symbolem funkcji, a t_1, t_2, \dots, t_n termami.

Jeżeli $m \leq 1$ to klauzula nazywa się klauzulą Horna. Ten rodzaj klauzuli może przyjmować cztery formy:

1. dla $m=1, n=0$ otrzymuje się fakt:

$$B \leftarrow \quad (3)$$

(dla każdego x_1, x_2, \dots, x_k zachodzi B),

2. dla $m=1, n \neq 0$ otrzymuje się regułę:

$$B \leftarrow A_1, A_2, \dots, A_n \quad (4)$$

(dla każdego x_1, x_2, \dots, x_k zachodzi $(A_1 \wedge A_2 \wedge \dots \wedge A_n) \Rightarrow B$),

3. dla $m=0, n \neq 0$ otrzymuje się zaprzeczenie

$$\leftarrow A_1, A_2, \dots, A_n \quad (5)$$

(dla żadnego x_1, x_2, \dots, x_k nie zachodzi $A_1 \wedge A_2 \wedge \dots \wedge A_n$),

4. dla $m=0, n=0$ otrzymuje się sprzeczność:

$$\leftarrow \quad (6)$$

(klauzula pusta o wartości logicznej fałsz).

Wprowadzenie klauzul zaprzeczenia i sprzeczności umożliwia dowodzenie nie wprost, co ma zastosowanie w dedukcyjnym wnioskowaniu zstępującym.

3. Mechanizm wnioskowania

Wiedza zgromadzona w bazie wiedzy systemu ekspertowego jest przetwarzana przez mechanizm wnioskowania. Mechanizm wnioskowania systemu odzwierciedla proces rozumowania specjalisty naprawiającego urządzenie. Czynności wykonywane przy rozwiązywaniu zadań związanych z lokalizacją uszkodzenia można traktować jako poszukiwanie uszkodzeń obiektów o zadanej charakterystyce. Kompletne przebadanie przestrzeni wszystkich uszkodzeń urządzenia np. stacji radiolokacyjnej jest czasochłonne. W celu poprawienia efektywności poszukiwania w prezentowanym systemie korzysta się z heurystyk, tj. praktycznych reguł często intuicyjnych i empirycznych, wynikających z doświadczenia specjalistów. Specjaliści, często oprócz rozumowania opartego o dedukacyjne wnioskowanie, wykorzystują swoje doświadczenie do stawiania hipotez opartych na obserwacjach określonych zjawisk w naprawianym urządzeniu. Hipotezy są stosowane do ograniczenia przestrzeni poszukiwania [7] i stanowią podstawową wiedzę w poszukiwaniu uszkodzenia.

W mechanizmie wnioskowania proponowanego systemu diagnostycznego, zastosowano dwukierunkowe, oparte o wnioskowanie dedukcyjne, strategie poszukiwań:

- wstępujące (nazywane także w przód albo progresyjne) do lokalizacji uszkodzonego podzespołu;
- zstępujące (nazywane także wstecz albo regresyjne), oraz stawianie hipotez [8-9].

W pierwszym etapie wnioskowania wstępującego z posiadanej wiedzy zgromadzonej w bazie wiedzy i symptomów system ustala, które z podzespołów są na pewno uszkodzone i stawia hipotezy o rodzaju uszkodzenia. Każda z hipotez posiada listę elementów powodujących dany rodzaj uszkodzenia w podzespołe, uporządkowaną według malejącego prawdopodobieństwa ich występowania. W drugim etapie realizowany jest proces wnioskowania zstępującego do czasu, aż jedna z hipotez zostanie potwierdzona przez ustalenie, który element z listy jest uszkodzony.

Użytkownik pracuje z systemem w sposób interakcyjny. W przypadkach, gdy wnioski nie mogą być wyciągnięte z faktów i reguł, system zadaje szczegółowe pytania użytkownikowi: dotyczą one zachowania układu, podukładów, przebiegów, prądów czy napięć w określonych węzłach bądź gałęziach układu. Odpowiedź użytkownika na pytanie powiększa bazę wiedzy. Ponadto system posiada wbudowany mechanizm wyjaśniania podjętych decyzji na podstawie śledzenia procesu wnioskowania.

Wnioskowanie wstępujące i zstępujące są wnioskowaniami dedukcyjnymi wykorzystującymi rezolucję oraz metodę porównywania ze wzorcem.

Rezolucja $[(p \vee q) \wedge (\sim q \vee r)]$, jest uogólnieniem dwóch następujących praw dedukcji:

1. modus ponens: $[p \wedge (p \Rightarrow q)] \Rightarrow q$ jeżeli prawdziwe jest zdanie (fakt) p oraz prawdziwa jest implikacja $p \Rightarrow q$ to logiczną konsekwencją jest prawdziwość zdania (faktu) q ,

2. modus tollens: $[\sim q \wedge (p \Rightarrow q)] \Rightarrow \sim p$ jeżeli prawdziwe jest zdanie $\sim q$ oraz prawdziwa jest implikacja $p \Rightarrow q$ to logiczną konsekwencją jest prawdziwość zdania $\sim p$

Metoda porównywania ze wzorcem, nazywana także metodą dopasowania, porównuje dwa zdania (w tym jeden wzorzec) w procesie unifikacji.

Unifikacja, nazywana także uzgadnianiem jest takim ciągiem podstawień termów za zmienne w porównywanych zdaniach, który powoduje, że zdania stają się identyczne.

Jeżeli unifikacja jest wykonalna tzn. istnieje podstawienie uzgadniające, to wynikiem porównania dwóch zdań w metodzie porównania ze wzorcem jest wartość logiczna prawda, a w przeciwnym razie fałsz.

W opisie przyjęto następujące oznaczenia:

1. podstawienie termu t za zmienną x oznacza się $\{t/x\}$,
2. ciąg podstawień $\{t_1/x_1, t_2/x_2, \dots, t_k/x_k\}$ oznacza się Θ ,
3. rezultat podstawienia Θ w zdaniu L nazywa się konkretyzacją zdania L i oznacza się L^Θ .
4. jeżeli dla dwóch porównywanych zdań L_1 i L_2 istnieje podstawienie uzgadniające, takie że są one identyczne, to oznacza się to $L_1^\Theta = L_2^\Theta$.

Powyższe prawa i metoda porównania z wzorcem są stosowane wielokrotnie do wyprowadzania nowych faktów ze znanych faktów metodą wnioskowania wstępującego lub do zweryfikowania postawionych hipotez metodą wnioskowania zstępującego.

Mechanizm wnioskowania wstępującego z prawdziwych faktów oraz reguł opisujących konsekwencje różnych faktów wyprowadza nowe, prawdziwe fakty. Do tego celu wykorzystuje wielokrotnie regułę dedukcji modus ponens i metodę porównywania z wzorcem tj. porównania prawdziwych faktów z przesłankami reguły. Konkluzja B reguły

$$B \leftarrow A_1, A_2, \dots, A_n \quad (7)$$

może stać się nowym faktem jeżeli wśród starych faktów w bazie wiedzy są fakty:

$$A'_1, A'_2, \dots, A'_n \leftarrow \quad (8)$$

takie, że zastosowanie metody porównania z wzorcem pozwala zastąpić wszystkie przesłanki A_1, A_2, \dots, A_n faktami $A'_1 \leftarrow, A'_2 \leftarrow, \dots, A'_n \leftarrow$. Podstawienie Θ w porównaniu z wzorcem zamienia zmienne termami tak, że dla każdego i zachodzi

$$A_i \Theta = A'_i \Theta \quad (9)$$

Równość ta stwierdza, że wszystkie przesłanki są identyczne z faktami, a więc prawdziwe. Stosując przyjętą notację każda przesłanka A_i zunifikowana (uzgodniona) z faktem $A'_i \leftarrow$ metodą porównania z wzorcem, jest usuwana (zastępowana pustym zbiorem przesłanek faktu $A'_i \leftarrow$) z reguły $B \leftarrow A_1, A_2, \dots, A_n$. Gdy wszystkie przesłanki zostaną usunięte, to pozostaje fakt $B \leftarrow$. Tak wyprowadzony w jednym kroku wnioskowania wstępującego fakt jest dołączany do bazy wiedzy i

proces wyprowadzania kolejnych faktów jest powtarzany aż do osiągnięcia poszukiwanego celu lub wyprowadzenia wszystkich faktów. W celu zredukowania wyprowadzania zbędnych faktów system przeprowadza zwięzły wywiad z użytkownikiem. Odpowiedzi na pytania systemu odnośnie objawów zewnętrznych uszkodzeń stają się faktami powiększającymi wiedzę systemu ekspertowego. Z posiadanej wiedzy w postaci reguł i faktów system wnioskuje, które z podzespołów są na pewno uszkodzone lub ustawia hipotezy o uszkodzonych podzespołach według kolejności malejącego prawdopodobieństwa ich występowania.

Do wnioskowania zstępującego została zastosowana strategia dowodzenia nie wprost. Z negacji zdania (hipotezy do potwierdzenia) wyprowadza się następstwa tak długo, aż dojdzie się do sprzeczności tj. do negacji zdania (faktu) uznanego za prawdziwe. Ostatecznym rezultatem jest stwierdzenie, że zdanie zanegowane na początku jest prawdziwe. Ta strategia jest podstawą wnioskowania zstępującego, które wykorzystuje wielokrotnie regułę dedukcji modus tollens i metodę porównywania ze wzorem. Wnioskowanie zstępujące potwierdza hipotezę jako cel. Załóżmy, że dane są: cel w postaci zanegowanej hipotezy (10), fakt (11) i reguła (12).

$$\leftarrow A_1, A_2, \dots, A_n \quad (10)$$

$$A'_1 \leftarrow \quad (11)$$

$$B \leftarrow B_1, B_2, \dots, B_m \quad (12)$$

Stosując prawo modus tollens i metodę porównywania z wzorcem:

$$A_i \ominus = B \ominus, \quad (13)$$

$$A_1 \ominus = A'_1 \ominus \quad (14)$$

uzyskuje się nowy cel w postaci zanegowanej:

$$\leftarrow (A_2, \dots, B_1, B_2, \dots, B_m, \dots, A_n) \ominus \quad (15)$$

gdzie przesłanka A_1 okazała się identyczna z faktem $A'_1 \leftarrow$ i usunięta (zastąpiona pustym zbiorem przesłanek faktu), a przesłanka A_i identyczna z konkluzją B reguły i zastąpiona jej przesłankami B_1, B_2, \dots, B_m . W jednym kroku wnioskowania zstępującego podcel A_i celu $\leftarrow A_1, A_2, \dots, A_n$ został rozłożony na podcele B_1, B_2, \dots, B_m . Wszystkie podcele są w kolejnych krokach rozkładane na dalsze podcele w wyniku uzgadniania ich z konkluzjami reguł, bądź redukowane przez uzgadnianie z faktami. Jeżeli w procesie wnioskowania zstępującego cel zredukuje się do klauzuli pustej czyli sprzeczności, to hipoteza (cel) jest potwierdzona. Natomiast proces kończy się porażką, jeżeli choć jedna z przesłanek (podcel) nie została uzgodniona.

4. Przykład praktyczny

Metoda lokalizacji uszkodzonych podzespołów i elementów z wykorzystaniem systemu ekspertowego została praktycznie sprawdzona dla jednej ze stacji radiolokacyjnych. Stacja posiada układ kontroli pozwalający ocenić jej stan techniczny. Objawy o ewentualnych uszkodzeniach powstają w wyniku obserwacji układu kształtowania sygnałów kontrolnych oraz lampek sygnalizacyjnych. Napięcia zespołów zasilania, moc nadajników, parametry torów odbiorczych określa się za pomocą wbudowanych mierników pomiarowych. Przykładem strategii poszukiwania ze stawianiem hipotez jest procedura stosowana w lokalizacji uszkodzenia stwierdzonego na podstawie kontroli funkcjonowania stacji radiolokacyjnej.

W wyniku wywiadu przeprowadzonego z użytkownikiem system diagnostyczny gromadzi fakty (objawy). W pierwszym etapie wnioskowania wstępującego, lokalizuje uszkodzony podzespół oraz stawia hipotezę o uszkodzonym elemencie. W drugim etapie wnioskowania zstępującego weryfikuje kolejno hipotezy z listy elementów powodujących dane uszkodzenie w celu identyfikacji uszkodzonego elementu. Przykładowe reguły do lokalizacji uszkodzonego układu przełączania prądnic w dwóch krokach wnioskowania wstępującego (reguły 20 i 21), do postawienia hipotez o uszkodzonych elementach (reguła 42) i do zweryfikowania hipotezy tj. identyfikacji uszkodzonego stycznika K1 w dwóch krokach wnioskowania zstępującego (reguły 84 i 85) przedstawiają się następująco:

REGUŁA LOKALIZUJĄCA 20

zespół_przełączający_WKA_333(brak_napięcia_220V) ←
układ_zasilania(brak_napięcia_220V),
silnik_turbinowy(obroty_w_normie),
prądnic_a_S-75-406/12(sprawna),
układ_zabezpieczający(sprawny).

REGUŁA LOKALIZUJĄCA 21

układ_przełączania_prądnic(brak_napięcia_220V) ←
zespół_przełączający_WKA_333(brak_napięcia_220V),
złącze_kontrolne_4Sz33(brak_napięcia_220V),
złącze_kontrolne_4Sz1(brak_napięcia_220V).

REGUŁA HIPOTEZUJĄCA 42

układ_przełączania_prądnic(brak_napięcia_220V) ←
stycznik_K1(zwarcie);
zespół_przełączników_P1-P9(rozwarcie);
przycisk_sterowania_KW1(rozwarcie);

REGUŁA IDENTYFIKUJĄCA 84

podukład_styczników(brak_napięcia_220V) ←
układ_przełączania_prądnic(brak_napięcia_220V),
punkt_kontrolny_Sz1/1(brak_napięcia_220V),

REGUŁA IDENTYFIKUJĄCA 85

stycznik_K1(zwarcie) ←
podukład_styczników(brak_napięcia_220V),
punkt_kontrolny_Sz1/2(brak_napięcia_220V),
punkt_kontrolny_Sz1/3(brak_napięcia_220V).

5. Uwagi końcowe

System lokalizacji uszkodzonego elementu został zaimplementowany na mikrokomputerze IBM PC XT w języku GC LISP. Ma około 200 reguł i w połączeniu z systemem lokalizacji uszkodzonego podzespołu, zawierającym około 100 reguł, jest w fazie testowania. Jest on typowym systemem szkieletowym, którego wiedza może być wymieniona, powiększona lub modyfikowana bez konieczności zmian w mechanizmie wnioskowania.

6. Literatura

- [1] Shortliffe E.H., "Computer-based medical decision making", New York, Elsevier, 1976.
- [2] Shortliffe E.H., et al., "ONCOCIN: An expert system for oncology protocol management", IJCAI, 1981, pp. 876-881.
- [3] Bennett J.S., et al., "SACON: A knowledge-based consultant for structural analysis", IJCAI, 1979, pp. 47-49.
- [4] Bennett J.S. and Hollander C.R., "DART: An expert system for computer diagnosis", IJCAI, 1981, 843-845.
- [5] Nelson W.R., "REACTOR: An expert system for diagnosis and treatment of nuclear reactor accidents", AAAI, 1982.
- [6] Minsky M., "Framework for representation knowledge" in Psychology of computer vision, P. Winston, McGraw Hill, 1975.
- [7] Goldberg A. and Pohl I., "Is complexity theory of use to AI?", Artificial Human Intelligence, edited by A. Elithorn, R. Banerji, North-Holland, 1984, pp. 43-56.
- [8] Wawryn K. and Zinka W., "A prototype expert system for fault diagnosis in electronic device", 9th ECCTD, 1989, pp. 677-680.
- [9] Wawryn K. and Drabarek J., "Fault diagnosis in electronic system based on AI techniques", 16th Nat. Conf. Network Theory and Electronic Circuits, 1993, pp. 548-553.

Troubleshooting in electro-mechanical systems

Summary

A prototype expert system using novel reasoning style in diagnostic knowledge is presented. It helps a troubleshooting engineer to locate faulty functional modules of an electromechanical system, assists in diagnosing faulty electrical and mechanical components, wirings, printed circuit boards or other chassis-mounted parts which could have caused failure and recommends appropriate repair in the system.

**APPLICATION OF NUMERICAL METHODS
FOR SHAPE OPTIMIZATION OF NOTCHES
IN MACHINE OR STRUCTURAL COMPONENTS**

1. Introduction

For many years, numerous numerical methods have brought a great help to engineers when designing machine or structural components of any kind. They have been applied to every branch of mechanics in order to solve analysis problems. The success of analysis techniques (Finite Element Method, Boundary Element Method) in determining the stress and displacement solutions to complex engineering problems has enabled engineers to use these and new numerical techniques to optimize design. An extensive literature developed on optimization of machine or structural components which are defined by cross-section and thickness variables (size optimization). A more important problem, from the point of view of mechanical (structural) design, is determination of the shape of 2D, or 3D components (shape optimization). For such problems, the shape of the machine or structural element must be treated as design variables.

As it is already known the presence of notches (holes, fillets, undercuts, cut-outs etc.) in a machine component results in increasing the local tensions which are called stress concentrations. One of the principal features of stress concentration is the so called Stress Concentration Factor (SCF) [8]. To decrease a SCF, a possibility exists to minimize stress by changing the shape of notch. Such a class of optimization problems is referred to as the shape notch optimization.

One of the first treatment (involving finite element method) of the general problem of selection shape of the structure as the design variable is presented by Zienkiewicz and Campbell [18]. Since then there have been many contributions in which mostly 2-D or 3-D finite or boundary element methods have been coupled with numerical optimization techniques [1,3-7,11-14,16-17]. Because of the lack of space, a review of literature in this field is omitted in this paper.

2. Problem statement and solution procedures

Problem definition

A shape optimization problem of notches can be stated mathematically as one of the minimization of the volume V of the component, or the minimization of the maximum effectivestress σ_e , i.e. $\min [\max \sigma_e]$ for a given load. The problem is stated as

*Wyższa Szkoła Inżynierska w Koszalinie, ul. Raclawicka 15-17, KOSZALIN

$$\min V(\mathbf{D}) \text{ or } \min [\max \sigma_e(\mathbf{D})] \quad (1)$$

subject to

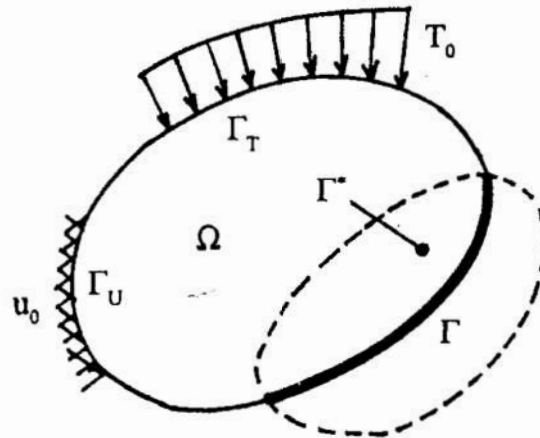
$$g_j(\mathbf{D}) \leq 0, \quad j = 1, k, \quad \text{and} \quad \mathbf{D}_l \leq \mathbf{D} \leq \mathbf{D}_u \quad (2)$$

where \mathbf{D} is n -dimensional vector of design variables, g_j are inequality constraints (such as : stress or displacement), \mathbf{D}_l and \mathbf{D}_u are the lower and upper limits on \mathbf{D} respectively; and k is the number of inequality constraints.

Bound formulation can be used to solve the min-max problem. The problem is restated as a simple minimization problem in terms of an unknown bound D_{n+1} on the value of $\max \sigma_e(\mathbf{D})$

$$\min D_{n+1} \quad \text{subject to} \quad \sigma_{ej}(\mathbf{D}) - D_{n+1} \leq 0, \quad j = 1, l \quad \text{and the constraints (2),}$$

where l is the number of critical points, in which stresses are evaluated.



Governing equations:

$$\operatorname{div} \boldsymbol{\sigma} = 0 \quad \text{in } \Omega$$

(body forces are omitted)

$$\boldsymbol{\sigma} \mathbf{n} = \mathbf{T}_0 \quad \text{on } \Gamma_T$$

$$\mathbf{u} = \mathbf{u}_0 \quad \text{on } \Gamma_u \quad (i, j = 1, 3)$$

Fig. 1. Machine component undergoing shape modification, Γ_T - the boundary with given tractions, Γ_u - the boundary with given displacements, Γ^* - the variation domain, Γ - the modified boundary

Selection of shape design variables

The shape of the variable boundary should be represented in a convenient way. To define the shape of the notch (notches) recent research has emphasized the solution of so-called "natural variables". In this paper we describe the variable (modified) boundary of notches using one segmented Bezier's curves with the coordinates of some control nodes (key points) treated as design variables [12,14-16]. It is convenient to give prescribed movement directions to the key points (Fig. 2).

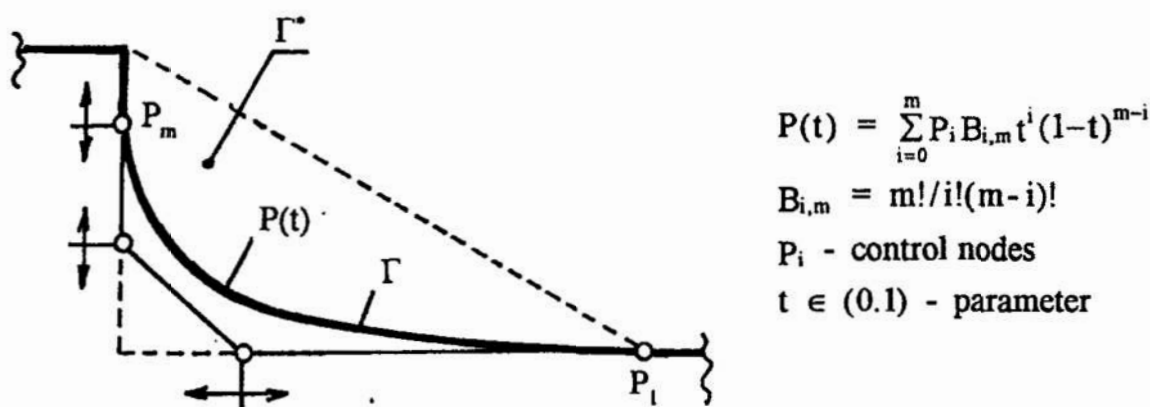


Fig. 2. Modified boundary: $P(t)$ - one segmented Bezier curve,
 → movement directions of control nodes

BEM or FEM method as a basis for optimization

Today the most efficient tools for the computation of stresses in a machine components are: the Boundary Element Method (BEM), the Finite Element Method (FEM) or coupling the BEM with the FEM. It is well known that different formulations of the BEM or the FEM lead to linear equations. The description of the BEM and the FEM is omitted in this paper. Figure 3 shows the BEM and FEM meshes for 2-D machine component (i.e. fillet problem).

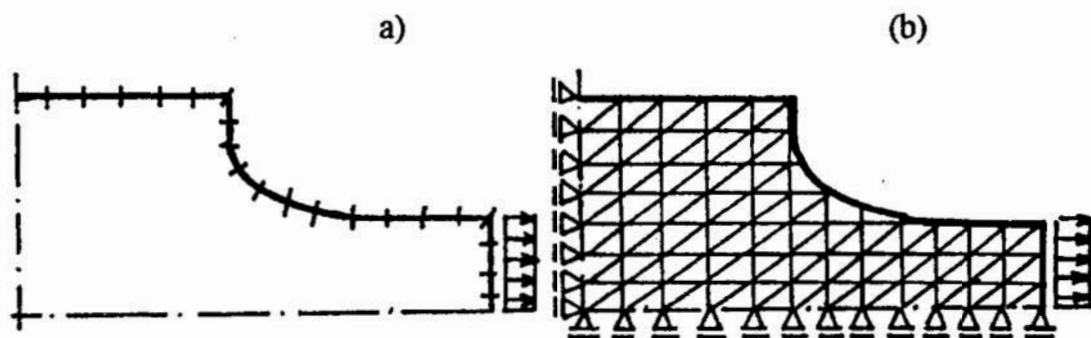


Fig. 3. Example of 2-D body (fillet): a) BEM mesh, b) FEM mesh

Sensitivity analysis

Having completed the BEM or FEM notch stress analysis, we now evaluate the sensitivity of the stresses of the current design to small changes in the design variables. Should be mentioned that the objective function is linear, and only there, there is a need to evaluate the gradients and of the behavioural constraints with respect to design

variables. Methods for evaluating the sensitivities may be purely analytical, semi-analytical or can be based on finite differences [3]. The above concepts of the sensitivity analysis are used in this paper.

Selection of optimization algorithm

Because objective function is linear, in this paper a sequential linear programming (SLP) has been used to obtain the optimal solution [9,10,12-14,17]. The behaviour constraints are sequentially linearized and the optimum solution will be obtained by using linear programming (LP), which can be efficiently solved by using the Simplex algorithm. A set of so called 'move limits' are added to the constraints of the SLP problem in order to control a stability and convergence of the algorithm. These move limits (additional constraints) may be specified as $\Delta \mathbf{D}_l \leq \Delta \mathbf{D} \leq \Delta \mathbf{D}_p$ where $\Delta \mathbf{D}_l$ and $\Delta \mathbf{D}_p$ are the lower and upper limits on $\Delta \mathbf{D}$ respectively. These move limits are reduced during the optimization process so that a solution is found. The optimization procedure is considered to have converged to the final minimum stress concentration if the condition $|\sigma_i^{\max} - \sigma_{i-1}^{\max}| / \sigma_{i-1}^{\max} \leq \varepsilon$ is satisfied for the two successive iterations, where σ_i and σ_{i-1} are the effective stresses after i and $i-1$ iterations, and ε is a user defined tolerance value (e.g. $\varepsilon = 0.01$ corresponds to a 1% convergence criterion).

3. Numerical examples

Example 1. Optimum shape of pressure vessel head under internal pressure

The axisymmetric thin shell (cylindrical pressure tank (vessel) with unknown shape of the head) is considered (Fig. 4). The head and the cylinder have been modelled with 59 conical finite elements (formulation based on Hellinger-Reissner principle is used). Figure 4 shows design variable assumed (standard Bezier curve), the starting profile and optimal profile of the pressure vessel head. Figure 5 shows the effective stress distribution in the pressure vessel. For details see Ref. [12]

Example 2. Optimum shapes of interacting holes in infinite plate under uniaxial tension

There is a way of the stress reduction by introducing additional (auxiliary, supplemental or defense) holes in the direction of loading on either side of the original hole (see papers cited in [14,16]). Further reduction is possible with the simultaneous shape optimization of the both kind of holes, with the object minimum stress concentration. The boundary of the both kind of holes is described using special Bezier curves. Owing to the symmetry of the problem only a quarter of the plate is considered. Taking into account that the infinite regions are easy to solve by FSM (indirect variant of the BEM) [2], the case of infinite plate is examined. A variation domain for both holes is shown in Fig. 6. A boundary of the central hole was modelled using 25 constant stress boundary elements, and a boundary of the supplemental hole using 50 boundary elements. That means about 40% reduction of stresses with comparison to a circular hole. Fig 6. shows the optimal design of the central and supplemental holes,

and Fig. 7 shows the tangential stress distribution along the design boundaries before, and after optimization. About 40% reduction of stresses with comparison to a circular hole, is obtained. For details see Ref. [16].

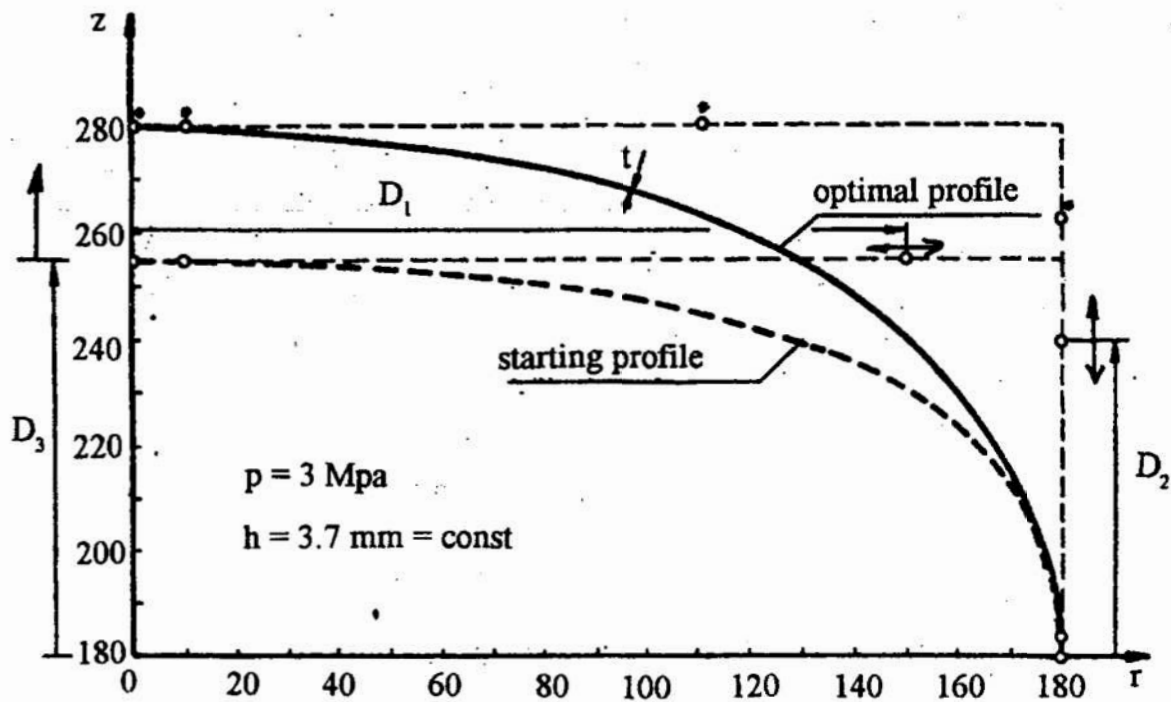


Fig. 4. Optimal (heavy line) and starting profile (dashed line) of pressure vessel head. \circ - the position of control points (design variables), \rightarrow the movement direction of control points, \circ with star - the position of control points for optimal solution

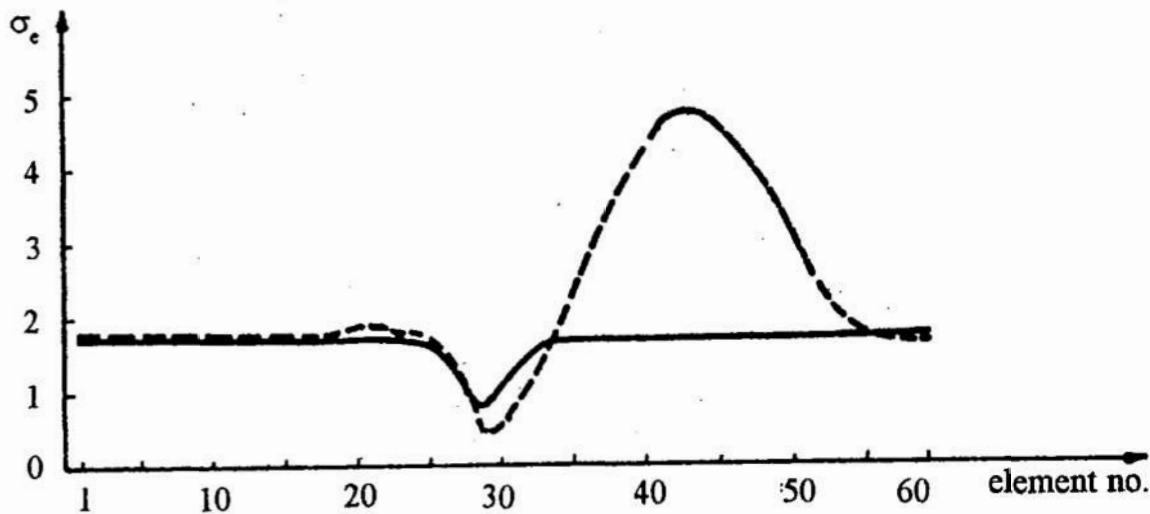


Fig. 5. Effective stress distribution in pressure vessel: for starting profile (dashed line), and for optimal profile (heavy line)

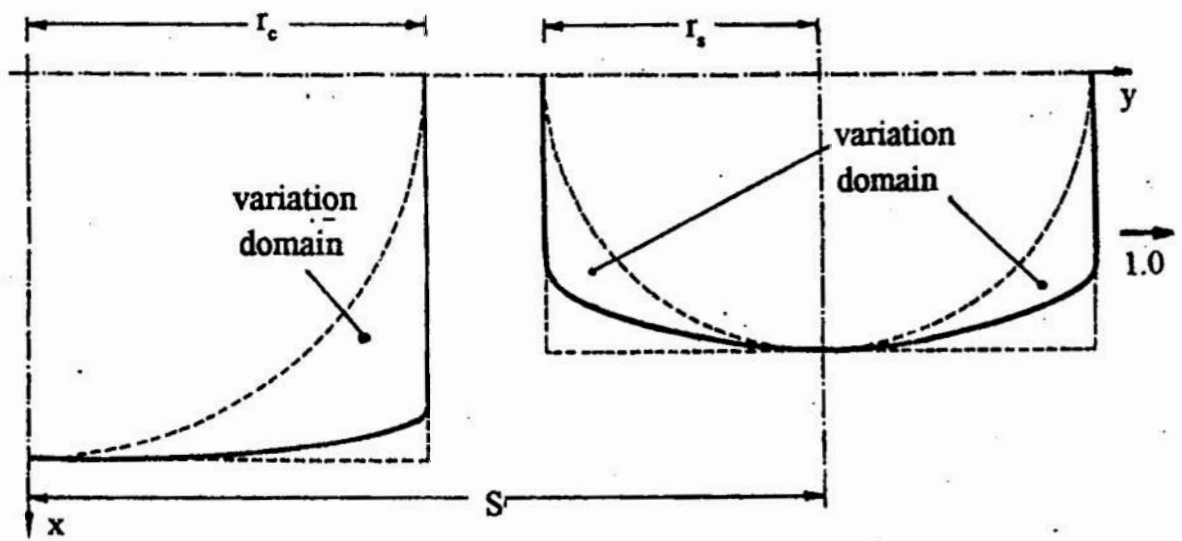


Fig. 6. Optimized geometries of central and supplemental holes in infinite plate subjected to uniaxial loading

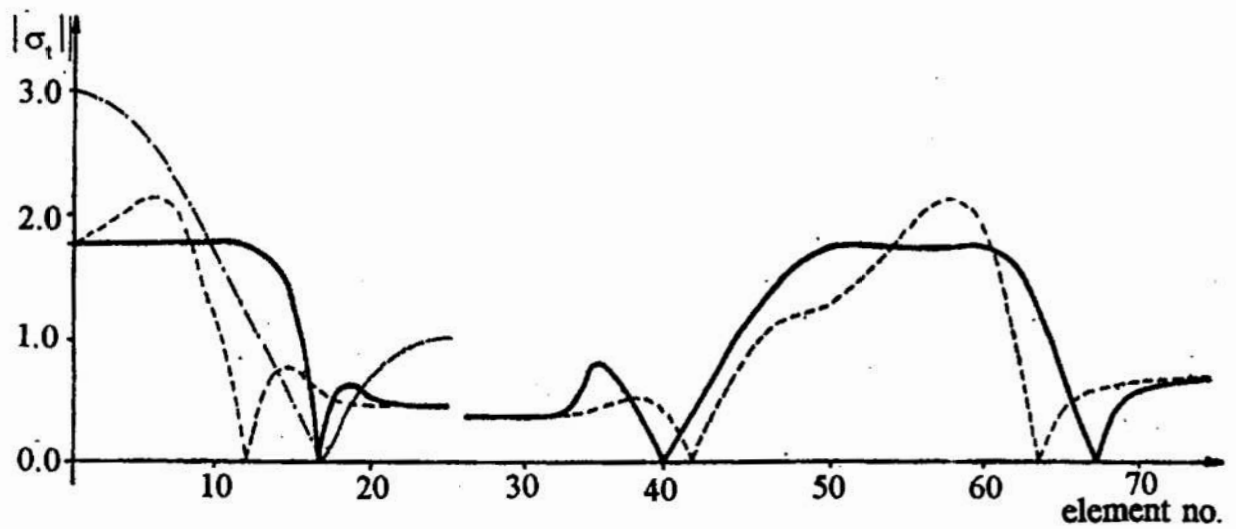


Fig. 7. Stress distribution along the design boundaries before (----), and after (—) optimization, and for a single hole (-·-·-)

5. Conclusion

The numerical methods combination of stress analysis of machine components, like BEM or FEM, numerical optimization methods, like sequential linear programming, and mathematical methods of computer graphics, like Bezier's curves, constitutes an effective and reliable approach by which an optimal shape of the notch (notches) with the minimum stress concentration is obtained.

References

1. Burczyński T.: *The boundary element method for selected analysis and optimization problems of deformable systems*, Scientific Reports, (in Polish), Silesian Technical University, Z. 97, 1989
2. Crouch S.L., Starfield A.M.: *Boundary element methods in solids mechanics*, George Allen and Unwin, London, 1983
3. Dems K., Mróz Z.: *On shape sensitivity approaches in the numerical analysis of structures*, Struct. Optim., Vol. 6, No. 2, 86- 93, 1993
4. Dybbro J.D., Holm N.C.: *On stress minimization for three-dimensional models*, Comput. Struct., Vol. 24, No. 4, 637-643, 1986
5. Kamiya N., Kita E.: *Local shape optimization of a two-dimensional elastic body*, Fin. Elem. Anal. Design, Vol. 6, 207-216, 1990
6. Mattheck C. Erb D., Bethge K. and Begemann U.: *Three-dimensional shape optimization of a bar with a rectangular hole*, Fatigue Fract. Engng Mater. Sci., Vol. 15, No. 4, 347-351, 1992
7. Moghaddasi-Tafreshi A. Fenner R.T.: *Design optimization using boundary element method*, Jnl Strain Anal., Vol. 26, 231-241, 1991
8. Olesiak Z.: *Stress concentration. Contact stresses*, In Strength of Structural Elements, Ed. Życzkowski M., PWN, Warszawa, Elsevier, Amsterdam, 527-612, 1991
9. Pedersen P., Carsten L.L.: *Design for minimum stress concentration by finite elements and linear programming*, J. Struct. Mech., Vol. 10, 375-391, 1982/83
10. Schittkowski K., Zillober C., Zotemantel R.: *Numerical comparison of nonlinear programming algorithms for structural optimization*, Struct. Optim., Vol. 7, No. 1/2, 1-19, 1994
11. Xu Ch., Yy M.: *Shape optimization of structures to minimize stress concentration*, Comput. Struct., Vol. 36, 491-497, 1990
12. Wilczyński B.: *Shape optimization of axisymmetric shells*, Proc. VIII Conf. Comp. Meth. in Mechanics, May 1987, Jadwisin, Vol. 2, 435-442, 1987 (in Polish)
13. Wilczyński B.: *Minimizing stress concentration around quasi-ovaloid hole infinite plate*, Scientific Reports, Silesian Technical University, Z.113, 423-428, 1993 (in Polish)
14. Wilczyński B.: *Stress minimization using Fictitious Stress Method*, II Int. Seminar on Boundary Element Techniques and Singularity Methods in Engineering, Wroclaw, December 1992, Scientific Reports, Technical University of Wroclaw, No. 74, 165-176, 1993
15. Wilczyński B.: *Accuracy of Fictitious Stress Method for stress concentration problems*, Scientific Reports, Technical University at Koszalin, No. 15, 192-210, 1993

16. Wilczyński B.: *Shape optimization of notches in 2-D machine components*, Scientific Reports, Technical University at Koszalin, No. 16, 61-70, 1994
17. Vanderplaats G.N.: *Effective use of numerical optimization in structural design*, Fin. Elem. Anal. Design, Vol. 6, 97-112, 1989
18. Zienkiewicz O.C., Campbell J.S.: *Shape optimization and sequential linear programming*, In Optimum Structural Design, Wiley and Sons, 109-126, 1973

Summary

This paper presents the application of numerical methods of analysis, like BEM or FEM, for shape optimization of notch (notches) in elastic machine (structural) components. The design objective is to minimize a maximum effective stress for a given load. A standard and a special concept of segmented Bezier interpolants is adopted for defining geometry of the machine component, and the Sequential Linear Programming is used as optimization procedure.

ZASTOSOWANIE METOD NUMERYCZNYCH DO OPTYMALIZACJI KSZTAŁTU KARBÓW W ELEMENTACH MASZYN I KONSTRUKCJI

Streszczenie

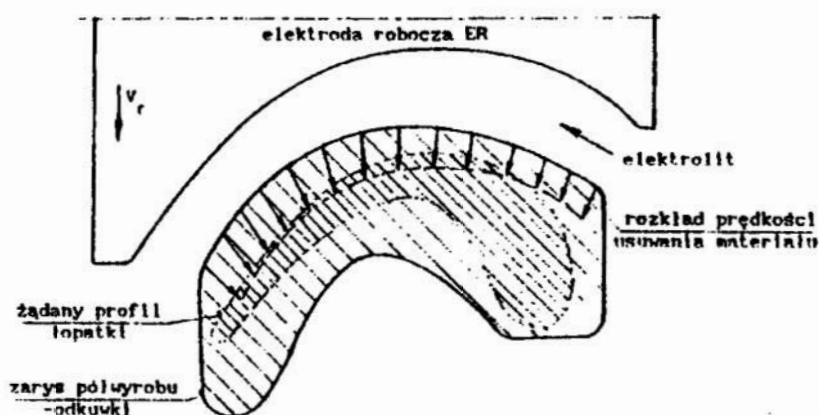
W artykule rozpatruje się zastosowanie numerycznych metod analizy konstrukcji: metody elementów brzegowych lub metody elementów skończonych do poszukiwania optymalnego kształtu brzegu karbu (kARBÓW) w elementach maszyn lub konstrukcji. Celem jest minimalizacja maksymalnych naprężeń zastępczych w otoczeniu karbu dla danego przypadku obciążenia. Do opisu kształtu brzegu wykorzystuje się standardowy i specjalny wariant krzywych Beziera. Zadanie optymalizacji sprowadza się do sekwencyjnego programowania liniowego.

PROJEKTOWANIE KOMPUTEROWE ELEKTROD W OBRÓBCE ELEKTROCHEMICZNEJ (ECM)

W referacie przedstawione zostały zagadnienia związane z korekcją elektrody roboczej oraz modelowaniem obróbki elektrochemicznej. Sformułowany został model matematyczny procesu obróbki oraz zaprezentowano wyniki symulacji komputerowej ECM i weryfikacji doświadczalnej.

1. Wprowadzenie

W obróbce elektrochemicznej (ECM) żądany kształt przedmiotu obrabianego otrzymuje się poprzez roztworzenie elektrochemiczne zbędnego materiału-naddatku, aż do uzyskania wymiarów w granicach wymaganych tolerancji. Jednocześnie dąży się do spełnienia wymagań jakości powierzchni obrabianej, najczęściej w odniesieniu do parametrów jej struktury mikrogeometrycznej (chropowatości).



Rys. 1. Przykład elektrochemicznego kształtowania pióra łopatk

W procesie kształtowania, na powierzchni poddanej roztwarzaniu anodowemu, występuje zmienny rozkład prędkości usuwania materiału (rys.1). Można mówić tu o wektorowym polu prędkości roztwarzania $V(A,t)$, gdzie: A - punkt na anodzie (o współrzędnych x_A, y_A, z_A), t - czas obróbki. Pole $V(A,t)$, określone w danym przedziale czasu $0 < t < t_{\text{obrobki}}$, decyduje o kształcie końcowym przedmiotu, a więc o wyniku obróbki.

Ze względu na istotę elektrochemiczną procesu, $V(A,t)$ zależy od gęstości prądu na anodzie i właściwości elektrochemicznych układu materiał obrabiany-elektrolit. Uzależnienie intensywności usuwania materiału od gęstości prądu, a więc od

^a Instytut Technologii Mechanicznej PW, al. Niepodległości 222, Warszawa

^b NMRC, University of Nebraska-Lincoln, USA

parametrów pola elektrycznego w obszarze, w którym zachodzi proces obróbki, powoduje, że o kształcie obrobionym poza czynnikami "geometrycznymi" decyduje szereg innych. Pole elektryczne uwarunkowane jest bowiem zarówno chwilową geometrią obszaru obróbki (ograniczonego z jednej strony powierzchnią elektrody roboczej, z drugiej w ogólnym przypadku zmienną w czasie powierzchnią anody), jak i warunkami fizykochemicznymi panującymi między elektrodami, zależnymi w szczególności od parametrów obróbki. Przy projektowaniu elektrody roboczej należy jej kształt ściśle uzależnić zarówno od żądanej geometrii przedmiotu obrabianego (PO), jak i od warunków obróbki. Ta cecha ECM powoduje, że przy opracowaniu technologii elektrochemicznej największe efekty uzyskuje się, gdy projektowanie elektrody roboczej (ER) i dobór warunków obróbki rozpatruje się łącznie (jako jedno zadanie).

2. Projektowanie elektrody roboczej i analiza dokładności obróbki

Po wstępnym wyborze warunków obróbki można przystąpić do projektowania elektrody roboczej. Powinno ono obejmować:

- wyznaczenie profilu części roboczej,
- rozmieszczenie wlotów-wylotów elektrolitu (należy mieć na uwadze, że mogą występować zarówno wloty, jak i wyloty, przy podawaniu "zewnętrznym" elektrolitu),
- konstrukcję elektrody (spawana, integralna itd.),
- warunki techniczne (materiał, dokładność, gładkość itp.).

Zasadnicze znaczenie ma przy tym rodzaj stosowanej obróbki, tj. obrabiarka o procesie ciągłym lub obrabiarka impulsowa (PECM). W przypadku procesu ciągłego niezbędne jest wyznaczenie rozkładu grubości szczeliny między PO (powierzchnią obrabianą) a ER, co równoznaczne jest z wyznaczeniem kształtu ER.

Istotne jest również określenie, czy kształt końcowy uzyskuje się w stanie ustalonym (grubość szczeliny nie zmienia się w funkcji czasu) czy przejściowym (ostatni przypadek ma miejsce, gdy grubość naddatków jest mała, nie wystarczająca do wystąpienia stanu ustalonego). Rozwiązanie zadania związanego z poszukiwaniem kształtu ER wiąże się z koniecznością wyznaczenia warunków fizykochemicznych panujących w szczelinie międzyelektrodowej. Warunki te zależą od parametrów ECM.

W związku z powyższym, projektowaniu ER towarzyszy sprawdzanie przyjętych kryteriów ograniczających warunki ECM. W przypadku ich naruszenia należy przeprowadzić korektę przyjętych wartości parametrów ECM. Pozostaje przy tym zagadnienie wyboru "strategii" procedury korekcji, tj. hierarchii parametrów przy ich zmianach, a więc czy np. dokonywać w pierwszej kolejności zmiany wielkości napięcia roboczego U , czy ciśnienia p_{in} , czy też prędkości ruchu posuwowego V_f . Opracowanie takiej procedury ma istotne znaczenie dla wskaźników użytkowych obróbki, a więc np.: uzyskania maksymalnej wydajności, czy też minimalnej energochłonności lub jednoczesnego spełnienia obu warunków. Mamy więc do czynienia z zagadnieniem optymalizacji warunków obróbki.

3. Ewolucja kształtu powierzchni obrabianej

Zmiana kształtu powierzchni obrabianej uwarunkowana jest chwilowymi rozkładami prędkości roztwarzania anodowego. W przypadku ogólnym powierzchni anody $F(x, y, t) = 0$, ewolucja kształtu opisana jest równaniem [1, 2]:

$$\frac{\partial \bar{V}_n}{\partial t} + (\bar{V}_n, \nabla F) = 0, \quad (1)$$

gdzie: $(\bar{V}_n, \nabla F)$ - iloczyn skalarny, \bar{V}_n - wektor prędkości roztwarzania.

Usuwanie nadmiaru roztwarzaniem elektrochemicznym podlega prawom Faradaya, z których wynika, że wielkość prędkości roztwarzania wyraża się zależnością:

$$V_n = \eta \frac{k}{\rho} j_a, \quad (2)$$

gdzie: η - wydajność prądowa procesu roztwarzania anodowego,
 k, ρ - równoważnik elektrochemiczny i gęstość materiału obrabianego,
 j_a - gęstość prądu na anodzie.

Można wyrazić opinię, że najbardziej uzasadnione jest posługiwanie się w technologii elektrochemicznej jedynie wielkością:

$$k_v \stackrel{def}{=} \eta \frac{k}{\rho}, \quad (3)$$

równą objętości materiału roztworzonego przy przepływie jednostkowego ładunku elektrycznego. Dalej k_v będzie określane mianem współczynnika obrabialności elektrochemicznej. Może on być wyznaczony jedynie na drodze eksperymentalnej.

Wartość k_v dla danego materiału zależy nie tylko od rodzaju stosowanego elektrolitu (tj. jego składu, stopnia "zużycia" - mierzonego np. wielkością pH), lecz również od warunków roztwarzania, a mianowicie: gęstości prądu j_a , temperatury elektrolitu T , średniej prędkości przepływu elektrolitu w . Dlatego przy opracowaniu technologii ECM, czy też w badaniach nad ECM, niezbędna jest dla danego układu materiał-elektrolit znajomość funkcji: $k_v = k_v(j_a, T, w)$, przy wyznaczaniu której najbardziej racjonalne jest zastosowanie planowanego eksperymentu.

W celu rozwiązania równania ogólnego (1), lub jego przypadków szczególnych, niezbędna jest znajomość chwilowych rozkładów gęstości prądu na anodzie.

3.1. Pole elektryczne w obszarze obróbki

Z elektrochemicznego punktu widzenia gęstość prądu na anodzie uwarunkowana jest dla danego układu materiał-elektrolit wielkością potencjału anody E_a i może być wyznaczona z krzywej polaryzacyjnej $j_a = j_a(E_a)$, otrzymanej eksperymentalnie dla danych warunków roztwarzania.

To teoretycznie słuszne ujęcie nie może być jednak przy programowaniu zrealizowane "wprost" nawet wówczas, gdy znamy wiarygodną rodzinę krzywych polaryzacyjnych, gdyż potencjał anody nie jest bezpośrednio dany.

Z tych względów niezbędne jest wyznaczenie j_a rozwiązując, metodą kolejnych przybliżeń, równania opisujące pole elektryczne w obszarze międzyelektrodowym:

$$\text{div}[\kappa \text{ grad } u] = 0, \quad (4)$$

gdzie: κ - przewodność elektryczna ośrodka, u - potencjał elektryczny.

Pamiętając, że potencjały anody i katody powiązane są z gęstością prądu na elektrodach krzywymi polaryzacyjnymi, warunki brzegowe na elektrodach są nieliniowymi warunkami III rodzaju (tj. potencjał brzegu jest funkcją gradientu potencjału). W niektórych układach materiał-elektrolit, nawet przy znacznych gęstościach prądu, polaryzacja elektrod może być opisana równaniem Tafela:

$$E_{a,k} = a_{a,k} + b_{a,k} \ln j_{a,k}. \quad (5)$$

Równanie to dotyczy zwłaszcza procesu katodowego, gdzie jest ściśle w bardzo szerokim przedziale: $j_k = 10^{-9} \div 10^2 \text{ A/cm}^2$ [1].

Metoda kolejnych przybliżeń jest następująca:

–w I przybliżeniu zakłada się, że potencjały anody i katody są stałe, równe wielkościom średnim:

$$u^{(1)}(A) = U - \bar{E}_a, \quad u^{(1)}(K) = U - \bar{E}_k, \quad (6)$$

–rozwiązuje się równanie (4) i wyznacza się gęstość prądu na elektrodach,

–w II przybliżeniu koryguje się i uzmiennia potencjały na brzegach na podstawie krzywych polaryzacyjnych (z tych względów dla potrzeb ECM celowe jest przedstawienie krzywych polaryzacyjnych w układzie współrzędnych odwróconych, tzn. przy aproksymacji krzywych zmienną zależną powinien być potencjał elektrody, a niezależną gęstość prądu),

–rozwiązuje się ponownie równanie (4) dla zmienionych uprzednio warunków brzegowych.

Cykl obliczeń powtarza się, aż do osiągnięcia przyjętego kryterium dokładności czy to w odniesieniu do potencjału, np. $\sup |u^{(k+1)} - u^{(k)}| < \varepsilon_u$, czy też gęstości prądu na anodzie $\sup |j_a^{(k+1)} - j_a^{(k)}| < \varepsilon_j$, gdzie: k - nr iteracji.

Praktyka metod numerycznych, stosowanych w elektrochemii do obliczeń pól elektrycznych w elektrolitach wykazuje, że opisana metoda iteracyjna jest zbieżna. W obróbce ECM istotną cechą jest, że wielkości E_a i E_k są najczęściej wielokrotnie mniejsze od omowego spadku potencjału (czy też od wielkości stosowanego napięcia międzyelektrodowego) oraz, że zmieniają się stosunkowo niewiele wzdłuż powierzchni elektrod. Niezbędna liczba iteracji podczas obliczeń jest niewielka i często nie przekracza liczby 3.

3.2. Zmiany warunków w obszarze obróbki

Powyższe zadanie komplikuje fakt, że warunki fizykochemiczne panujące między elektrodami są zmienne. Nagrzewanie elektrolitu ciepłem, głównie Joule'a, oraz wydzielanie się gazu na elektrodach powoduje powstawanie nierównomiernego pola temperatury T i koncentracji objętościowej gazu β , na skutek czego ośrodek staje się dwufazowy o zmiennej przewodności elektrycznej κ . Z tych względów rozwiązaniu równania (4) towarzyszy konieczność znalezienia rozkładów T i β .

Wymiana ciepła oraz wydzielanie się fazy gazowej odbywa się w warunkach przepływającego elektrolitu. W związku z tym pojawia się kolejne zadanie, a mianowicie obliczenie rozkładów prędkości przepływu elektrolitu w i ciśnienia statycznego p w obszarze międzyelektrodowym (to ostatnie decyduje o koncentracji objętościowej gazu β).

W warunkach ECM przepływający elektrolit jest najczęściej wielofazowy a więc ściśliwy i ma zmienną lepkość (wpływ T i β). W szczególnych przypadkach mogą wystąpić zjawiska przepływu krytycznego oraz kawitacja. W zależności od rodzaju regulacji układu obiegu elektrolitu zadawane są ciśnienia na wlocie i wylocie lub wydatek elektrolitu.

3.3. Uwagi dodatkowe

Wymienione sprzężone zadania, w których wyznacza się rozkłady w , p , T , β , j , k_v , nawet w przypadku prowadzenia obliczeń dla stanu ustalonego, rozwiązuje się w zmiennym w czasie obszarze, z tym że w odniesieniu do procesu ciągłego ECM czas spełnia rolę parametru (tj. lokalne pochodne po czasie można pominąć). Z tych względów można "zamrozić" geometrię na okres poszukiwania wymienionych pól.

Niezależnie od rodzaju zadania przy jego rozwiązywaniu występuje konieczność wyznaczenia ewolucji kształtu anody i warunków fizycznych panujących w obszarze między elektrodami. Jest to równoznaczne z symulacją procesu ECM.

W zadaniu dotyczącym doboru parametrów obróbki występują kryteria odnoszące się do warunków fizycznych (T_{max} , β_{max} , p_{min} , w_{max}) oraz wskaźników użytkowych obróbki (dokładność, wydajność, energochłonność).

W zadaniu dotyczącym projektowania ER stosowana jest metoda prób i korekcji ER na podstawie wyników uzyskanych z modelowania matematycznego zmian kształtu anody i porównania tak uzyskanego kształtu z zadaniem.

4. Projektowanie elektrody roboczej

Dla stanu ustalonego, przy wyznaczaniu kształtu ER, najczęściej był wykorzystywany warunek, wg którego rozkład gęstości prądu na stacjonarnej anodzie opisany jest zależnością:

$$j = \frac{(\vec{V}_f, \vec{n}_F)}{k_v} \quad (7)$$

gdzie: (\vec{V}_f, \vec{n}_F) -iloczyn skalarny, \vec{n}_F -wektor normalny do żądanej powierzchni anody.

Przy bezpośrednim stosowaniu powyższego warunku w odwrotnych zagadnieniach brzegowych ECM [3, 4, 5, 6] powstaje kwestia jednoznaczności rozwiązania, a mianowicie nielato odpowiedzieć na pytanie, czy otrzymane rozwiązanie jest jedynie prawdziwe. Przy poszukiwaniu ścisłego rozwiązania główna trudność związana jest z warunkiem początkowo-brzegowym, bowiem nie znana jest a priori początkowa grubość szczeliny na wlocie (lub w innym punkcie, który można uznać za początkowy, np. przy ECM-RE jest nim wylot). Aby uniknąć powyższych trudności, zastosowano metodę korekcji wykorzystując symulację procesu kształtowania.

4.1. Wyznaczanie geometrii elektrody roboczej metodą prób i korekcji

W praktyce ECM przy wykonywaniu elektrody roboczej często stosuje się metodę kolejnych korekcji. Polega ona na tym, że po wykonaniu elektrody roboczej o kształcie przybliżonym (np. stosując "prawo cosinusa" [1, 7]) przeprowadza się próbne drażnienie, po czym dokonuje się pomiarów geometrii otrzymanej powierzchni, niezbędnych do

wyznaczenia rozkładu jej odchyżeń od żadanego kształtu. Następnie poprawiany jest kształt elektrody przez usunięcie nadmiaru lub naniesienie metalu w zależności od znaku i rozkładu odchyłek. Po takiej korekcji ER znów przeprowadza się próbę drążenia; operacje są powtarzane aż do uzyskania obrobionego kształtu w zadanym polu tolerancji. Procedura powyższa jest pracochłonna i kosztowna, ale jej zasadę można wykorzystać przy projektowaniu.

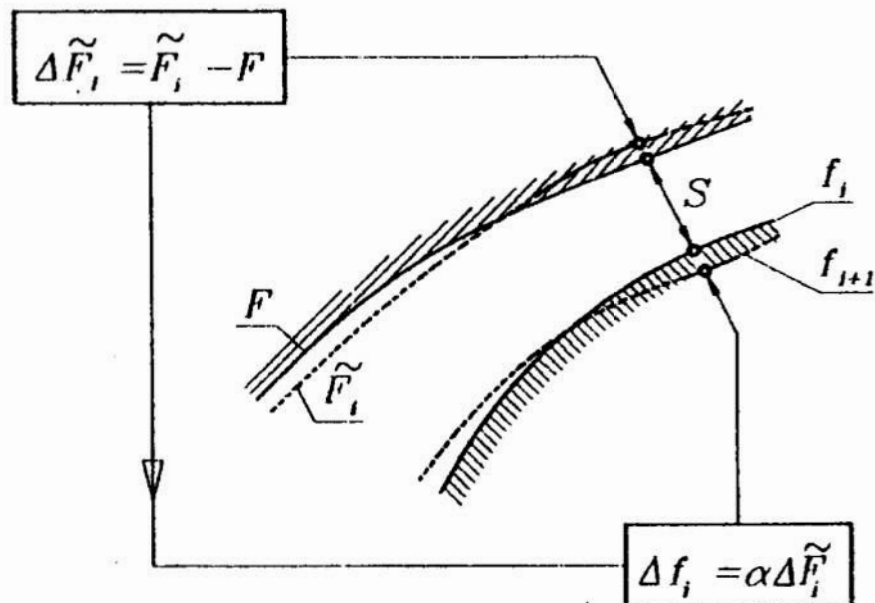
Po wyznaczeniu i -tego przybliżenia kształtu ER (przy czym dla $i = 1$ można założyć stałość κ , k_w , E_a , E_k , a przy określaniu ER zastosować metodę opisaną w [1, 7]) należy przeprowadzić symulację komputerową kształtowania elektrochemicznego. W wyniku uzyskany zostanie "obrobiony" elektrodą f_i kształt końcowy \tilde{F}_i (w stanie ustalonym) oraz rozkład odchyłek kształtu \tilde{F}_i od kształtu żadanego: $\Delta F_i = \tilde{F}_i - F$. Następnie poprawia się elektrodę roboczą f_i , przesuwając punkty jej zarysu, np. w kierunku grubości szczeliny o odcinki (rys.2):

$$\Delta f_i = \alpha \Delta \tilde{F}_i, \quad (8)$$

gdzie: $0 < \alpha < 1$ - współczynnik warunkujący szybkość zbieżności procesu iteracji.

Procedurę powtarza się do momentu uzyskania kryterium dokładności rozwiązania.

Przy symulacji kształtowania ECM w celu wyznaczenia ustalonego kształtu \tilde{F}_i odpowiadającej danej f_i przyjmuje się, że początkowym kształtem anody jest kształt zadany F , co znacznie upraszcza obliczenia.

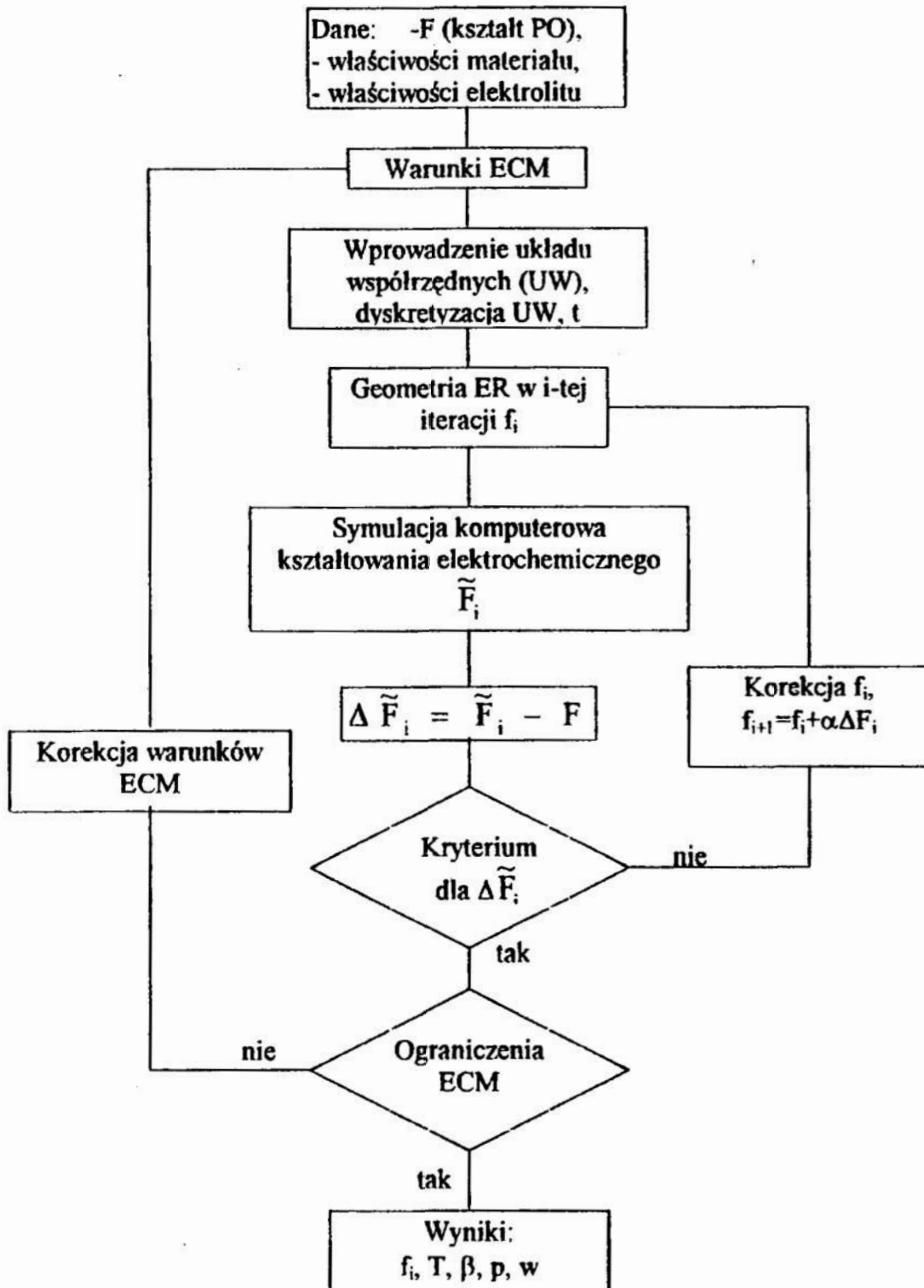


Rys.2. Schemat korekcji elektrody roboczej

4.2. Algorytm projektowania elektrody roboczej

Jak zaznaczono uprzednio, symulację komputerową kształtowania ECM i korekcję ER można prowadzić w układzie współrzędnych krzywoliniowych lub prostokątnych związanych z ruchomą elektrodą. Przy rozwiązywaniu układu równań różniczkowych określających p , w , T , β , F stosuje się metodę różnic skończonych. Schemat algorytmu projektowania przedstawiony jest na rys.3.

Symulacja kształtowania w i -tym cyklu iteracyjnym prowadzona jest aż do ustalenia się \tilde{F}_i , natomiast proces iteracyjny wyznaczania ER zostaje zakończony po spełnieniu kryterium odnośnie $\Delta\tilde{F}_i$, np. w postaci $\sup |\tilde{F}_i - F| < \varepsilon_F$. Jednocześnie powinny być spełnione ograniczenia nakładane na warunki realizacji obróbki, np. w postaci $T_{max} < T_{dop}$, $\beta_{max} < \beta_{dop}$, $w < w_{dop}$ i innych. Należy zauważyć, że wyznaczony kształt ER zależy nie tylko od żądanej geometrii powierzchni obrabianej, lecz również od warunków obróbki. W związku z tym zadania projektowania ER i wyboru warunków obróbki (elektrolitu, parametrów) są ze sobą ściśle związane.



Rys.3. Algorytm projektowania ER metodą korekcji

Podana metodyka projektowania ER stanowiła podstawę do opracowania programu na IBM PC w oparciu o przedstawiony poniżej model matematyczny:

$$\begin{aligned}
 \xi &= \int_{x_0}^L \sqrt{1+f'^2} dx, \\
 \alpha(\xi) &= \operatorname{arctg} f'(x) \\
 r(\xi) &= \frac{(1+f'^2)^{3/2}}{f' f''}, \\
 \frac{\partial S}{\partial a} &= \frac{k_v j_A - V_f \cos(\alpha + \delta)}{\cos \delta}, \\
 S(\xi, 0) &= S_0(\xi), \\
 j &= \kappa_0 (1 + \alpha_T (T - T_0)) (1 - \beta)^{3/2} \frac{U - E}{S}, \\
 \delta &= \operatorname{arctg} \left(\frac{1}{1 - S/r} \frac{\partial S}{\partial \xi} \right), \\
 \frac{\partial T}{\partial \xi} &= \frac{Uj}{w_0 \rho_e (1 - \beta_0) C_p S_0}, \\
 T(0, t) &= T_0, \quad S(0, t) = S_0, \quad w(0, t) = w_0, \\
 \frac{\partial \left(\beta \frac{273}{T} \frac{p}{10^5} S w \right)}{\partial \xi} &= \eta_H k \\
 \beta(0, t) &= \beta_0, \\
 \frac{\partial p}{\partial \xi} &= - \frac{C}{S \operatorname{Re}^n} \frac{\rho w^2}{4} - \frac{1}{S} \frac{\partial (\rho w^2 S)}{\partial \xi}, \\
 p(0, t) &= p_{in} - \zeta_i \frac{\rho_e w_0^2}{2}, \\
 p(\zeta_L, t) &= p_{out} + \zeta_0 \frac{\rho_e (1 - \beta) w^2}{2}, \\
 \operatorname{Re} &= \frac{2 S w \rho}{\mu} = \frac{2 S_0 w_0 \rho (1 - \beta_0)}{(1 - \beta) \mu}, \\
 \rho &= \rho_0 (1 - \beta), \\
 \mu &= \mu_0 (1 - \mu \beta) \exp(-b(T - T_0)).
 \end{aligned} \tag{9}$$

Wynikiem rozwiązania powyższego układu równań (9) są rozkłady grubości szczelin, temperatur, ciśnień, prędkości oraz koncentracji objętościowej fazy gazowej.

5. Weryfikacja doświadczalna

Wyznaczenie stopnia zgodności wyników uzyskiwanych z programu, z wynikami rzeczywistego procesu ECM może być dokonane jedynie na drodze doświadczalnej.

Badania prowadzono na zmodyfikowanej obrabiarce elektrochemicznej *DEL-2B* o następujących danych technicznych:

- napięcie robocze	do 30 V
- natężenie prądu roboczego	do 6000 A
- prędkość posuwu roboczego (regulacja bezstopniowa)	0.3+6.5 mm/min
- posuw skokowy (wg zadanego programu)	2 mm/impuls
- częstotliwość impulsów posuwowych	do 50 Hz
- wymiary komory roboczej	600x400x300 mm
- ciśnienie elektrolitu	do 3 MPa
- pojemność zbiornika elektrolitu	1 m ³

Stanowisko badawcze wyposażone zostało w tor pomiarowy, umożliwiający zapis natężenia prądu, napięcia, prędkości posuwu roboczego, temperatury i ciśnienia elektrolitu na wlocie i wylocie ze szczeliny międzyelektrodowej. Bardziej szczegółowy opis stanowiska *DEL-2B* podany jest m.in. w pracy [8].

Próbki do obróbki elektrochemicznej wykonane zostały ze stali *WNL*, o składzie chemicznym: 0.55% C, 0.7% Mn, 0.2% Si, 0.7% Cr, 1.6% Ni, 0.25% Mo, maksimum 0.03% S (reszta Fe). Wybór stali *WNL* można uzasadnić stosowaniem jej na matryce obrabiane elektrochemicznie (m.in. w FSC Lublin, PZL Rzeszów). Obróbkę prowadzono w roztworze wodnym azotanu sodu (NaNO_3) o stężeniu masowym 12%. Wybór elektrolitu uzasadniony był szerokim zastosowaniem roztworu NaNO_3 w obróbce elektrochemicznej.

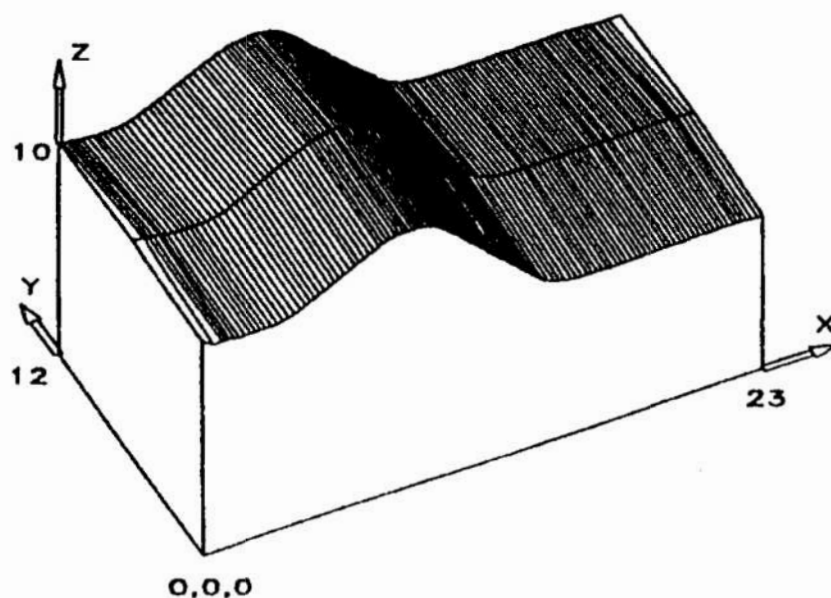
5.1. Weryfikacja komputerowego projektowania elektrody roboczej

Weryfikacji projektowania elektrody kształtowej dokonano na podstawie porównania profili kształtów obrobionych elektrodą zaprojektowaną, z zadanym profilem, dla którego tę ER zaprojektowano. Badania przeprowadzono przy użyciu elektrody walcowej, przy czym dla stworzenia "gorszych" warunków obróbki nie stosowano dławienia elektrolitu na wylocie ze szczeliny.

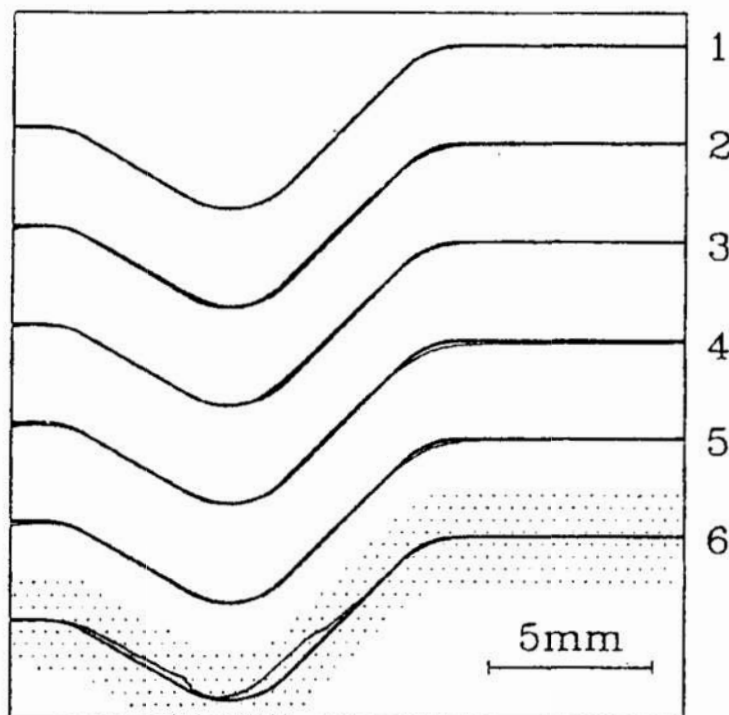
Na rys.4 pokazano przykład kształtu powierzchni otrzymanej w wyniku obróbki elektrodą walcową. Każdą powierzchnię mierzono na maszynie pomiarowej, uzyskując dla różnych przekrojów zbiory współrzędnych jej punktów, które zapisywano na dysku PC, a następnie poddawano analizie, w której wykorzystywano programy profesjonalne. Analiza sprowadzała się do porównywania kształtu uzyskiwanego z eksperymentu z kształtem zadanym dla zaprojektowania ER. Wyniki porównania wyprowadzono w postaci profilów, których przykładowy zbiór przedstawia rys.5 oraz wykresów rozkładów odchylek powierzchni otrzymanej od wymaganej (teoretycznej), mierzonych

wzdłuż normalnych do niej. W każdym przypadku zestawienia krzywych dokonywano powiększeń tych fragmentów, których odchyłki były największe, jak to zilustrowano na rys.6. Typowy rozkład odchyłek, dla przypadku gdy były one duże, przedstawia rys.7.

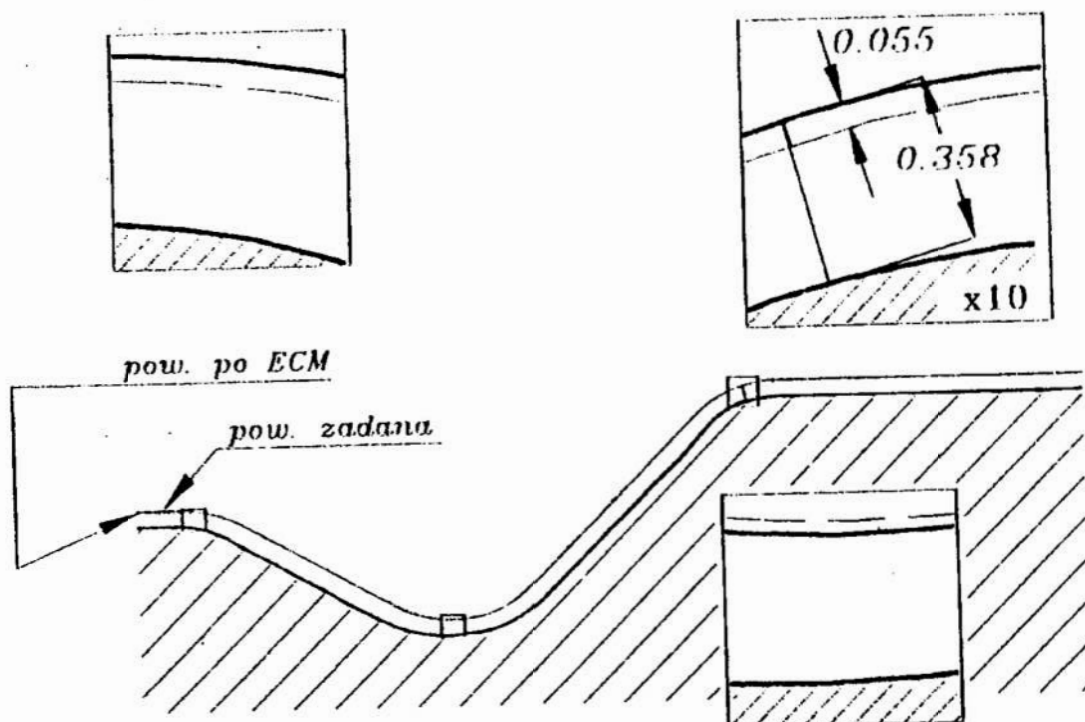
Z rozkładu odchyłek (rys.7) między profilami otrzymanymi z badań i zadanymi wynika, że maksymalne występują na odcinkach będących łukami okręgów (o promieniu $r = 2.5$ mm). Lokalne grubości szczeliny w tych miejscach były w granicach $S = 0.35 \div 0.4$ mm, co odpowiada $S/r = 0.15$. Powyższe wielkości S/r są więc wielokrotnie większe od tych, przy których dopuszczalna jest linearyzacja rozkładu potencjału wzdłuż S .



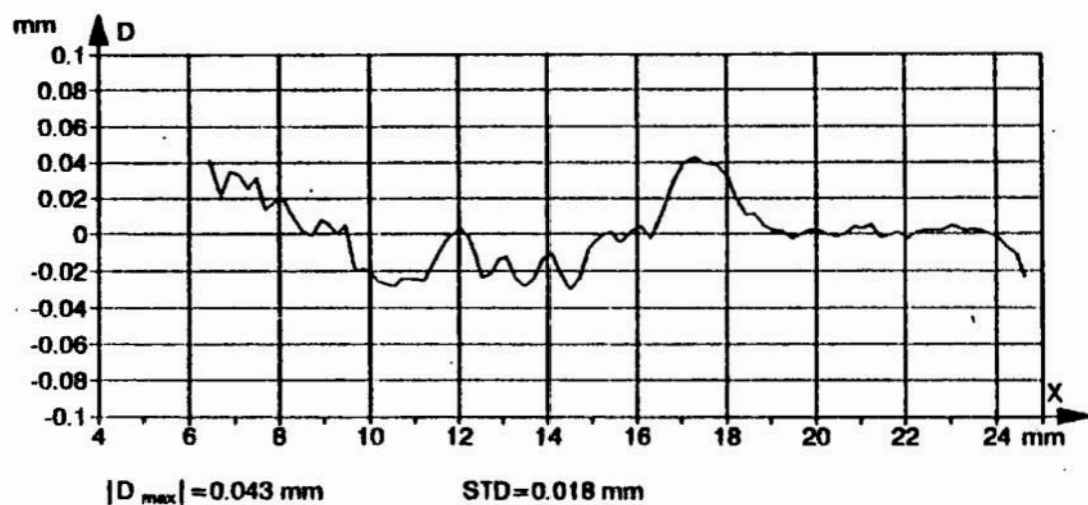
Rys.4. Kształt powierzchni po obróbce elektroda walcowa



Rys.5. Profile zadane (linia gruba) i otrzymane po ECM (linia cienka)



Rys.6. Powiększony profil nr 1 z rys.5



Rys.7. Rozkład odchylek normalnych do powierzchni dla profilu nr 1 z rys.5

Odchyłki kształtu na odcinkach powierzchni pochylonych i o małej krzywiznie mieściły się w granicach niedokładności maszyny pomiarowej, wynoszącej $\pm 5 \mu\text{m}$ (rys.7).

Na rys.5 profil nr 6, który otrzymano w innych warunkach, uwidacznia jakie mogą być skutki przy prowadzeniu obróbki z użyciem innych parametrów niż nominalne (przyjęte do obliczeń ER). Widoczne są duże rozbieżności kształtu obrobionego od zadanego, a ponadto podczas obróbki występowały zakłócenia stabilności procesu.

Reasumując można stwierdzić: weryfikacja dla przypadku obróbki elektrodą walcową wykazała, że niedokładność powierzchni obrobionej zaprojektowaną elektrodą nie przekraczała 0.04 mm, przy czym największe odchyłki występowały lokalnie jedynie na odcinkach dużej krzywizny (większej od zalecanej ze względu na błąd linearyzacji rozkładu potencjału). Na pozostałych fragmentach, w tym zakrzywionych, niedokładność mieściła się w granicach niedokładności maszyny pomiarowej. Podobnie jak w poprzednim przypadku należy mieć na uwadze niedokładność wykonania elektrody na obrabiarce (pomiar wykazały, że jest mniejsza od 0.021 mm).

Pracę wykonano częściowo w ramach projektu badawczego Nr 0256/P4/93/05 finansowanego przez Komitet Badań Naukowych.

LITERATURA

- [1] Davydov A. D., Kozak J. *Wysokoskorostnoje elektrochimičeskoje formoobrazowanie*, Nauka, Moskwa, 1990;
- [2] Dąbrowski L., *Podstawy komputerowej symulacji kształtowania elektrochemicznego*, Prace Naukowe PW, Mechanika, z.154, Warszawa 1992.
- [3] Kozak J.: *Geometria elektrod w procesie kształtowania elektrochemicznego*. Praca doktorska, Politechnika Warszawska, Warszawa 1969.
- [4] Krylov A.L.: *Zadača Cauchy dlja uravnenija Laplace v teorii elektrochimičeskoj obrabotki metallov*. Dokl. AN ZSRR, t. 178, No. 2, 1968
- [5] Narayanan O.H., Hinduja S., Noble C.F.: *Design of Tools for ECM by Boundary Element Method*. Proc Inst. Mech. Eng., 200(C2) 1986, s.195÷205.
- [6] Zhu D., Yu C.Y.: *CAD of ECM Tool Shape*. Proceed. Computer-Aided Production Enginnering, Edinburg 1986, s.397÷400.
- [7] Tipton H.: *The Determination of Tool Shape for ECM*. Machinery and Production Eng., 14, No. 2, 1968, s.325÷328.
- [8] *Podstawy teoretyczne obróbki elektrochemicznej oraz komputerowe wspomaganie projektowania technologii ECM*. Praca zbiorowa. Sprawozdanie ITM PW z tematu CPBP 02.04-13 dla Politechniki Wrocławskiej, Wrocław 1990.

COMPUTER CALCULATING THE WORKING TOOL OF THE ELECTROCHEMICAL MACHINING

Summary

The problems of shape correction of the working electrode and computer modeling of the electrochemical machining are described in this paper. The mathematical model, results of computer simulation and experimental verification are presented.

ANALIZA ZMIAN STANU NAPRĘŻEŃ WŁASNYCH I INNYCH CECH EKSPLOATACYJNYCH WARSTWY WIERZCHNIEJ PODDANEJ PROCESOWI ZUŻYCIA

Omówiono określenie sposobu wyznaczania funkcji obiektu badań określających wybrane własności eksploatacyjno - użytkowe badanych warstw wierzchnich konstytuowanych na drodze obróbki nagniataniem naporowym tocznym. Badania przeprowadzono dla dwu i trójwartościowego zbioru wielkości wejściowych - parametrów obróbki wykorzystując plany badań kompletny ($PS/DK - 3^2$) i Hartleya ($PS/DS - P:Ha_3 - hS$). Otrzymane wyniki pozwalają na celowe i świadome sterowanie parametrami obróbki w celu uzyskiwania pożądanych cech użytkowych otrzymywanych powierzchni.

1. Wprowadzenie

W trakcie prowadzenia badań doświadczalnych można wyróżnić [1, 2] etap przygotowania pomiarów, to znaczy planowanie badań, etap realizacji badań, oraz analizę statystyczną i merytoryczną uzyskanych wyników pod kątem statystycznym i merytorycznym. Przystępując do badań określa się obiekt badań za pomocą właściwych dla niego wielkości, przy czym między tymi wielkościami mogą występować zależności o charakterze przyczynowo - skutkowym, których opis matematyczny można nazwać modelem matematycznym obiektu badań. Wśród wielkości opisujących badany obiekt wybiera się wielkości wejściowe (wielkości niezależne, czynniki badane), których wartości można dobierać tworząc plan doświadczenia oraz wielkości wyjściowe (wielkości wynikowe, badane) i wielkości stałe (parametry które w trakcie badań celowo pozostawia się niezmiennione). Ponadto w trakcie badań mogą wystąpić wielkości zakłócające, których wpływ należy minimalizować. Następnym krokiem jest przygotowanie planu przygotowywanych badań. Na tym etapie trzeba zdecydować o ilości wielkości wejściowych (co może zadecydować o niepowodzeniu całych badań w wypadku pominięcia któreś z istotnych wielkości), o liczbie wartości wielkości wejściowych (co z kolei decyduje o postaci możliwego do uzyskania modelu matematycznego) oraz o przedziałach badanych wartości wielkości wejściowych, tak by dla wszystkich skojarzeń wartości wejściowych możliwe było przeprowadzenie pomiarów. W tym punkcie badań podejmuje się decyzję o wyborze najwłaściwszego planu badań (jednego z wielu proponowanych przez teorię badań naukowych) charakteryzującego się właściwą ilością wielkości wejściowych i ilością wartości wielkości wejściowych oraz możliwą do zaakceptowania ilością układów planu. Plan wybrany powinien się cechować maksymalną informatywnością i efektywnością. Kolejnym etapem badań jest przeprowadzenie zaplanowanych badań doświadczalnych. Uzyskuje się wyniki pomiarów o liczbie równej ilości układów planu badań, przy czym każdy z wyników pomiarów wykonany jest z jednym lub wieloma powtórzeniami. Umożliwia to wyznaczenie miar statystycznych wielkości wyjściowej. Kolejnym etapem badań jest przeprowadzenie analizy statystycznej otrzymanych wyników. Następnie

* Katedra Technologii Budowy Maszyn. Politechnika Częstochowska
42 - 200 Częstochowa ul. Armii Krajowej 21

opracowane wyniki badań można wykorzystać do wykonania aproksymacji otrzymanych wyników funkcją obiektu badań. Po wykonaniu badań adekwatności otrzymanej funkcji, wykonaniu obliczeń istotności poszczególnych wielkości wejściowych i przeprowadzeniu analizy merytorycznej badanego zjawiska możliwe jest potraktowanie otrzymanej funkcji jako modelu matematycznego badanego zjawiska. Model taki może posłużyć do przeprowadzenia dalszych badań rozpatrywanego obiektu (np. przeprowadzenie optymalizacji, komputerowej symulacji badanego obiektu, czy też badanie wpływu zmian poszczególnych czynników badanych na rozpatrywany obiekt badań). W omawianych badaniach obiektem badań doświadczalnych były własności fizykochemiczne, takie jak: stan, rozkład i zmiany w stanie naprężeń wewnętrznych w warstwie wierzchniej, parametry stereometrii powierzchni R_a ; n ; H_{max} ; H_{min} oraz t_p , stan i zmiany w głębokości i wartości umocnienia warstwy wierzchniej wyrobów nagniatanych poddanej procesowi zużywania. W przedstawionym opracowaniu ograniczono się do omówienia badań dotyczących stanu naprężeń własnych w warstwie wierzchniej.

2. Badania własne

2.1. Przygotowanie badań

Badaniom poddano dwie grupy powierzchni stali 55: nagniatane naporowo tocznie kulką i nagniatane naporowo tocznie krążkiem. Ponadto w celach porównawczych dokonano badań powierzchni toczonych i szlifowanych. Parametry toczenia wykańczającego poprzedzającego nagniatanie $n = 800$ obr/min, $v = 100,5$ m/min, $f = 0,102$ mm/obr. Nagniatanie przeprowadzono narzędziem umożliwiającym sprężysty docisk elementu nagniatającego (kulki, krążka) do powierzchni nagniatanej z możliwością regulacji siły docisku.

Analiza teoretyczna badanego obiektu pozwoliła założyć przy badaniach powierzchni nagniatanych kulką badania wpływu dwóch wielkości wejściowych (promień kulki R_k i siła nacisku narzędzia nagniatającego F) na przebieg i uzyskiwane efekty zużycia. Dla tak założonego obszaru badań przyjęto statyczny zdeterminowany kompletny plan doświadczenia ($PS/DK - 3^2$) o $n_k = \text{const} = 3$. Plan ten przy stosunkowo niewielkiej ilości układów $n = 9$ zapewnia maksymalną informatywność i wysoką efektywność prowadzonych badań. Parametry zmienne przy nagniataniu kulką zestawiono w tab.1. Pozostałe parametry nagniatania były stałe i wynosiły $v = 10$ m/min i $f = 0,102$ mm/obr.

Tab.1. Parametry zmienne nagniatania kulką. (plan kompletny $PS/DK - 3^2$)

Układ planu	1	2	3	4	5	6	7	8	9
Sila nagniatania [kN]	0.5	2.75	5	0.5	2.75	5	0.5	2.75	5
Promień kulki [mm]	5	5	5	13.5	13.5	13.5	9.5	9.5	9.5

Drugą grupę powierzchni stanowiły powierzchnie nagniatane krążkiem. Dla powierzchni nagniatanych krążkiem na podstawie analizy obiektu badań przyjęto za

zmienne trzy parametry obróbki - trzy wielkości wejściowe (promień krzywizny krążka R_k , siła docisku narzędzia F i posuw f). Pozostałe parametry nagniatania były stałe w trakcie badań i wynosiły $v = 10$ m/min średnica krążka nagniatającego $d = 39$ mm. W badaniach wykorzystano statyczny zdeterminowany poliselekcyjny specjalny plan doświadczeń Hartleya ($PS/DS - P:Ha_3$) na hipersześcianie (hS). Wybór planu podyktowany był jego szczególnie dobrą efektywnością połączoną z dość dobrą informatywnością. Wybór planu opartego na hipersześcianie (hS) a nie bardziej informatywnego opartego na hiperkuli (hK) podyktowany był dążeniem do zmniejszenia do minimum liczby wielkości wejściowych. Pozwoliło to zmniejszyć liczbę narzędzi przygotowanych do przeprowadzenia badań. Parametry zmienne w trakcie nagniatania krążkiem zestawiono w tab.2.

Tab.2. Parametry zmienne nagniatania krążkami (plan Hartleya $PS/DS - P:Ha_3 - hS$)

Układ planu	1	2	3	4	5	6	7	8	9	10	11
Promień krzywizny zarysu [mm]	5	15	5	15	5	15	10	10	10	10	10
Posuw [mm/obr]	0.07	0.07	0.45	0.45	0.26	0.26	0.07	0.45	0.26	0.26	0.26
Siła nagniatania [kN]	5	0.5	0.5	5	2.75	2.75	2.75	2.75	0.5	5	2.75

Przedziały zmienności wartości wejściowych dobrano w oparciu o ogólne zasady doboru parametrów obróbki nagniataniem w ten sposób że dowolne ich skojarzenie było możliwe do realizacji i sensowne z punktu widzenia prowadzonych badań. Na etapie planowania badań zbiory wartości wejściowych [x_{min} , x_{max}] zostały zastąpione przedziałami bezwymiarowych wartości unormowanych [-1, 1]. Taki sposób podejścia znacznie ułatwił planowanie badań. Wartości wielkości wejściowych były dla przyjętych planów badań równe -1, 0 lub 1 odpowiadających rzeczywistym wartościom wielkości wejściowych podanych w tab.1 i tab.2. Dzięki takiemu uniwersalnemu podejściu możliwe było bezpośrednie korzystanie z gotowych planów badań dostępnych w literaturze [1, 2]. Wykonywanie wszystkich obliczeń i przekształceń dla wartości normowanych znacznie ułatwia również sam proces dalszej obróbki otrzymanych wyników doświadczenia. Relacje normowania zestawiono w tab.3.

2.2. Przeprowadzenie badań doświadczalnych

Pomiary stanu naprężeń [3] dokonano metodą dyfraktografii na dyfraktometrze rentgenowskim typu *DRON-2* (lampa chromowa *BSW21 Cr*; $U = 35$ kV; $I = 5$ mA). Pomiaru naprężeń dokonano metodą krokową mierząc natężenie promieniowania w odstępach $2\theta = 0,5^\circ$. Przeprowadzona analiza naprężeń metodą dyfraktografii rentgenowskiej umożliwiła wyznaczenie przede wszystkim naprężeń I rodzaju (makronaprężeń). Analiza naprężeń objęła wyznaczenie sumy naprężeń głównych (przy osi próbek ustawionej pionowo i licznika obiegającego próbkę w płaszczyźnie prostopadłej do osi próbek i przy osi próbek ustawionej poziomo i liczniku obiegającym próbkę w płaszczyźnie osi próbek) oraz wyznaczenie naprężeń w wybranych kierunkach: wzdłuż obwodu próbki i wzdłuż osi próbki. Głębokość z której uzyskiwano informację wynosiła około 20 μ m licząc od powierzchni materiału.

W celu prześledzenia zmian stanu naprężeń w trakcie eksploatacji w warstwie wierzchniej wyrobów nagniatanych badane powierzchnie poddano [4, 5] zużyciu (w warunkach tarcia suchego) na drodze zużycia 9,5 tys. metrów. Badania te przeprowadzono na stanowisku do badań zużycia wykonanym w oparciu o tokarkę uniwersalną CU - 500.

2.3. Analiza statystyczna otrzymanych wyników pomiarów

Ze względu na znaczną pracochłonność każdego z przeprowadzanych pomiarów oraz znaczny koszt wyznaczania naprężeń wykonywanych poza naszym ośrodkiem badawczym zdecydowano się wykonać powtórzenia tylko dla części z przeprowadzonych pomiarów (układów wartości wielkości wejściowych). Na tej podstawie wyznaczono wariancję tej części wyników pomiarów naprężeń $S^2(z) = 1,022 \cdot 10^3$ MPa. Wartość tę przyjęto do dalszych obliczeń zakładając że jest ona zbliżona dla wszystkich wykonanych pomiarów.

2.4. Wyznaczenie funkcji obiektu badanego

Dokonując wyboru funkcji aproksymującej trzeba mieć na uwadze by otrzymana funkcja obiektu badań w możliwie prostej postaci opisywała badany obiekt. Wybrana funkcja musi zapewnić wystarczającą zgodność (adekwatność) otrzymanej funkcji z uzyskanymi wynikami badań. Dla opisywanych tu badań powierzchni nagniatanych dokonano aproksymacji wyników otrzymanych pomiarów naprężeń wielomianami algebraicznymi drugiego stopnia (ze składnikami liniowymi, kwadratowymi i interakcyjnymi). Wybór takiej postaci funkcji aproksymującej podyktowany został jej stosunkowo prostą postacią, łatwością dokonania aproksymacji i łatwością dokonania późniejszej obróbki matematycznej otrzymanych funkcji. Wybór stopnia wielomianu był kompromisem pomiędzy dążeniem do jak najlepszej zgodności otrzymanej funkcji z otrzymanymi wynikami, a maksymalną efektywnością badań. Zwiększenie stopnia wielomianu musiało by zaowocować koniecznością przyjmowania planów badań o większej ilości układów planu. Liczba występujących w wielomianie współczynników nie może być bowiem większa od ilości układów przyjętego planu doświadczenia.

Przy dokonaniu aproksymacji posłużono się najczęściej chyba stosowaną metodą najmniejszych kwadratów. Obliczenia wykonano z zastosowaniem rachunku macierzowego. Obliczenia rozpoczęto od wyznaczenia macierzy planu doświadczenia X typu $n \times N_b$ (n - liczba układów planu, N_b - liczba współczynników funkcji aproksymującej) i wyznaczenia macierzy transponowanej X^T typu $N_b \times n$ oraz macierzy diagonalnej ilości poszczególnych pomiarów R typu $n \times n$. Pozwala to wyznaczyć macierz informacyjną M typu $N_b \times N_b$, $M = (X^T R) X$. Jeżeli macierz M nie jest macierzą osobliwą ($\det M \neq 0$) to można wyznaczyć macierz odwrotną M^{-1} . Pozwala to wyznaczyć macierz aproksymacji $A = M^{-1} (X^T R)$. Po zestawieniu uzyskanych wyników pomiarów w postaci wektora kolumnowego - macierzy Z typu $n \times 1$ uzyskuje się wynik przeprowadzonej aproksymacji w postaci wektora kolumnowego współczynników funkcji aproksymującej $B = AZ$ typu $N_b \times 1$.

Po dokonaniu aproksymacji uzyskano dla powierzchni nagniatanych kulką funkcję obiektu badań $N = f(x, y)$, to znaczy zależność badanych w warstwie wierzchniej naprężeń w zależności od wielkości wejściowych (siły nagniatania F i promienia kulki R_k). Podobnie dla przyjętego trójwymiarowego obszaru badań powierzchni

nagniatanych krążkiem uzyskano funkcje obiektu badań $N = f(x, y, z)$, to znaczy zależność badanych w warstwie wierzchniej naprężeń w zależności od wielkość wejściowych (promienia krzywizny krążka R_k , siły nagniatania F i posuwu f).

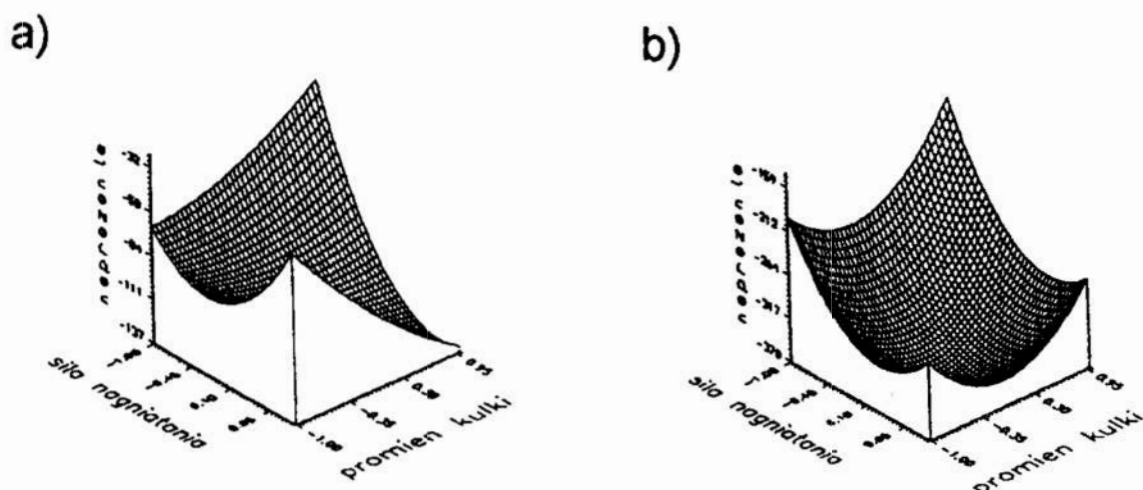
Funkcje wyznaczone zostały w postaci unormowanej w przedziałach normowania $[-1; 1]$. Otrzymane funkcje obiektu badań zestawiono w tab.3, a obraz graficzny dwóch z funkcji opisujących stan naprężeń dla powierzchni nagniatanych kulką przedstawiono na rys.1. W tab.3. zestawiono również funkcje aproksymujące zmiany stanu naprężeń w wyniku badanego zużycia oraz relacje normowania pozwalające na wyznaczenie funkcji aproksymujących dla wartości rzeczywistych badanych parametrów obróbki..

Tab.3. Funkcje obiektu badań opisujące stan naprężeń po nagniataniu

Lp	Badane naprężenia	Nagniatanie kulką Relacje normowania		Nagniatanie krążkiem Relacje normowania	
		$X = 0,44 (F - 0,5) - 1$	$F \text{ [kN]}$ $R_k \text{ [mm]}$	$X = 0,20 (R_k - 5) - 1$	$R_k \text{ [mm]}$ $f \text{ [mm/obr]}$ $F \text{ [kN]}$
		N [MPa]		N [MPa]	
1	suma naprężeń głównych	- 600,9	- 4,5X - 15,7Y + 127,8X ² + 3,8XY + 48,3Y ²	- 535,8 + 224X + 34Y + 2Z + 47,6X ² - 148,3XY - 46,8XZ + 11,1Y ² - 188,3YZ + 124,6Z ²	
2	naprężenia obwodowe	- 113,1 - 19,7X - 14Y + 38,7X ² - 36XY + 9,7Y ²		- 95,5 + 43X + 18,5Y + 49,5Z - 29,2X ² - 80,8XY - 34,3XZ + 70,3Y ² + 44,3YZ + 41,3Z ²	
3	naprężenia osiowe	- 360,7 - 51,3X + 16,8Y + 72,7X ² - 10XY + 69,2Y ²		- 195,3 + 133,5X - 4,5Y - 36Z - 27,1X ² - 72,5XY + 52XZ - 26Y ² - 38,5YZ + 26,4Z ²	
4	zmiana sumy nap. głównych w wyniku zużycia	475,5 + 81,8X + 19,3Y - 162,8X ² - 27XY - 155,3Y ²		177,7 - 228X - 55,5Y + 30,5Z + 12,7X ² + 154,6XY + 62,4XZ - 17,3Y ² + 162,6YZ - 32,3Z ²	
5	zmiana naprężeń obwodowych w wyniku zużycia	- 54,7 + 53,3X + 47,5Y - 76X ² + 22,3XY - 1,5Y ²		- 127,5 - 43,5X - 13Y - 58Z + 23,8X ² + 89XY + 31,5XZ - 43,7Y ² - 61YZ - 31,7Z ²	
6	zmiana naprężeń osiowych w wyniku zużycia	390,5 + 82,2X + 12,2Y - 114,8X ² - 45,5XY - 101Y ²		25,4 - 221,5X + 13Y + 66Z + 17,5X ² + 62XY - 92,5XZ + 62Y ² + 130YZ + 86Z ²	

2.5. Statystyczna weryfikacja otrzymanych funkcji

Po wykonaniu aproksymacji wyznaczono dla każdego punktu pomiarowego (układu planu) wartości funkcji aproksymującej w tym punkcie i porównanie ich z wartościami rzeczywistymi. Pozwoliło to wyznaczyć błędy bezwzględne i względne aproksymacji, maksymalny błąd bezwzględny i względny, błąd średniokwadratowy i błąd przeciętny oraz współczynnik korelacji wielowymiarowej. Umożliwiło to na podjęcie decyzji o przyjęciu lub odrzuceniu otrzymanych funkcji.



Rys.1. Graficzne przedstawienie funkcji obiektu badań (w postaci unormowanej) dla powierzchni nagniatanych kulką; a - naprężenia osiowe, b - naprężenia obwodowe
 Fig.1. The function of investigation object (normal form) for burnishing of surface by nodule; a - axis stresses, b - circumferential stresses

Przed podjęciem decyzji przeprowadzono formalną weryfikację adekwatności za pomocą testu F - Snedecora [1, 2, 6], to znaczy sprawdzono, czy stosunek F wariancji adekwatności $S^2(z)_a$ do wariancji wykonanych pomiarów $S^2(z)$ jest mniejszy od wartości krytycznej $F_{\alpha, f1, f2}$ (α - przyjęty poziom istotności, $f1$ - stopnie swobody związane z wariancją niedokładności pomiarów, $f2$ - stopnie swobody odniesione do wariancji adekwatności). Przyjęto poziom istotności 0,01 i 0,05.

Po zaakceptowaniu otrzymanej funkcji pod względem jej adekwatności względem wyników pomiarów można przystąpić do statystycznej weryfikacji istotności współczynników otrzymanych funkcji. Umożliwia to usunięcie z funkcji aproksymującej składników z tymi współczynnikami. W krańcowym wypadku może to doprowadzić do wyeliminowania (jako nieistotnych) niektórych z badanych wielkości wejściowych. Weryfikację istotności przeprowadzono wykorzystując test t - Studenta. Przed ewentualnym usunięciem nieistotnych współczynników wykonano również badanie ich skorelowania z innymi współczynnikami.

Praktycznie w każdej z funkcji otrzymanych znaleziono współczynniki nieistotne, jednak na skutek wystąpienia skorelowania z innymi współczynnikami ich usunięcie wymagałoby ponownego wyznaczenia współczynników dla zmodyfikowanej funkcji aproksymującej i ponowne przeprowadzenie weryfikacji adekwatności. Uzyskane w ten sposób funkcje nie ułatwiały dalszych obliczeń dlatego od eliminacji odstąpiono.

2.6. Merytoryczna analiza otrzymanych funkcji obiektów badań

Przeprowadzenie oceny adekwatności otrzymanych funkcji, dokonanie ewentualnej eliminacji nieistotnych współczynników i merytoryczna ocena otrzymanych funkcji pozwala potraktować je jako funkcje opisujące rozpatrywane zjawiska. Oczywiście funkcje te nie stają się modelem matematycznym obiektu badań opisującym w jakiejś logicznej formie teoretyczne zależności przyczynowo skutkowe właściwe opisywanemu obiektowi. Otrzymane funkcje stanowią jedynie funkcje obiektu

badania, a otrzymanych wyników nie można w żadnym wypadku bezkrytycznie przenosić poza obszar wartości wielkości wejściowych przyjęty w badaniach (dlatego tak istotnym jest właściwy wybór przedziałów zmienności badanych wielkości). Pozwalają one jednak na prześledzenie części zależności przyczynowo - skutkowych występujących w trakcie opisywanych tu badań. Możliwe staje się wyznaczenie tych parametrów wyjściowych, dla których wartości otrzymanych w warstwie wierzchniej naprężeń są największe i najmniejsze, lub dla których zmiany w trakcie badanej eksploatacji (w warunkach tarcia suchego) były największe i najmniejsze. Możliwą staje się również analiza wpływu poszczególnych czynników badanych (analiza jednoczynnikowa) na zachowanie się badanego obiektu badań.

Analiza otrzymanych wyników dla funkcji opisujących stan i zmiany stanu naprężeń własnych dla powierzchni nagniatanych kulką, opisanych funkcjami dwóch zmiennych (promienia kulki R_k i siły nacisku narzędzia F) nie nastrocza żadnych problemów. Wystarczy w tym celu sporządzić ich wykresy "trójwymiarowe" - dwa z nich przedstawiono na rys.1 lub warstwicowe. Obecnie istnieje wiele komputerowych programów narzędziowych pozwalających na wykonanie tego typu wykresów w zadanym zakresie wartości na podstawie podanej funkcji dwóch zmiennych. W ostateczności bez problemów wykres taki (choćby uproszczony) można wykonać ręcznie podstawiając do otrzymanej funkcji zbiór kilku interesujących nas wartości wielkości wejściowych. Na podstawie takich wykresów łatwo można odnaleźć położenie punktów, w których funkcje uzyskują wartości najmniejsze i największe. Można też prześledzić wpływ zmian wartości poszczególnych badanych czynników na zachowanie się badanego obiektu (zmiany właściwości warstwy wierzchniej). Dla przyjętego w opisywanych badaniach powierzchni nagniatanych kulką kompletnego planu badań podobna analiza była by możliwa również bez wyznaczania funkcji obiektu badanego jedynie na podstawie odpowiednio usystematyzowanych wyników pomiarów.

Odmienne przedstawia się analiza wyników pomiarów dla powierzchni nagniatanych krążkiem. Tu przy analizie otrzymanych wyników pomiarów praktycznie nie sposób posłużyć się bezpośrednio otrzymanymi w pomiarach wartościami. Wynika to ze specyfiki planów poliselekcyjnych (jednym z nich jest zastosowany tu plan Hartleya). Plany te z samego założenia (niekompletność planu) nie pozwalają na przeprowadzenie analizy wpływu poszczególnych badanych czynników na badany obiekt bez użycia otrzymanej funkcji obiektu, mocno problematyczne staje się również wychwycenie wartości ekstremalnych badanego obiektu.

Znalezienie wartości ekstremalnych przy użyciu otrzymanych funkcji również nie jest sprawą oczywistą. Zawodzą tu bowiem proste metody badania przebiegu zmienności stosowane dla funkcji mniejszej ilości zmiennych. Ekstremum funkcji może bowiem leżeć wewnątrz badanego przedziału i jest to ekstremum lokalne dla którego spełniony jest warunek zerowania się pierwszych pochodnych (ekstremum takie może być jedno lub więcej). Funkcja może również posiadać ekstremum na którejś z granic badanego przedziału zmienności i ten przypadek w rozpatrywanych tu badaniach występuje najczęściej. Jedynym wyjściem wydaje się sięgnięcie do jednej z wielu metod optymalizacji funkcji, to znaczy poszukiwania maksimum lub minimum badanych funkcji w założonym obszarze zmienności. W opisywanych tu badaniach dla powierzchni nagniatanych krążkiem posłużono się metodami Complex (Boxa), Hooke'a - Jeeves'a i Gaussa - Seidela.

Wykorzystanie otrzymanych funkcji pozwala ponadto na poszukiwania zależności opisujących zachowanie się badanego obiektu pod wpływem zmian jednego z badanych czynników wejściowych (promienia krzywizny krążka R_p , siły nagniatania F i posuwu f) przy stałych wartościach pozostałych wielkości badanych. Uzyskuje się w ten sposób przekroje badanej funkcji aproksymującej będące łatwymi do przeanalizowania funkcjami badanej zmiennej. Oczywiście konieczne jest przeanalizowanie takich przebiegów dla kilku wartości każdej z wielkości wejściowych. Można również uzyskaną funkcję $N = f(x, y, z)$ trzech zmiennych zamienić, poprzez przyjęcie któreś z badanych wielkości wejściowych za stałą, na funkcje (jedną lub kilka) dwóch zmiennych, które łatwo przedstawić w formie graficznej.

3. Podsumowanie

Przedstawiony w niniejszym opracowaniu sposób podejścia do przeprowadzenia badań stanu wybranych własności użytkowych warstw wierzchnich umożliwia znaczne przyspieszenie prowadzonych pomiarów przy jednoczesnym znacznym zwiększeniu ilości możliwych do uzyskania w ich wyniku informacji. Dzięki zastosowaniu współczesnych metod planowania doświadczeń i wykorzystywania selekcyjnych planów badań naukowych znacznie spada czas i co bardzo istotne koszt prowadzonych badań. Przeprowadzenie wyznaczania funkcji badanego obiektu znacznie zwiększa ilość możliwych do uzyskania informacji, przyspiesza i ułatwia ich opracowanie. Otrzymane funkcje umożliwiają łatwe przeprowadzenie komputerowej symulacji badanych procesów. Dlatego przedstawiony tu sposób prowadzenia badań staje się powoli codziennością.

4. Literatura

1. Polański Z.: Planowanie doświadczeń w technice. PWN. Warszawa 1984.
2. Polański Z.: Cadex esdet. Wyd. Cermet. Kraków 1992.
3. Tubielewicz K.: Analiza naprężeń powstających w warstwie wierzchniej podczas procesu nagniatania. Politechnika Częstochowska. Częstochowa 1993.
4. Tubielewicz K., Zaborski Z.: Zużycie powierzchni stali poddanej procesowi nagniatania. Mat. XVII Szkoły Trybologicznej KBM PAN. ZN Pol. Rzeszowskiej nr 28/91. Rzeszów 1991.
5. Tubielewicz K. i inni: Prace Zespołu Obróbki Powierzchniowej. Monotematyczny Zeszyt Naukowy. Politechnika Częstochowska. Nr 143. Mechanika 23. Częstochowa 1990.
6. Volk V.: Statystyka stosowana dla inżynierów. WNT. Warszawa 1965.

THE ANALYSIS OF CHANGES OF THE STATE OF OWN STRESSES AND THE OTHER OPERATING ATTRIBUTES OF SURFACE LAYER DURING WEAR PROCESS

Summary

In this work the way of calculation of the investigation function object, which specified the chosen operating - useful attributes of the surface layer the pressure rolling burnishing process was presented. The investigations were made for bivalent and trivalent ensembles input quantity - surface parameters, used the scheme of completely investigations ($PS/DK - 3^2$) and Hartley's investigations ($PS/DS - P:Ha, - hS$). The given results make possibilities for intentionally and voluntary control of treatment parameters. In this way it can get the required useful attributes of surface.