

KOMITET METROLOGII I APARATURY NAUKOWEJ
POLSKIEJ AKADEMII NAUK

PROBLEMY METROLOGII ELEKTRONICZNEJ I FOTONICZNEJ 7

Praca zbiorowa
pod redakcją Janusza Mroczi



Oficyna Wydawnicza Politechniki Wrocławskiej

PROBLEMY METROLOGII
ELEKTRONICZNEJ I FOTONICZNEJ

KOMITET METROLOGII I APARATURY NAUKOWEJ
POLSKIEJ AKADEMII NAUK

**PROBLEMY METROLOGII
ELEKTRONICZNEJ
I FOTONICZNEJ**
7

*Praca zbiorowa
pod redakcją Janusza Mroczki*



Oficyna Wydawnicza Politechniki Wrocławskiej
Wrocław 2014

Opracowanie redakcyjne i korekta

Hanna JUREK

Projekt okładki

Janusz MROCZKA

Wszelkie prawa zastrzeżone. Żadna część niniejszej książki, zarówno w całości, jak i we fragmentach, nie może być reprodukowana w sposób elektroniczny, fotograficzny i inny bez zgody wydawcy i właścicieli praw autorskich.

© Copyright by Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 2014

OFICYNĄ WYDAWNICZĄ POLITECHNIKI WROCŁAWSKIEJ

Wybrzeże Wyspiańskiego 27, 50-370 Wrocław

<http://www.oficyna.pwr.edu.pl>

e-mail: oficwyd@pwr.edu.pl

zamawianie.ksiazek@pwr.edu.pl

ISBN 978-83-7493-867-9

Druk i oprawa: EXPOL, P. Rybiński, J. Dąbek, sp.j., ul. Brzeska 4, 87-800 Włocławek
tel. 54 232 37 23, e-mail: sekretariat@expol.home.pl

SPIS RZECZY

Przedmowa	7
Patrycja RUŚKOWSKA, Michał MOSIĄDZ, Włodzimierz POPIOLEK, Janusz MROCZKA, Kwantowy Układ Jednostek Miar SI – ewolucja czy rewolucja?	9
Adam BIENKOWSKI, Roman SZEWCZYK, Magnetoelastyczne właściwości materiałów magne- tycznie miękkich jako sensorów naprężeń i sił	81
Grzegorz WIELGOSZEWSKI, Maciej RUDEK, Grzegorz JÓŹWIAK, Teodor GOTSZALK, Skaningowa mikroskopia bliskiego pola termicznego w metrologii mikro- i nanostruktur	149
Sławomir CIĘSZCZYK, Wybrane zagadnienia pomiaru zawartości gazów w zmieniających się wa- runkach	205

Pomiar bezpośrednio mówi nam równie mało o fizycznym obrazie świata, jak i o świecie realnym. Każdy pomiar oznacza pewien proces w narzędzie zmysłowym fizyka, dokonującego pomiaru lub w użytych przez niego przyrządach, o którym to procesie wiadomo tylko tyle, że pozostaje w jakimś związku z mierzonym, realnym procesem.

M. Placzek, *Jedność fizycznego obrazu świata*,
KIW, Warszawa 1970

PRZEDMOWA

Metrologia należy do działu nauki doświadczalno-informacyjnego. Informacje o otaczającym nas świecie fizycznym czerpiemy z obserwacji przy pomocy naszych zmysłów. To dzięki obserwacjom bezpośrednim, które docierają do naszej świadomości, budujemy w naszych myślach wrażenie przestrzenności otaczającego świata. Tak uzyskany obraz świata przestrzenno-czasowy, wynikający z obserwacji bezpośredniej jest oparty na doświadczeniach i wrażeniach indywidualnych obserwującego, stąd jest on jakościowy, subiektywny, a pod względem odczucia czasu – ciągły. Dopiero przyjęcie jednego z własnych wrażeń obserwatora za jednostkę pomiarową umożliwia powstanie w jego świadomości ilościowego obrazu świata oraz przedstawienie swych własnych wrażeń za pomocą liczb. Tak powstały obraz ilościowy jest subiektywny z racji wyboru jednostki na wrażenia obserwatora.

Obiektywizm w ocenie występujących w świecie zjawisk uzyskamy wówczas, gdy za jednostkę miary przyjmujemy zjawisko zewnętrzne niezależne od obserwatora i potrafimy porównać go ze zjawiskami docierającymi do naszej świadomości.

W tak zarysowanym procesie naturalny obraz świata zastępujemy abstrakcyjnym schematem zastępczym, którego właściwości wyrażamy w sposób dyskretny za pomocą liczb. Stąd nietrudno wywnioskować, że stworzenie odpowiedniego systemu jednostek i ich wzorców jest podstawą naukowej części metrologii; natomiast działem miernictwa jest porównanie wskazań przyrządów użytkowych, używanych w codziennym życiu, tak aby uzyskaną informację o zjawiskach i rzeczach z wzorami podstawowymi umożliwiało obiektywność uzyskanych informacji.

Tak tworzony obraz tych zewnętrznych zjawisk, powstający w naszej świadomości, zbudowany jest z pojęć abstrakcyjnych stworzonych przez nas samych: ze słów

i zdań języka naturalnego, z liczb oraz z symboli logiki tworzących języki sztuczne, wśród nich również dialekty i gwary poszczególnych dyscyplin naukowych. To między dziedziną zjawisk i dziedziną symboli działa wiele dyscyplin ogólnie uznawanych za naukowe, które to, przetwarzając energię na inną jej postać, umożliwiają uzyskanie danego zjawiska w życiu człowieka lub łatwiejsze zinterpretowanie ich istoty – jak np. chemia bądź też zmieniających formę symboli abstrakcji w celu zbliżenia wyniku do znanych nam pojęć. To typowy sposób działania matematyka i informatyka. Naturalne wydaje się, że powinno istnieć ogniwo, w którym rzeczywiście następuje przetworzenie energii w symbole. Tym ogniwem jest metrologia. Oczywiście jest, że to przejście od naukowych dyscyplin zajmujących się procesami energetycznymi do nauk abstrakcyjnych jest też naukowe, dlatego metrologia jest oddzielną dyscypliną naukową. To metrologia tworzy w swoich poczynaniach zespół twierdzeń uznawanych aktualnie powszechnie za prawdziwe, stanowiących usystematyzowaną całość sprawdzonych doświadczalnie. To proces precyzowania i udowadniania poszczególnych twierdzeń oraz powiązania ich w spójny system jest treścią wydawanej serii „Problemy Metrologii Elektronicznej i Fotonicznej”.

Wydając kolejny już siódmy tom ukazujemy, my jako grupa uczonych publikująca w tej serii, że chęć poznania pobudza naszą fantazję do dalszych dokonań. To fantazja ludzka wydaje się niemierzalną z natury wielkością. Prawdopodobnie to ona wraz z opartym na niej działaniem stanowi indywidualny i subiektywny świat wewnętrzny każdego z nas, występujący obok świata zewnętrznego, ilościowego i zobiektywowanego z pomocą pomiarów.

Janusz Mroczka

Patrycja RUŚKOWSKA*, Michał MOSIĄDZ*,
Włodzimierz POPIOŁEK*, Janusz MROCZKA**

KWANTOWY UKŁAD JEDNOSTEK MIAR SI – EWOLUCJA CZY REWOLUCJA?

Niniejsze opracowanie przedstawia zagadnienia związane z weryfikacją, ustanowionego w 1960 roku i obowiązującego nadal, Międzynarodowego Układu Jednostek Miar SI. Omówiono w nim ewolucyjny proces kształtowania się definicji jednostek podstawowych zachodzący na przestrzeni wieków XIX i XX. W pracy przedstawiono nowe założenia definicyjne systemu definiowania jednostek miar. Zgodnie z Uchwałą nr 1 XXIV. Generalnej Konferencji Miar (CGPM) (Paryż, 2011), Międzynarodowy Układ Jednostek SI (franc. *Système International d'unités*) zostanie wkrótce poddany weryfikacji. W nowym ujęciu wszystkie jednostki podstawowe, łącznie z kilogramem (obecnie definiowanym za pomocą materialnego prototypu), będą definiowane na podstawie numerycznych wartości uniwersalnych stałych fizycznych, jak na przykład jednostka długości jest realizowana na podstawie wartości prędkości światła w próżni. Nowe podejście definicyjne jednostek, choć technicznie nader skomplikowane i bardzo kosztowne, umożliwi w przyszłości realizowanie bardziej precyzyjnych pomiarów.

1. WSTĘP

Burzliwemu rozwojowi w nauce i przemyśle w wiekach XIX–XX towarzyszyło wprowadzanie nowych definicji i jednostek miar opartych na odkrywanych zjawiskach fizycznych [1]. Przyjmowane jednostki były w wielu przypadkach od siebie niezależne i nie tworzyły spójnego układu. Dopiero prace metrologiczne prowadzone w latach 1930–1960 umożliwiły opracowanie układu o wielkich walorach użytkowych. Intencją jego twórców było to, aby nowy układ jednostek miar był logiczny, spójny i oparty na solidnych podstawach naukowych [1].

W okresie ostatniej dekady na międzynarodowym forum toczy się intensywna dyskusja dotycząca redefinicji jednostek układu SI (franc. *Système International*

* Główny Urząd Miar, ul. Elektoralna 2, 00-139 Warszawa.

** Katedra Metrologii Elektronicznej i Fotonicznej, Politechnika Wroclawska, ul. B. Prusa 53/55, 50-317 Wrocław.

d'unités). Weryfikacja, przyjętego i zaleconego przez XI. Generalną Konferencję Miar (CGPM) w 1960 roku, systemu wiąże się przede wszystkim z dążeniem do sprostania wyzwaniom dwudziestego pierwszego wieku [1, 5].

Postęp techniczny wraz z zaawansowaniem technologicznym umożliwia wykonanie wzorców dla pozostałych jednostek podstawowych układu SI, które również będą oparte na fundamentalnych zjawiskach fizycznych, tak by można było realizować je niezależnie w różnych laboratoriach z taką samą dokładnością.

XXIV posiedzenie Generalnej Konferencji Miar, które odbyło się w dniach 17–21 października 2011 r., było odpowiednią platformą dla omówienia ważkich z perspektywy nauki i przemysłu zagadnień redefinicji jednostek podstawowych układu SI. Opracowana i sformułowana przez Komitet Doradczy ds. Jednostek Miar (CCU) propozycja weryfikacji systemu SI w formie projektu Rezolucji A została przedłożona Konferencji CGPM. Projekt ten powstał w efekcie długotrwałych dyskusji w komitetach doradczych Międzynarodowego Komitetu Miar (CIPM) oraz kręgach naukowych i jest wyrazem poparcia dla idei redefinicji kilograma, ampera, kelwina i mola. Z powodu niewystarczającego stopnia zaawansowania badań, zwłaszcza nad redefinicją jednostki masy, niemożliwe jest określenie terminu ukończenia weryfikacji całego systemu SI. Dodatkowo sytuację komplikuje fakt, iż zmiana definicji kilograma może wpłynąć na definicje innych jednostek fizycznych, w tym jednostek podstawowych SI. Trzy jednostki tego układu są bezpośrednio powiązane z definicją kilograma: amper odnosi się do jednostki siły (niutona), kandela związana jest z jednostką mocy (watem) i mol – do pewnej ilości materii układu zawierającego liczbę cząstek równą liczbie atomów w masie 0,012 kg izotopu węgla C^{12} [3].

Celem niniejszej pracy jest omówienie idei redefinicji podstawowych jednostek miar należących do międzynarodowego układu SI i przedstawienie ich powiązania z fundamentalnymi stałymi fizycznymi: prędkością światła w próżni, c , częstotliwością stanu podstawowego struktury nadsubtelnej atomu cezu $^{133}\text{Cs} - (\Delta\nu^{133}\text{Cs})_{hfs}$, stałej Boltzmanna, k_B , Avogadra, N_A , światłości, K_{Cd} oraz Plancka h .

Rozdział 2 został poświęcony ewolucji systemu miar, natomiast w rozdziale 3 omówiono uniwersalne stałe fizyczne i ich powiązanie z podstawowymi jednostkami miar układu SI. Rozdziały 4, 5, 6 i 7 zawierają kolejno obowiązujące realizacje definicji oraz propozycje redefinicji: kilograma, mola, kelwina i ampera, dotyczące nowego systemu definiowania miar. W rozdziale 8 opisano prawnoadministracyjną problematykę systemów miar.

2. W KIERUNKU KWANTOWEGO UKŁADU JEDNOSTEK SI

Międzynarodowy Układ Jednostek Miar (SI) (franc. *Système International d'unités*), będący z założenia tworem nieskończonym, ustanowiony w 1960 roku był następnie modyfikowany i udoskonalany przez kolejne Generalne Konferencje Miar

(CGPM) (franc. *Conférence Générale des Poids et Mesures*). Został stworzony na podstawie metrycznego systemu miar, a jednostki w nim zawarte dzielą się na podstawowe i pochodne. W układzie tym dzięki przestrzeganiu zasady spójności każdej wielkości jest przyporządkowana jedna, jednoznacznie określona jednostka miary [1]. Układ SI został oficjalnie przyjęty przez wszystkie kraje świata, z wyjątkiem Liberii oraz Birmy. W Polsce układ SI obowiązuje od 1966 r.

Decyzją XXIV. Generalnej Konferencji Miar (Rezolucja 1) przewiduje się dalsze zmiany układu jednostek SI, realizowane w ramach stałego procesu jego udoskonalania. Wśród jednostek podstawowych Międzynarodowego Układu Jednostek Miar SI znajdują się obecnie: metr, kilogram, kandela, kelwin, amper, sekunda i mol (tab. 1).

Tabela 1. Jednostki podstawowe Międzynarodowego Układu jednostek Miar (SI) [2]
Table 1. Base units of the International Units System (SI) [2]

Wielkość podstawowa	Jednostka podstawowa SI	
Nazwa	Oznaczenie	
Długość	metr	m
Masa	kilogram	kg
Czas	sekunda	s
Prąd elektryczny	amper	A
Temperatura termodynamiczna	kelwin	K
Liczność materii	mol	mol
Światłość	kandela	cd

Podstawowe jednostki miar układu SI umożliwiają tworzenie jednostek pochodnych dla wielkości mechanicznych, elektromagnetycznych, cieplnych, fizykochemicznych oraz świetlnych.

Uchwałą XIV. Generalnej Konferencji Miar w 1971 roku do układu SI dla wielkości fizykochemicznych został wprowadzony mol, reprezentujący licznosc materii i będący *de facto* odmianą jednostki masy. Masa mola, wyrażona w kilogramach, zależy od rodzaju cząstek (atomów, jonów itd.).

Wiek XXI niesie za sobą nowe pomysły i rozwiązania prowadzące do udoskonalenia przyjętego w 1960 roku układu SI. Brana pod rozwagę była również możliwość redukcji uporządkowanego zbioru siedmiu podstawowych jednostek miar. Usunięciu mogłyby ulec mol i kandela. Rozwiązanie to związane było z tym, iż we wczesnym wszechświecie w warunkach ekstremalnie wysokiej temperatury (por. teoria „wielkiego wybuchu”) w czasie i przestrzeni znajdowały się cząstki elementarne obdarzone masą i ładunkiem elektrycznym. Stąd też status kompletności uzyskują: długość, masa, czas, prąd elektryczny i temperatura, których pochodzenie może mieć swoje uzasadnienie w ewolucji i budowie wszechświata [5, 6]. Obecnie nie przewiduje się rezygnacji ani z mola, ani z kandel. Stosowanie mola jest niezbędne w chemii podczas ilościowego określania składników reakcji chemicznych, umożliwia także wyeliminowanie

nowanie niezbyt ściśle określonych pojęć chemicznych takich, jak gramoatom czy gramocząsteczka. Choć w ujęciu energetycznym światłości odpowiada natężenie promieniowania świetlnego (z jednostką wat na steradian – W/Sr), kandela jest jednostką powszechnie stosowaną w pomiarach światłości zarówno realizowanych w fotometrii, jak i radiometrii [1]. Ich obecność wśród jednostek podstawowych układu SI jest uzasadniona ze względów praktycznych.

Postęp techniczny i technologiczny, zachodzący w XX i XXI wieku, wywarł ogromny wpływ również na rozwój współczesnej metrologii, stając się motorem do działania dla uczonych, którzy, odnajdując we wszechświecie stabilne odniesienie, postanowili zweryfikować obowiązujący globalnie system miar SI. Choć kwantowe podejście definicyjne do jednostek układu SI zaczęło się od metra, krokiem milowym w tym podejściu było osiągnięcie znacznej poprawy niepewności pomiaru dla wielkości elektrycznych, których definicyjne realizacje oparte są na kwantowym efekcie Halla czy Josephsona [8]. Dalej omówiono wybrane jednostki, uwzględniając odkrycia naukowe, które stały się motorem rozwoju ich definicji.

3. UNIWERSALNE STAŁE FIZYCZNE

Stałe fizyczne i jednostki miar; które z tych pojęć pojawiło się jako pierwsze w nauce? Wydawać by się mogło, że stałe fizyczne są wartościami bezwzględными, niezależnymi od warunków otoczenia, pierwotnymi. Umożliwiają zdefiniowanie pojęć wtórnych i dokładny opis ilościowy otoczenia, zjawisk fizycznych. Jednak, spoglądając wstecz na rozwój nauki, pojawiają się wątpliwości dotyczące oczywistości tego stwierdzenia [7].

Jeszcze niedawno, 100–200 lat temu, stałe fizyczne, używane powszechnie w dzisiejszej nauce i technice, były nieznanne. Były one wartościami opisującymi zjawiska, które zostały odkryte stosunkowo niedawno; zastąpiły stałe, poprzednio wykorzystywane w nauce i technice odpowiadające ówczesnemu poziomowi wiedzy, bliższe rzeczywistości odbieranej przez człowieka zmysłami, makroskopowe. Dzisiaj stałe, takie jak stała Plancka, ładunek i masa elektronu są dla przeciętnego czytelnika wartościami abstrakcyjnymi i trudno sobie wyobrazić, że dzisiejsza technika umożliwia wyznaczenie, zmierzenie ich wartości wprost, a nie jedynie za pomocą bardzo skomplikowanych obliczeń teoretycznych.

Do wyrażenia wartości, do stworzenia ilościowego opisu każdej z tych stałych potrzebna jest jednostka miary; odniesienie, bezwzględne, umożliwiające określenie, ile dokładnie wynosi wartość stałej. I tutaj pojawia się zasadnicze pytanie – jak określić, do czego odnieść wartość stałej fizycznej, wielkości bezwzględnej, niezmiennej i podstawowej? Jaki sens ma odnośnienie ich do powszechnie stosowanych jednostek miar układu SI bądź jakiegokolwiek innego, w którym jednostki miar są wartościami umownymi, często stworzonymi sztucznie, w odniesieniu do makroskopowych wzor-

ców fizycznych, które charakteryzuje duża zmienność spowodowana chociażby koniecznością zachowania odpowiednich warunków środowiskowych ich przechowywania czy efektami starzeniowymi. Przecież to wartości bezwzględne, niezmiennie, powinny być podstawą układu jednostek miar, punktami odniesienia, w stosunku do których określane są inne wielkości. Podejście takie wymaga, wraz z rozwojem nauki, redefinicji jednostek miar, która powinna być procesem permanentnym, nieprzerwanym. Ścisła współpraca środowisk metrologicznych z nauką wymusza zmiany, doskonalenie koncepcji stosowanego układu miar do rozwoju nauki, podążania za nim. A co pewien czas postęp wymusza rewolucję, zupełną zmianę podejścia do definicji jednostek miar. Dzięki temu układ miar nie pozostaje archaicznym zabytkiem przeszłości, lecz ułatwia coraz doskonalszy i dokładniejszy pomiar, a co za tym idzie – opis rzeczywistości.

Należy jednak pamiętać, że uniwersalne układy jednostek miar nie są stosowane jedynie w ośrodkach naukowych i laboratoriach. Ich istotą jest uniwersalność, możliwość wykorzystania w możliwie wielu obszarach działalności człowieka. Dlatego opierają się na zjawiskach i wzorcach makroskopowych o częstokroć historycznym pochodzeniu.

Nikt sobie nie wyobraża, aby w handlu i rozliczeniach ekonomicznych odnosić się do masy elektronu zamiast kilograma, aby czas mierzyć w jednostkach o zbliżonej wartości do czasu połowicznego rozpadu substancji promieniotwórczych. Zresztą naukowcy w swej codziennej pracy często stosują własne, inne od powszechnie stosowanych w codziennym życiu, układy jednostek miar – często wymienianym przykładem mogą tu być astronomowie i astrofizycy, środowiska związane z fizyką i chemią na poziomie molekularnym, marynarze itp. W ich codzienności i opisie zjawisk, którymi się zajmują ze względu na skalę wielkości, ze względów praktycznych wygodniejsze jest posługiwanie się zupełnie innymi jednostkami, takimi jak lata świetlne, parseki czy odległości atomowe. Umożliwia im to uproszczenie obliczeń po utrzymaniu wymaganej ich dokładności lub posługiwanie się prostszymi formalnie wzorami, podczas wprowadzenia odpowiednich współczynników proporcjonalności, specyficznych dla danej dziedziny.

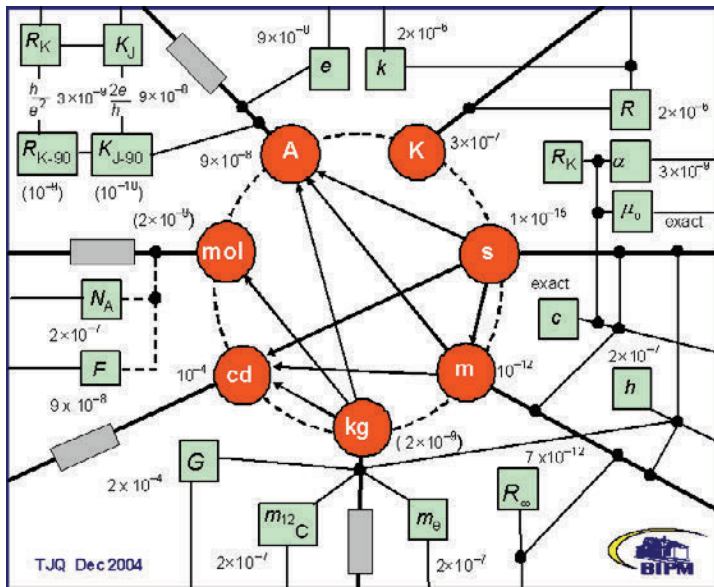
3.1. REKOMENDACJE XXIV. GENERALNEJ KONFERENCJI MIAR

W 2011 roku na mocy uchwalonej przez XXIV. Generalną Konferencję Miar, Rezolucji 1 definicje siedmiu podstawowych jednostek Międzynarodowego Układu Jednostek Miar SI (zwanego *Quantum SI*) realizowane będą według następujących stałych fizycznych [4, 6], (rys. 1):

- częstotliwości stanu podstawowego struktury nadsubtelnej atomu cezu,
- $^{133}\text{Cs} - (\Delta\nu^{133}\text{Cs})_{\text{hf}}$, wynoszącej dokładnie 9 192 631 770 Hz,
- prędkości światła w próżni c wynoszącej dokładnie $299\,792\,458\text{ m}\cdot\text{s}^{-1}$,
- stałej Plancka h , wynoszącej dokładnie $6,62606\text{X}\cdot 10^{-34}\text{ J}\cdot\text{s}$,

- ładunku elementarnego e , wynoszącego dokładnie $1,60217 \times 10^{-19} \text{ C}$,
- stałej Boltzmanna k wynoszącej dokładnie $1,3806 \times 10^{-23} \text{ J} \cdot \text{K}^{-1}$,
- stałej Avogadro N_A , wynoszącej dokładnie $6,02214 \times 10^{23} \text{ mol}^{-1}$,
- skuteczności świetlnej K_{cd} dla promieniowania monochromatycznego o częstotliwości $540 \cdot 10^{12} \text{ Hz}$,

gdzie symbol X określa dodatkowe cyfry wyznaczone i zatwierdzone przez Komitet (ang. *Committee on Data for Science and Technology, CODATA*) przed przyjęciem nowych definicji.



Rys. 1. Schemat wzajemnych powiązań między podstawowymi jednostkami miar oraz ich powiązań z fundamentalnymi stałymi fizycznymi w kwantowym układzie SI (źródło: http://www.bipm.org/en/si/si_constants.html)

Fig. 1. Diagram of relations between base units and their links to the fundamental physical constants in the Quantum SI (source: http://www.bipm.org/en/si/si_constants.html)

W tabeli zestawiono porównanie realizacji definicji podstawowych jednostek Międzynarodowego Układu Jednostek Miar SI, realizowanych za pomocą wzorców materialnych i stałych fizycznych (tab. 2), [4, 6, 7, 12].

Zgodnie z założeniem, aby wyrazić wartość stałej fizycznej Q , należy przedstawić ją jako wynik wartości liczbowej $\{Q\}$ i jednostki $[Q]$ [7]

$$Q = \{Q\} \cdot [Q] \quad (1)$$

Przykładowo dla stałej Plancka h równanie (1) przyjmie postać

$$h = \{h\} \cdot [h] = 6,626069 \dots \cdot 10^{-34} \text{ J} \cdot \text{s} \quad (2)$$

Tabela 2. Zestawienie sposobów realizacji jednostek podstawowych układu SI i Quantum SI
 Table 2. Comparison of realization methods of base units of the International System SI and the Quantum SI

Obecny układ SI		Nowy układ SI	
Jednostka	Realizacja	Jednostka	Realizacja
Sekunda, s	$(\Delta \nu^{133}\text{Cs})_{hfs}$	sekunda, s	$(\Delta \nu^{133}\text{Cs})_{hfs}$
Metr, m	C	C	prędkość światła w próżni
Kilogram, kg	masa wzorca materialnego	h	stała Plancka
Amper, A	μ_0	e	elementarny ładunek elektryczny
Kelwin, K	$T_0 \text{ } ^\circ\text{C}$	k_B	stała Boltzmannna
Mol	$M(^{12}\text{C})$	N_A	stała Avogadro
Kandela	K_{cd}	K_{cd}	skuteczność świetlna

Podczas definiowania jednostki $[Q]$, konieczne jest wyznaczenie eksperymentalne wartości liczbowej $\{Q\}$. Natomiast w celu określenia wartości liczbowej $\{Q\}$, należy zdefiniować jednostkę $[Q]$. Przykładowo, jeśli liczbową wartość $\{h\}$ wynosi dokładnie $6,626069 \dots \cdot 10^{-34}$, to w rezultacie definiowana jest jednostka J·s. Przez wybór wartości liczbowej stałej fizycznej można zatem definiować jej jednostkę. Celem takiego sposobu definiowania, podobnie jak w przypadku metra i sekundy, jest wyrażenie jednostki masy w kategoriach jednostki częstotliwości przez fundamentalne równanie (3), opisujące energię [15]:

$$E = h \cdot \nu \quad (3)$$

Niepewności wyznaczenia wartości wielu innych stałych uniwersalnych oraz współczynników konwersji energii (ang. *energy conversion factor*) zostaną wyeliminowane lub znacznie zredukowane w przypadku, gdyby stałe: h , e , k_B i N_A miały przypisane dokładne wartości numeryczne w definicjach powiązanych z nimi jednostek.

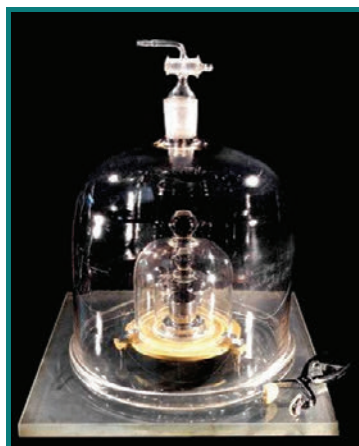
Dalej zostaną omówione cztery wybrane jednostki miar układu SI, ewolucja ich definicji oraz aktualny stan wiedzy dotyczący postępów w badaniach prowadzonych nad ich kwantową redefinicją, zapowiedzianą przez ostatnią XXIV. Generalną Konferencję Miar (CGPM) [6].

4. KILOGRAM

4.1. WZORZEC MATERIALNY KILOGRAMA

Spośród siedmiu jednostek podstawowych Międzynarodowego Układu Jednostek Miar SI tylko kilogram nie jest definiowany w sposób uniwersalny, czyli za pomocą stałej fizycznej odniesienia. Natomiast odtworzenie wzorca masy (składającego się z 90% platyny i 10% irydu) wymaga bezpośredniego porównania danej próbki z oryginalnym wzorcem.

nałem lub jedną z jego oficjalnych kopii. Pierwsze trzy prototypy kilograma, oznaczone symbolami KI, KII i KIII, były porównywane w Obserwatorium Paryskim z Kilogramme des Archives (pierwszym wzorcem jednostki masy, ustanowionym w 1791 r., którego definicja oparta była na objętości – 1 dm^3 wody destylowanej w temp. $4 \text{ }^\circ\text{C}$). Ostatecznie za międzynarodowy wzorec kilograma przyjęto prototyp KIII, co zostało usankcjonowane uchwałą I. Generalnej Konferencji Miar. Od tej pory jest on przechowywany w sejfie Międzynarodowego Biura Miar (BIPM) i stanowi odniesienie dla wzorców niższego rzędu w dziedzinie pomiarów masy. Podczas ostatniego stulecia wykonano kilka porównań z wzorcami-kopiami, utrzymywanymi w krajowych instytutach metrologicznych (z sejfu wzorec wyjmowany był tylko trzykrotnie: w 1939 r., 1946 r. oraz w latach 1988–1989) [9]. Rezultaty wymienionych porównań wykazały pewne zmiany masy prototypu MPK, nazywanego przez Francuzów *La Grande K*, względem jego kopii w czasie. Zmiany jego masy względem narodowych kopii kształtowały się na poziomie $5 \cdot 10^{-8}$ – w przybliżeniu $50 \text{ }\mu\text{g}$ od pierwszego wzorcowania, jakie odbyło się ponad 100 lat temu. Na podstawie podanych obserwacji wykazano, iż kopie MPK regularnie przybierały na masie, podczas gdy on stawał się coraz lżejszy (rys. 2) [10].

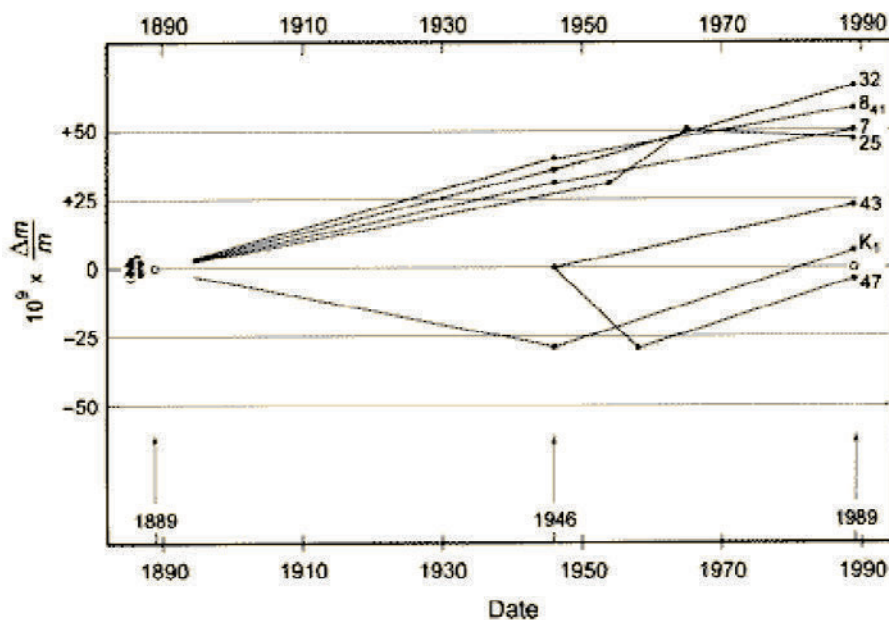


Rys. 2. Międzynarodowy prototyp kilograma (tzw. MPK), przechowywany obecnie w sejfie BIPM
(źródło: <http://www.bipm.org/en/scientific/mass/prototype.html>)

Fig. 2. International Prototype of Kilogram, kept in the safe of the BIPM
(source: <http://www.bipm.org/en/scientific/mass/prototype.html>)

W związku z rosnącą niepewnością zapewnienia spójności pomiarowej w dziedzinie pomiarów masy, związanej z niestabilnością masy MPK, podjęto decyzję o rozpoczęciu badań nad redefinicją kilograma (mającego niepewność względną równą zeru). Zmiany definicyjne mogą jednak wpłynąć także na definicje innych jednostek fizycznych, w tym jednostek podstawowych SI. Trzy jednostki tego układu są bezpośrednio

powiązane z definicją kilograma: amper odnosi się do jednostki siły (niutona), kandela związana jest z jednostką mocy (watem) i mol – do pewnej ilości materii układu zawierającego liczbę cząstek równą liczbie atomów w masie 0,012 kg izotopu węgla C^{12} [10, 11].



Rys. 3. Wyniki z porównań wybranych siedmiu kopii Międzynarodowego Prototypu Kilogramu (MPK) [11]

Fig. 3. Results of the comparisons of seven copies of the International Prototype of Kilogram (IPK) [11]

4.2. KWANTOWA DEFINICJA KILOGRAMA

Opracowanie redefinicji jednostki kg opartych na stałych fizycznych powinno być realizowane zgodnie z rekomendacjami Komitetu Doradczego ds. Jednostki Masy CCM przy Międzynarodowym Komitecie Miar (CIPM) [12]: G1 (2010) i G2 (2010). Według wytycznej CCM G1 – na podstawie uzyskanych rezultatów w trzech niezależnych eksperymentach, wykonanych za pomocą wagi Watta (Watt balance) oraz w ramach projektów międzynarodowych dotyczących wyznaczenia stałej Avogadra (ang. *International Avogadro Coordination*, IAC), opartych na kryształach krzemu. Wartości właściwych stałych powinny zostać wyznaczone ze względną niepewnością standardową nie większą niż $5 \cdot 10^{-8}$ i przynajmniej jeden z tych wyników powinien być obarczony niepewnością nie większą niż $2 \cdot 10^{-8}$. Ponadto dla każdej stałej fizycznej wartości ustalone na podstawie różnych eksperymentów powinny być stałe na pozio-

mie ufności 95% oraz będzie potwierdzona spójność pomiarowa prototypów BIPM z międzynarodowym prototypem kilograma. Według wytycznej CCM G2 – zastosowanie międzynarodowego prototypu kilograma ma potwierdzić spójność pomiarową z prototypami BIPM [12].

Od 2003 roku w ramach międzynarodowego projektu Avogadro [10, 13, 15–17], którego koordynatorem jest Physikalisch-Technische Bundesanstalt (PTB) w Niemczech, prowadzone są prace nad opracowaniem definicji kilograma opartej na stałej Avogadra, N_A . Jako taka stała Avogadra łączy właściwości atomowe i makroskopowe materii [13]. Kilogram określa tutaj masę dokładnie wyznaczonej liczby atomów związku lub pierwiastka w formie monokryształu o bardzo wysokim stopniu dokładności (liczonej w nanometrach) [13]. W projekcie Avogadra za materiał budulcowy wzorca kilograma zastosowano, najczęściej występujący w naturze, izotop ^{28}Si tego pierwiastka, który został wybrany zdaniem naukowców dlatego, że może powstawać w postaci niezwykle czystych, dużych i niemal doskonałych kryształów. Ponadto krzem jest materiałem bardzo stabilnym i łatwym w obróbce. Idealny odważnik kilogramowy powstał dzięki międzynarodowej współpracy instytutów NMI z Rosji, Niemiec, Włoch, Belgii, Japonii, Australii i USA. Pomiary krzemowego wzorca odbywają się w dużej próżni i w temperaturze odniesienia utrzymywanej w reżimie stałej kontroli.

W ramach badań międzynarodowy zespół naukowców dążył do policzenia liczby atomów N w pojedynczym, jednokilogramowym kryształku krzemu. Ostatecznym celem projektu Avogadra jest pomiar stałej Avogadra w takim kryształku z niepewnością pomiaru zaledwie $2 \cdot 10^{-8}$. W projekcie czerpano z wiedzy eksperckiej instytutów metrologicznych z całego świata. Naukowcy wykorzystali dwie kule wypolerowane w Australii. Następnie poddano ocenie doskonałość kryształów oraz wpływ wad w sieci krystalicznej. Włoski instytut metrologiczny (INRIM) zastosował interferometrię rentgenowską do ustalenia parametru sieci krystalicznej, a jego wyniki zostały potwierdzone na podstawie porównania pomiarów w naturalnym kryształku krzemu przeprowadzonych w amerykańskim Narodowym Instytucie Standaryzacji i Technologii (NIST). Masa dwóch kul krzemowych została porównana z masą wzorców kilograma przechowywanych w BIPM, PTB i japońskim krajowym instytucie metrologicznym (NMIJ). Warstwa powierzchniowa kul z ditlenku krzemu została poddana serii testów, w tym promieniowaniu rentgenowskiemu i promieniowaniu synchrotronowemu. Zmierzone również zanieczyszczenie powierzchni krzemkami miedzi i niklu oraz oceniono jego wpływ na objętość i masę kul. Wreszcie zespół z PTB zastosował nową metodę spektroskopii mas do ustalenia masy molowej kul.

Po zastosowaniu metody krystalograficznej XCRD (X-Ray Crystal Density), skutecznej do zastosowania przy stałej sieci krystalicznej a i stałej objętości kryształu krzemu policzono liczbę atomów N w próbkach z niepewnością pomiaru zaledwie $3 \cdot 10^{-8}$ [13, 16].

Naukowcy podkreślają jednak, że ich definicja nadal nie jest wystarczająco precyzyjna, aby zastąpić Międzynarodowy Prototyp Kilogramu z BIPM. W tym celu muszą

osiągnąć poziom niepewności rzędu $2 \cdot 10^{-8}$, jak wymaga tego Komitet Doradczy ds. Masy i Wielkości Pochodnych (CCM) przy Międzynarodowym Biurze Miar (BIPM) w Paryżu – Sévres.

Liczbę Avogadra można zdefiniować jako stosunek średniej (tj. uśrednionej po składzie izotopowym) objętości molowej V_{mol} krzemu do średniej objętości jego atomu V_a w badanym monokryształe krzemu ^{28}Si [10]:

$$N_A = \frac{V_{\text{mol}}}{V_a} \quad (4)$$

Ponieważ średnia objętość atomu jest wprost proporcjonalna do objętości komórki elementarnej (sześciianu o krawędzi a), i odwrotnie proporcjonalna do liczby atomów krzemu w komórce kryształu ($n = 8$) dla monokryształu o masie m oraz objętości V po prostych przekształceniach otrzymamy równanie na stałą Avogadra:

$$N_A = \frac{nM_{\text{mol}}V}{mV_0} \quad (5)$$

Po uwzględnieniu $A_r(\text{Si})$ – stosunku atomowego Si do ^{12}C oraz gęstości monokryształu ^{28}Si równanie na stałą Avogadra przyjmie postać następującą [19]:

$$N_A = \frac{A_r(\text{Si})M_{\text{mol}}}{\sqrt{8d_{220}^3} \rho(\text{Si})} \quad (6)$$

$M_{\text{mol}} - 1000 \text{ kg mol}^{-1}$, d_{220} – typ płaszczyzny kryształu krzemu, $\alpha_0^3 = \sqrt{8d_{220}^3}$ – jednostka objętości komórki atomowej, ρ – gęstość kryształu krzemu.

Równanie (5) umożliwia wyznaczenie liczby Avogadra, czyli stałej fundamentalnej o wymiarze mikroskopowym, na podstawie pomiaru zrealizowanego dla makroskopowej próbki monokryształu krzemu.

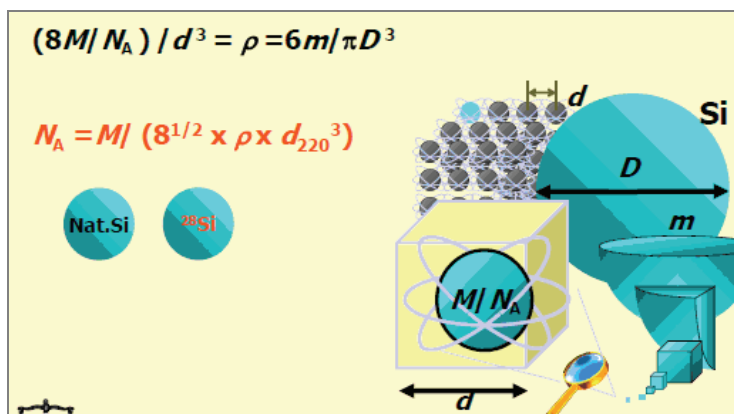
W tym celu trzeba jak najdokładniej wyznaczyć jej gęstość, czyli masę molową i objętość oraz parametr sieci a . Po zastosowaniu zaawansowanych technik spektroskopowych oraz rentgenografii dokładne wyznaczenie stałej Avogadra staje się coraz bardziej osiągalne. Nowa definicja kilograma mogłaby brzmieć następująco:

Kilogram jest masą $5,0184458 \cdot 10^{25}$ (równej $10^3 N_A / 12$) niezwiązanych, pozostających w spoczynku atomów ^{12}C w stanie podstawowym.

W ramach projektu pomiary wartości stałej Avogadra, N_A ustalano według pomiarów masy molowej oraz średnicy d kuli (kryształu) krzemu, zbudowanego z atomów ^{28}Si [12].

Prace nad wyznaczeniem stałej Avogadra realizowane w ramach Projektu IAC na podstawie pomiaru masy molowej kryształu krzemu 28 prowadzone są w kilku ośrodkach NMI. Uzyskiwane uprzednio rezultaty w instytucie PTB [15], czy NRC [16] są

potwierdzone kolejnymi, między innymi przeprowadzonymi w 2014 r. [17] w japońskim NMIJ. W eksperymencie wyznaczono wartość masy molowej ^{28}Si równą: $27,97697009(14) \text{ g mol}^{-1}$ z względną niepewnością standardową $5,2 \cdot 10^{-9}$. Masę molową obliczono pośrednio między innymi na podstawie wyznaczonych techniką spektrometrii MC-ICP-MS (ang. *Multicollector-Inductively Coupled Plasma Mass Spectrometry*) stosunków izotopowych $^{29}\text{Si}/^{30}\text{Si}$ [17]. Masa ta następnie posłuży do obliczenia stałej Avogadra.



Rys. 4. Schemat wyznaczenia stałej Avogadra – projekt Avogadro [12]

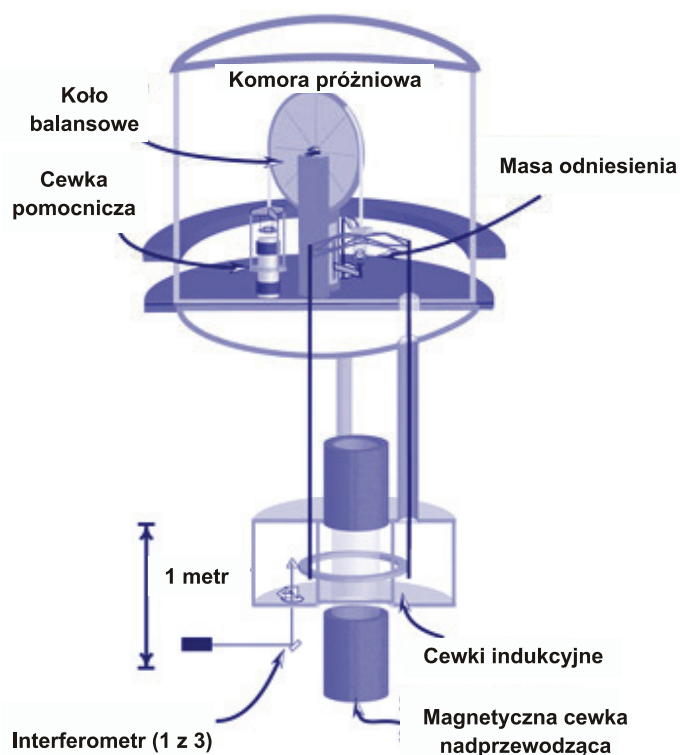
Fig. 4. Scheme of the Avogadro constant determination – Avogadro project [12]

Rezultaty eksperymentów przeprowadzonych w instytutach metrologicznych NIST (USA) i NPL (Wielka Brytania) można zaobserwować w okresie 1990–2012 wyraźną tendencję dążenia do wartości referencyjnej stałej Avogadra $N_A = 6,022\,141\,29(27) \cdot 10^{23} \text{ mol}^{-1}$ (wg CODATA 2010). Obiecujące, zarówno ze względu na uzyskiwaną wartość N_A , jak i deklarowaną niepewność względną, wydają się badania prowadzone w instytucie NIST (USA). Zespół prowadzący eksperyment w PTB również zapowiada działania mające na celu poprawę wyników – obniżenie niepewności względnej do wartości $1,5 \cdot 10^{-8} N_A$. We wnioskach z dotychczasowego eksperymentu zwrócono uwagę na dużą niepewność względną podczas określania jednorodności powierzchni (ok. 22% całkowitego budżetu niepewności) oraz objętości (ok. 55% całkowitego budżetu niepewności) badanego kryształu krzemu. Stąd podjęte działania i plany na lata 2014–2015 powiązane ze sprowadzeniem monolitycznych, wysokiej jakości kryształów krzemu ^{28}Si z Rosji oraz zastosowaniem nowej techniki pomiaru topografii powierzchni – unowocześnionego interferometru [12].

W ramach prac nad kwantową definicją kilograma w Międzynarodowym Biurze Miar (BIPM) oraz kilku wiodących instytutów NMI (w tym NIST, USA) realizowany jest równolegle projekt wykorzystujący prądową wagę Watta. Waga ta – choć, wymyślona znacznie wcześniej, bo w XIX w., przez lorda Kelvina w celu wyznaczenia

wartości ampera, udoskonalona w 1975 r. przez Bryana Kibble'a [14] stała się najodpowiedniejszym przyrządem, dzięki któremu możliwe jest wyznaczenie stałej Plancka, h , a tym samym zdefiniowanie jednostki masy. Jej unowocześniona wersja to aparat wysokości dwupiętrowego budynku, skonstruowany ze złożonego układu elektromagnesów i cewek (rys. 5).

W BIPM projekt budowy kriogenicznej wagi Watta nadal jest w realizacji (wizualizacja projektu wagi BIPM dostępna jest pod następującym adresem: http://www.bipm.org/en/scientific/elec/watt_balance/wb_cryogenic.html).



Rys. 5. Schemat wagi Watta NIST (źródło: <http://www.nist.gov/pml/newsletter/siredef.cfm>)
 Fig. 5. Diagram of the NIST Watt balance (source: <http://www.nist.gov/pml/newsletter/siredef.cfm>)

Wytwarzane w jej wnętrzu pole magnetyczne równoży ciężar jednego kilograma. Aby zwiększyć precyzję pomiaru, badacze umieścili najważniejsze elementy w komorze próżniowej wykonanej z włókna szklanego. Waga ta określa z dużą dokładnością siłę potrzebną do zrównoważenia przyciągania kilogramowego odważnika przez pole grawitacyjne Ziemi.

W pierwszej, statycznej fazie doświadczenia siła grawitacyjnego przyciągania ciężarka o masie m (sprężonego z górną cewką, przez którą płynie prąd stały o natężeniu I_2) jest

równoważona przez odpychającą siłę indukcji wzajemnej w strumieniu pola magnetycznego Φ wytworzonym przez cewkę dolną [11]:

$$mg = I_2 \frac{d\Phi}{dz} \quad (7)$$

W fazie dynamicznej doświadczenia górna cewka dodatkowo oscyluje wokół położenia równowagi z prędkością $v = dz/dt$, w wyniku czego indukowane w niej jest napięcie

$$U = -\frac{d\Phi}{dt} \quad (8)$$

Po przekształceniu otrzymujemy równanie, na podstawie którego wyznaczana jest masa

$$mgv = UI_2 \quad (9)$$

Parametry równania (9) mierzone są z dużą dokładnością i precyzją, co umożliwi wyznaczenie masy z jak najmniejszą niepewnością pomiaru. Przyspieszenie ziemskie g mierzone jest za pomocą grawimetrów z dokładnością względną wynoszącą obecnie 10^{-9} . Dokładność ta może ulec zwiększeniu dzięki zastosowaniu w tym celu przyrządów opartych na interferometrii atomów chłodzonych laserowo [10]. Prędkość v mierzy się za pomocą interferometru, który śledzi wiązkę odbitą od zwierciadła przymocowanego do cewki. Do pomiarów wielkości elektrycznych wykorzystuje się zjawisko Josephsona i kwantowe zjawisko Halla, wobec czego równanie na wyznaczaną masę można zapisać w następujący sposób:

$$m = h \frac{K_{J-90}^2 R_{K-90} UI_2}{4gv} \quad (10)$$

Uzależniając pomiar masy od wartości napięcia i natężenia prądu, na podstawie równania (13) kilogram można zdefiniować, opierając się na stałej fizycznej, zwanej stałą Plancka, h , stałej fizycznej pojawiającej się w większości równań mechaniki kwantowej. Stała h historycznie została wprowadzona przez niemieckiego uczonego, Maxa Plancka, który, prowadząc badania nad promieniowaniem emitowanym przez ciało doskonale czarne, stwierdził, że energia nie może być wypromieniowywana w dowolnych ciągłych ilościach, a jedynie w postaci kwantów o wartości $h\nu$, gdzie ν jest częstotliwością. Stała Plancka w układzie SI wynosi

$$h = 6,626\ 069\ 57(29) \cdot 10^{-34} \text{ J}\cdot\text{s (wg CODATA 2010)}$$

Nowa definicja kilograma oparta na stałej Plancka brzmiałaby:

Masa spoczynkowa ciała, którego energia równoważna jest energii fotonów, których częstotliwości sumują się do $1,356392733 \cdot 10^{50}$ Hz [10].

W celu zwiększenia dokładności pomiaru, konieczne jest jednoczesne zastosowanie kilku wag Watta. Zgodnie z dotychczasowym stanem wiedzy waga Watta będzie stosowana do realizacji jednostki masy oraz badań dryftu masy prototypów najwyższego rzędu, natomiast przekazywanie jednostki będzie odbywać się dalej z użyciem wzorców tradycyjnych platynowo-irydowych oraz w postaci wzorców fizycznych ulepszonych pod względem materiałowym, wzorcowanych za pomocą wagi Watta przez związek ze stałą Plancka [19, 20].

4.3. PRZEKAZYWANIE JEDNOSTKI MASY A REDEFINICJA KILOGRAMA

Z perspektywy kilku lat intensywnych prac nad redefinicją jednostki masy, po analizie uzyskanych do tej pory wyników badań i pomiarów, możemy spojrzeć na kwantową koncepcję definicji kilograma w zupełnie innym świetle. Deklaracje XXIV. Generalnej Konferencji Miar zostały diametralnie zweryfikowane przez środowiska metrologów z całego świata (reprezentantów instytutów metrologicznych, NMI) podczas ostatniego posiedzenia Komitetu Technicznego TC-M EURAMET (ds. Masy i Wielkości Pochodnych), jakie odbyło się w marcu 2013 r. w Dubrowniku w Chorwacji [21]. Na podstawie dotychczas uzyskiwanych rezultatów w projektach Avogadro i wagi Watta oraz opinii użytkowników przyrządów pomiarowych (np. Organizacji OIML) zebrano w formie Rekomendacji CCM G1 uwagi do początkowych założeń definicyjnych Rezolucji 1 XXIV. Konferencji CGPM. Wskazano także, że realizacja definicji kilograma może zostać oparta na stałej fizycznej, ale pod pewnymi warunkami. Najważniejsze z nich to:

- Co najmniej trzy niezależne realizacje projektów wagi Watta otrzymają wartości stałej Plancka ze względną niepewnością standardową nie większą niż $5 \cdot 10^{-8}$.
- Co najmniej jeden z tych wyników osiągnie względną niepewność standardową nie większą niż $2 \cdot 10^{-8}$.
- Prototypy (najwyższego rzędu) kilograma BIPM oraz wzorce 1 kg stosowane ww. projektach będą porównywane bezpośrednio z międzynarodowym prototypem MPK.
- Obecnie opracowywane *mise en pratique* nowego kilograma będzie zawierać procedury porównań kluczowych oraz ich pozytywnej walidacji, np. przez porównania bezpośrednio z MPK.
- Po zakończeniu wymienionych porównań i analizie ich rezultatów, zgodnie z zasadami określonymi w dokumencie CIPM MRA-D-05 Measurement comparisons in the CIPM MRA, zostanie przyjęta określona wartość stałej Plancka.
- Zostanie ustanowiony zestaw wzorców odniesienia BIPM, związany bezpośrednio jak to możliwe z IPK, stanowiący wartość referencyjną w porównaniach kluczowych po redefinicji.
- W Międzynarodowym Biurze Miar i odpowiedniej liczbie instytutów NMI nadal będą prowadzone prace badawczo-rozwojowe w dziedzinie redefinicji

jednostki masy tak, aby na nowo zdefiniowany kilogram mógł być realizowany i przekazywany ze względną niepewnością standardową nie większą niż $2 \cdot 10^{-8}$.

Tabela 3. Przejście do nowego układu jednostek SI [21]

Table 3. Transition to the New SI [21]

Niepewność względna $m_{\text{MPK}}/h = Q$	
Przed redefinicją	Po redefinicji
Niepewność względna $Q = u_r(Q)$	Niepewność względna $Q = u_r(Q)$
Wartość $m_{\text{MPK}} = 1 \text{ kg}$	Wartość $m_{\text{MPK}} = 1 \text{ kg}$
Niepewność względna $m_{\text{MPK}} = 0$	Niepewność względna $m_{\text{MPK}} = u_r(Q)$
Niepewność względna $h = u_r(Q)$	Niepewność względna $h = 0$

Przed redefinicją międzynarodowy prototyp kilograma MPK, o masie dokładnie 1 kg (bez przypisanej niepewności), przekaże jednostkę masy wzorcom transferowym, aby w trakcie porównań kluczowych między pierwotnymi realizacjami (wagi Watta) określić jak najdokładniej wartość liczbową stałej Plancka h z niepewnością względną rzędu $u_r(h)$ wynoszącą 10^{-8} . W momencie redefinicji ustalona zostanie wartość stałej Plancka o niepewności równej zero (rys. 4), a dotychczasową niepewność wyznaczenia stałej Plancka h przejmie MPK. Odtwarzanie i przekazywanie jednostki masy pomiędzy BIPM a instytucjami metrologicznymi NMI odbywać się będzie przez tak zwane realizacje pierwotne (czyli za pomocą wag Watta), monitorowane podczas porównań kluczowych, wykonywanych raz na 5–10 lat. W tych 5–10-letnich okresach wykorzystywany będzie międzynarodowy zestaw wzorców masy, zwany też kompletem wzorców materialnych (PoA – *pool of artefacts*). W celu zachowania ciągłości przekazywania jednostki masy należy przede wszystkim:

- ustalić najdokładniejszą wartość liczbową stałej Plancka h ,
- przyjąć, że wartości masy po redefinicji nie ulegają zmianie [$m_{(\text{przed})} = \text{masa}_{(\text{po})}$],
- po redefinicji przypisać międzynarodowemu prototypowi MPK niepewność stałej Plancka h .

Międzynarodowe Biuro Miar oraz instytucje krajowe NMI mające pierwotne realizacje będą miały wspólnie możliwość niezależnego przekazywania jednostki masy w sposób gwarantujący ciągłość, stosując międzynarodowy zestaw wzorców złożony z 12 wzorców masy [21].

W praktyce pierwotne realizacje definicji kilograma będą oparte na porównaniu mocy elektrycznej z mechaniczną oraz na zastosowaniu metody XRCD. W ramach przygotowań do wdrożenia nowej definicji jednostki masy przewiduje się również opracowanie sposobu jej przekazywania do laboratoriów niższego rzędu z poszczególnych NMI i BIPM (odtworzących jednostkę masy za pomocą pierwotnych realizacji) oraz zaprojektowanie zestawu wzorców masy w odniesieniu do rezultatów uzyskanych w porównaniach kluczowych. Zagadnienia związane z utrzymaniem wzorców masy

oraz przetwarzaniem i zastosowaniem średniej wagowej wzorców masy z danego zestawu wzorców. Niemniej jednak, pomimo prowadzonych badań nad nową definicją i w związku z uzyskiwaniem poprawy niepewności pomiaru stałych Plancka czy Avogadra, prace oparte na poprzedniej definicji kilograma będą kontynuowane [21–23].

5. MOL

5.1. DEFINICJA OPARTA NA ATOMIE WĘGLA 12

Po odkryciu podstawowych praw chemii do określenia liczby pierwiastków lub związków chemicznych zastosowano jednostki nazywane przykładowo „gramoatom” i „gramocząsteczka”. Jednostki te miały bezpośredni związek z „masą atomową” i masą cząsteczkową”, które w istocie są masami względnymi. Masy atomowe, o których mowa były początkowo odniesione do umownie przyjętej masy atomowej tlenu – 16, podczas gdy fizycy oddzielili izotopy tlenu za pomocą spektrometru masowego i przypisali wartość 16 do jednego z jego izotopów, chemicy przypisali tę wartość mieszaninie izotopów tlenu 16, 17 i 18, którą traktowali jako naturalny pierwiastek. Ostatecznie dzięki porozumieniu, jakie nastąpiło w latach 1959–1960 Międzynarodowej Unii Fizyki Czystej i Stosowanej (ang. *International Union of Pure and Applied Physics*, IUPAP) oraz Międzynarodowej Unii Chemii Czystej i Stosowanej (ang. *International Union of Pure and Applied Chemistry*, IUPAC) za wartość odniesienia w chemii przyjęto masę atomową – 12, izotopu węgla (^{12}C), dokładnie określaną jako względną masę atomową $A_r(12\text{C})$. Według tak powstałej zunifikowanej skali podawane masy atomowe i molekularne mają wymiar względny [5].

W 1961 roku Unie IUPAP i IUPAC, podając definicję mola, określiły wielkość fizyczną, której jednostką jest: „liczność materii”. W tej postaci wprowadziła ją XIV. Generalna Konferencja Miar (1971 r.), uznając mol za jedną z podstawowych jednostek miar układu SI [25].

Zgodnie z obowiązującą dotąd definicją: „mol jest to ilość substancji układu, który zawiera tyle samo elementarnych cząstek, ile zawartych jest atomów w 0,012 kg węgla 12”. Cząstki elementarne powinny być określone jako atomy, cząsteczki, jony, elektrony cząsteczki, jony, elektrony, inne cząstki lub ich swoiste konglomeraty w odniesieniu do mola. Stąd masa molowa węgla 12 wynosi dokładnie 12 gramów na mol, czyli $M(^{12}\text{C}) = 12 \text{ g/mol}$.

W 1980 r. Komitet CIPM, zatwierdzając raport Komitetu Doradczego ds. jednostek CCU, który wskazuje w nim, że w definicji mola niezwiązane atomy węgla 12 znajdują się w energetycznym stanie podstawowym i spoczynku.

Definicja mola jednocześnie determinuje wartość uniwersalnej stałej Avogadra, N_A , która dotyczy liczby podmiotów do ilości substancji dla każdej próbki. Stała Avogadra, a właściwie liczba Avogadra została odkryta w 1909 r. przez Jeana Baptiste’a

Perrina, francuskiego fizyka i uczonego, podczas realizacji badań nad ruchami cząsteczkowymi Browna [24]. Zdefiniował on liczbę Avogadra jako liczbę cząsteczek w jednej *gramocząsteczce* tlenu (wówczas dokładnie 32 g) [24]. Dawniej pojęcie *gramocząsteczka* określało ilość substancji/składnika równą jej masie cząsteczkowej wyrażonej w gramach, czyli równoważną dzisiejszemu molowi. Liczba Avogadra była używana jako jednostka ilości materii do października 1971 r., kiedy to XIV. Generalna Konferencja Miar wprowadziła pojęcie mola [25]. Pojęcia: „liczba Avogadra” i „stała Avogadra” były opisywane dawniej jako synonimy [26]. We współczesnej literaturze chemicznej w dalszym ciągu spotyka się pojęcie liczby Avogadra używane w znaczeniu stałej Avogadra [27].

Liczba Avogadra, liczbowo równa liczbie atomów, cząsteczek lub innych cząstek materii zawartych w jednym molu tej materii, ma identyczne oznaczenie i wartość liczbową ze stałą Avogadra, ale w odróżnieniu od niej ma jednostkę.

Stała fizyczna Avogadra, N_A – oznaczana jest przez N_A , a jej wartość wynosi [CODATA 2010]:

$$N_A = 6,02214129(27) \cdot 10^{23}$$

z odchyleniem standardowym $0,00000027 \cdot 10^{23} \text{ mol}^{-1}$ ($2,7 \cdot 10^{16}$).

Jednostkę licznosci materii cechuje bezwymiarowość, podobnie jak jest w przypadku innych jednostek miary licznosci (np. mendel, tuzin czy gros).

Użycie mola ze względu na jego wielkość ogranicza się do zliczania obiektów mikroskopowych. Mol jest jednostką tak dobraną, iż masy atomów oraz cząsteczek podane w jednostkach masy atomowej dają się łatwo przeliczać na masy moli w gramach.

Stała Avogadra występuje w relacji, w której jeżeli $N(X)$ oznacza liczbę podmiotów X w określonej próbce, a $n(X)$ oznacza ilość substancji X podmiotów w tej samej próbce, to [28]:

$$n(X) = \frac{N(X)}{N_A} \quad (11)$$

5.2. PRAKTYCZNA REALIZACJA MOLA

Definicja mola może być realizowana za pomocą różnych metod podstawowych [25, 30]. Między innymi ilość substancji czystej w badanej próbce można wyznaczyć na podstawie pomiaru masy, a następnie dzieląc ją przez masę molową substancji

$$n = \frac{m}{M} \quad (12)$$

Kolejną stosowaną w praktyce realizacją jednostki licznosci materii [28, 30] jest metoda kulometryczna oparta na prawie elektrolizy, sformułowanym przez Michaela Faradaya w latach 1833–1834, zgodnie z którym:

Chemiczne efekty prądu elektrycznego są wprost proporcjonalne do ilości elektryczności, która przechodzi przez układ.

W elektrolizie chemicznej ilość substancji wydzielająca się na elektrodzie jest wprost proporcjonalna do ładunku elektrycznego wytwarzanego w układzie pomiarowym

$$n = \frac{Q}{zF} = \frac{It}{zF} \quad (13)$$

gdzie: n – ilość substancji, mol, Q – ładunek elektryczny, I – natężenie prądu elektrycznego ($I = \text{const}$), A , t – czas elektrolizy, s, z – wartościowość jonu, F – stała Faradaya, C/mol.

W równaniu (13), będącym podstawą obliczeń w analizie kulometrycznej, wszystkie wielkości są znane, oprócz ładunku, który jest charakterystyczną w tej metodzie wielkością mierzoną. Mierząc zatem wielkość ładunku, można łatwo obliczyć masę oznaczanej substancji bez wstępnej kalibracji metody. Kulometria jest metodą bezwzględną, czyli absolutną w odróżnieniu od większości metod instrumentalnych, które są metodami porównawczymi wymagającymi kalibracji z użyciem roztworów wzorcowych. Ponadto kulometria jest metodą uniwersalną, nadaje się do oznaczania zarówno makroilości – podobnie jak metody klasyczne (grawimetria i wolumetria), jak i mikroilości substancji. Z użyciem nowoczesnych instrumentów istnieje bowiem możliwość dokładnego pomiaru małych natężeń prądu (rzędu nanoamperów) oraz czasu (rzędu milisekund).

W przypadku czystych gazów ilość substancji (mol) może być określana na podstawie równania wirialnego stanu gazu (14)

$$pV = nRT \left[1 + B(T) \left(\frac{n}{V} \right) + \dots \right] \quad (14)$$

gdzie: p – ciśnienie, V – objętość, T – temperatura, R – stała gazowa ciśnienia i objętości, B – drugi współczynnik wirialny.

W wielu zastosowaniach dla ciśnienia o wartościach poniżej 1 atmosfery, obliczenie liczności materii dla substancji gazowych możliwe jest na podstawie równania stanu gazu doskonałego Clapeyrona

$$pV = nRT \quad (15)$$

Współczynniki wirialne, B , będące funkcją temperatury, wyrażone w jednostkach SI zostały stabelaryzowane dla wielu prostych gazów. W niskich temperaturach przyjmują wartości ujemne, podczas gdy wraz ze wzrostem temperatury ich wartości szybko zwiększają się. W temperaturze T_B , zwanej temperaturą Boyle'a zmienia się znak współczynnika wirialnego na dodatni i wtedy już znacznie wolniej zwiększa się on wraz z temperaturą.

Molowa stała gazowa R znana ze swej względnej niepewności standardowej $1,7 \cdot 10^{-6}$ (CODATA 2006) w układzie SI ($\text{Pa m}^3 \text{ mol}^{-1} \text{ K}^{-1}$) prowadzi do wyznaczenia liczności materii w molu.

Niepewność pomiarów n wyznaczona w ten sposób uzależniona jest od niepewności pomiaru parametrów p , V i T oraz od korekcji niedoskonałości gazu. Ta metoda realizacji jednostki mola dla czystych substancji gazowych gazu uzależniona jest znacznie od stopnia czystości badanej próbki.

Chociaż $N(X)$ jest parametrem bezwymiarowym, a $n(X)$ definiuje jednostkę układu SI – mola, to stała Avogadra ma jednostkę spójną z SI, powiązaną z molem. Ponadto w definicji mola, czyli określenie „ilość substancji” (ang. *amount of substance*) powinno być zastąpione dla jasności sformułowaniem definiującym konkretny rodzaj substancji, np. „ilość chlorowodoru, HCl” z podaniem jej wzoru empirycznego.

5.3. DEFINICJA MOLA OPARTA NA STAŁEJ AVOGADRA

W przypadku redefinicji mola (powiązanej obecnie ściśle z definicją kilograma) koncepcja definicji opartej na stałej fundamentalnej wymaga przeprowadzenia wielu badań i porównań liczby podmiotów tworzących jeden mol, tj. liczbę *Avogadra*, N_A aby odpowiadała dokładnie wartości jednego grama – Daltona [28]. W nowym ujęciu kwantowym definicja mola brzmiałaby następująco:

Mol jest jednostką liczności materii wyspecyfikowanej elementarnej cząstki, którą może być atom, cząsteczka, jon, elektron, każda inna cząstka lub wyspecyfikowana grupa takich cząstek. Jej wielkość jest oparta na ustalonej wartości liczbowej stałej Avogadra równą dokładnie $6,022\ 14X \cdot 10^{23}$ (mol^{-1}).

W związku z tym mol będzie definiowany w odniesieniu do określonej liczby jednostek danej substancji, co w rezultacie doprowadzi do przerwania powiązania definicyjnego kilograma z molem, a dokładniej relacji między masą atomu ^{12}C – Daltona, kilograma i liczby Avogadro. Jednakże konieczne do zmiany będą zapisy takich relacji:

- masa atomu ^{12}C jest dokładnie 12 Daltonów,
- liczba 1 Daltona w gramach jest dokładnie wartością liczbową stałej Avogadra.

Propozycja redefinicji mola jest częścią procesu redefinicyjnego kilograma. Jest również konsekwencją powiązania definicyjnego tych jednostek w aktualnie obowiązującym systemie SI. Wynika także z potrzeby przeprowadzenia zmian definicyjnych opartych na fundamentalnych stałych fizycznych dla czterech jednostek układu SI, czyli kilograma, kelwina, ampera i mola. Zmiany te umożliwią usankcjonowanie nowego Quantum SI, w którym definicje siedmiu jednostek podstawowych będą realizowane na podstawie zjawisk kwantowych.

Dla wielu badaczy, metrologów problem stanowi brak mierzalnej jednostki dla stałej Avogadra, N_A . Według nich jest ona po prostu bezwymiarowa i powinna pozo-

stać czystą liczbą, a nie stałą fizyczną o wymiarze N^{-1} , gdzie N stanowi licznosc substancji. Ilość substancji jest i zawsze była mierzona jako stosunek innych fizycznych ilości, a nie jako specyficzna, czysta liczba elementarnych podmiotów. Stała Avogadra była zatem mierzona zawsze w odniesieniu do jednostki licznosci materii, a nie odwrotnie [25]. Proponowana redefinicja mola według ustalonej wartości liczby Avogadra, N_A , z pewnością nie wniesie żadnej metrologicznej korzyści, może jedynie wprowadzić zamieszanie na tym polu. Tak więc, według opinii niektórych badaczy, byłoby korzystniej nie zmieniać definicji mola według schematu definicyjnego opartego na stałych fizycznych, zgodnie z którym zmieniono dotychczas kilka podstawowych jednostek międzynarodowego układu SI [29]. Dopuszczalne byłoby jedynie wprowadzenie zmian w definicji w zakresie stosowanych sformułowań i nazw. Sformułowanie „ilość (licznosc) substancji” [28, 29] jest na tyle ogólne, że w przypadku konkretnych cząstek, cząsteczek, sformułowanie to przyjmie postać na przykład „ilość wodoru”, „ilość azotu”, czy „ilość etanu”. Stała Avogadra jest ściśle powiązana z definicją mola i jest stosowana w celu odniesienia poziomu makroskopowego do poziomu pojedynczych atomów przez następującą zależność [29]:

$$M(^{12}\text{C}) = N_A m(^{12}\text{C}), \quad (16)$$

gdzie $m(^{12}\text{C})$ jest masą pojedynczego atomu węgla – izotopu 12. Wprowadzając stałą masę atomową, na podstawie zależności 8 i 9, otrzymamy:

$$m(^{12}\text{C}) = A_r(^{12}\text{C}) m_u \quad (17)$$

$$M_u = N_A m_u \quad (18)$$

Powiązanie stałej Avogadra z masą prowadzi do pośredniej zależności ze stałą Plancka. W aktualnych raportach CODATA i równaniach sformułowanych na podstawie doświadczeń dla wykalibrowanych stałych fizycznych i obserwacji nie pojawia się stała Avogadra. W związku z tym warto jest zweryfikować dopasowanie CODATA zgodnie z nowym ujęciem fundamentalnych stałych układu SI. Punktem wyjściowym podczas określania stałej Avogadra, N_A , jest stała gazowa Rydberga, która wiąże masę elektronu (m_e) ze stałą Plancka, h :

$$m_e = \frac{2hR_\infty}{c\alpha^2} \quad (19)$$

gdzie c – prędkość światła, α – stała struktury subtelnej.

Po przekształceniu równania (18) z równaniem (19) otrzymamy wzór na stałą Avogadra

$$N_A = M_u \frac{A_r(e)c\alpha^2}{2hR_\infty} \quad (20)$$

Względna masa elektronu ($A_r(e)$), gazowa stała Rydberga oraz stała struktury subtelnej wyznaczone są eksperymentalnie z niepewnością względną mniejszą niż 10^{-9} . Niepewność wyznaczenia stałej Avogadra jest taka sama, jak niepewność wyznaczenia masy m_u .

5.4. PRZEKAZYWANIE JEDNOSTKI LICZNOŚCI MATERII A REDEFINICJA MOLA

Nie jest do końca jasne, w jaki sposób nowa definicja mola odniesiona do ustalonej liczby jednostek (cząstek, atomów, cząsteczek) stanie się „pojęciowo prostsza” od obowiązującej definicji opartej na ustalonej wartości masy M_u , równoważnej molowo masie atomu węgla 12. Proponowana redefinicja sugeruje, aby ilość substancji była mierzona przez liczenie elementarnych cząsteczek, podczas gdy jest to metoda całkowicie niepraktyczna. Chemicy będą nadal mierzyć ilość substancji przez stosunki masowe, nauczyciele jednak będą musieli przekonać swych uczniów, że masy oraz stosunki masowe są równoważne „zliczaniu” liczby jednostek rzędu $6 \cdot 10^{23}$ [31–33].

Liczność materii była mierzona znacznie wcześniej sto lat przed pierwszym precyzyjnym wyznaczeniem stałej Avogadra. Tak jak stosowanie ilości substancji nie jest zależne od teorii atomu, jej pomiar nie jest zależny od stałej Avogadra. Pierwsze jednostki opracowane specjalnie w tym celu pojawiły się w 1882 r., czyli 77 lat przed odkryciem Perrina i wyznaczeniem przez niego wartości liczby Avogadra [9] oraz prawie w tym samym czasie, kiedy nastąpił rozwój jednostek elektryczności. Ponieważ nigdy dotąd licznosc materii nie była mierzona przez liczenie jednostek (cząsteczek, atomów, jonów), stosunkowo łatwo będzie wykazać, że będzie to niemożliwe w praktyce w skali makroskopowej. Jest powszechnie akceptowane i potwierdzone, że definiowanie jednostki elektryczności opartej na ładunku elementarnym e , przyczyni się do poprawy precyzji i dokładności pomiarów licznych wielkości fizycznych w Międzynarodowym Układzie Jednostek Miar SI; żadne praktyczne korzyści niezadeklarowane w przypadku ustalonej numerycznej wartości stałej Avogadra. Według opinii niektórych uczonych aktualnie tocząca się debata nad redefinicją kilograma oraz pozostałych kilku podstawowych jednostek układu SI jest doskonałą okazją, aby otwarcie odrzucić redefinicję mola w odniesieniu do ustalonej numerycznej wartości stałej Avogadra [31, 33].

6. KELWIN

6.1. EWOLUCJA POMIARÓW TEMPERATURY

Współczesną termodynamiczną definicję temperatury zawdzięczamy odkryciu Lorda Kelwina, który w 1848 r. zbadał i scharakteryzował stan zera bezwzględnego, przyjmując za model do swych analiz – układ kryształu doskonałego. Temperatura,

będąc wielkością termodynamiczną, w ujęciu mikroskopowym opisuje stan cieplny cząsteczek w ruchu. W stanie zera bezwzględnej zaś, cząsteczki nie poruszają się wcale lub ich drgania ograniczone są do minimum energetycznego.

Temperaturę można ściśle zdefiniować tylko dla stanów równowagi termodynamicznej, ponieważ zgodnie z zerową zasadą termodynamiki temperatura jest wielkością reprezentującą wspólną właściwość dwóch układów A i B, pozostających w równowadze cieplnej z trzecim układem C pełniącym funkcję termometru [34]. Obowiązująca nadal jej termodynamiczna definicja umożliwia porównywanie wartości, ale nie określa skali temperatury. Skala temperatury przyporządkowuje określonej temperaturze. Do określenia skali temperatury potrzebne są stałe punkty termodynamiczne, odpowiadające odtwarzalnym stanom równowagi międzyfazowej, jak np. zmiana stanu skupienia pewnych substancji (początkowo wody). Punktom tym przypisane były i są pewne wartości liczbowe [36].

Pierwsze pomiary temperatury, uwzględniające potrzebę wyznaczenia punktów termometrycznych skali temperatury prowadził Robert Hook w 1664 r. Za standard skali temperatury wybrał on punkt zamarzania wody. Jeden stopień na wprowadzonej przez niego skali odpowiadał zwiększeniu objętości wody o 1/500. Narzędzie pomiarowe zastosowane przez Hooka w opisanym doświadczeniu stanowił termometr pracujący według rozszerzalności objętościowej barwionego alkoholu [36].

Prace nad wyznaczeniem najstarszej skali temperaturowej (do dziś stosowanej w krajach anglosaskich) zawdzięczamy holenderskiemu uczonemu Gabrielowi Danielowi Fahrenheitowi, który w 1715 r. opracował pierwszy, precyzyjny termometr, przyjmując za stałe punkty termometryczne: temperaturę mieszaniny lodu, wody i chlorku amonu (0°F) oraz mieszaniny lodu i wody (32°F), dzieląc przedział tych temperatur na 32 równe części. Skala termometru Fahrenheita (temperatura mieszaniny lodu i wody z salmiakiem lub solą kuchenną, uznana za jej punkt zerowy i temperatura mieszaniny wody z lodem), została uznana i ogłoszona w 1724 r. [36] W 1742 roku A. Celsjusz, astronom i fizyk szwedzki, przyjął temperaturę wrzenia wody za 0° , a temperaturę mieszaniny lodu i wody za 100° , dzieląc ten przedział na 100 równych części. Dopiero w 1850 r. astronom M. Stromer odwrócił te wartości.

Zarówno rtęć, którą stosował Fahrenheit oraz Celsjusz, jak również inne substancje termometryczne umożliwiają określenie skal termometrycznych tylko w określonym zakresie temperatury, ograniczonym wytrzymałością cieplną zastosowanej substancji, elementów konstrukcyjnych termometrów lub możliwością wykorzystania przyjętej właściwości substancji. Na przykład zastosowanie termometru rtęciowego jest ograniczone od dołu krzepnięciem rtęci, natomiast od góry wytrzymałością szkła kapilary. W 1848 roku angielski uczoney W. Thomson (późniejszy lord Kelvin), wykorzystując równanie opisujące sprawność odwracalnego obiegu Carnota, ustalił termodynamiczną skalę temperatury, całkowicie niezależną od wyboru substancji termodynamicznej.

Rozwój dokładnych metod pomiaru temperatury związany był w istotny sposób z postępowaniem nad opracowaniem wzorca jednostki miary długości, realizowanych

w XIX w. w Międzynarodowym Biurze Miar (BIPM) w Sèvres koło Paryża. Każdy z narodowych prototypów metra zaopatrzony był w dwa szklane termometry rtęciowe wykonane według projektu paryskiego konstruktora Tonnelota i wywzorcowane w BIPM. Termometry te zostały wykonane z wysokiej jakości szkła, zapewniającego powtarzalność pomiarów temperatury na poziomie kilku tysięcznych stopnia. Termometry szklane rtęciowe były przedmiotem badań Chapiussa, który w latach 1884–1887 w BIPM za ich pomocą planował dokonać pomiaru temperatury absolutnej. Rozwazał także, aby w tym celu zastosować termometr gazowy o stałej objętości, stosując wodór, azot i dwutlenek węgla jako płyny robocze. Termometry gazowe są nadal stosowane do wyznaczania termometrycznych punktów potrójnych.

Pierwotnie skale temperatury (zawierające charakterystyczne wartości temperatury i odpowiadające im zjawiska określające stan cieplny) były konstruowane na podstawie charakterystycznych wartości temperatury zmian stanów skupienia i po założeniu, że rozszerzalność cieczy jest liniowa. Późniejsze konstrukcje opierano na właściwościach gazów. Współcześnie zaś temperaturę definiuje się przez odwołanie do fizyki statystycznej. Pierwszą Międzynarodową Skalę Temperatur ustanowiono podczas VII. Generalnej Konferencji Miar w 1927 r. Dzięki wprowadzeniu skali temperatury możliwe było pokonanie praktycznych trudności związanych z bezpośrednią realizacją gazową termometrii oraz zastąpienie istniejących wówczas narodowych skal temperatury [35, 36]. Skala wprowadzona została między temperaturą wrzenia tlenu i topnienia złota – ustalono kilka punktów temperaturowych i dwa przyrządy pomiarowe, kalibrowane w tych punktach: oporowy termometr platynowy w zakresie niższych temperatur i termoparę rod-platyna dla wartości temperatury powyżej 660 °C. Powyżej temperatury topnienia złota temperatura zdefiniowana została przez prawo Wiena i mierzona metodą optyczną pirometrem.

Poprawki do tej skali wprowadzono następnie w roku 1948, implementując Międzynarodową Praktyczną Skalę Temperatury 1948 (IPTS-48).

Skala ta utrzymywała punkt potrójny wody jako jedyny punkt definiujący kelwina, jednostkę skali termodynamicznej temperatury.

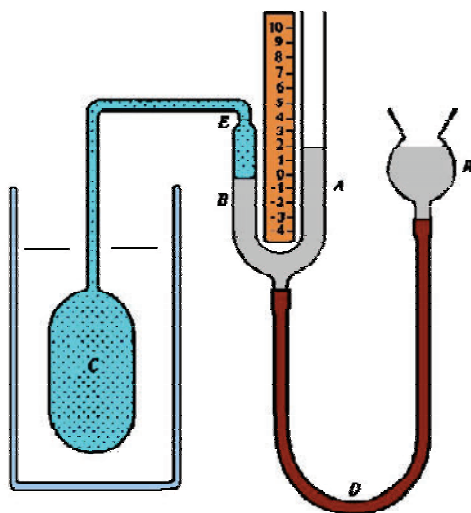
Punktowi potrójnemu wody przypisano zaś 0,01 °C. Skala ta do 1954 roku była oparta na założeniu, że różnica temperatury wrzenia wody i topnienia lodu w warunkach normalnych wynosi 100°, a od 1954 r. skalę termodynamiczną temperatury określono, przyjmując za podstawę jeden punkt termometryczny – temperaturę równowagi wody w stanie gazowym, ciekłym i stałym, tzw. punkt potrójny wody, który leży powyżej temperatury topnienia lodu o wartości 0,01 °C. Punktowi temu przyporządkowano wartość temperatury termodynamicznej równą 273,16 K [36].

6.2. TERMODYNAMICZNA DEFINICJA KELWINA A SKALA TEMPERATURY

Dopiero od 1954 roku, zgodnie z uchwałą X. Konferencji CGPM, podstawową jednostkę temperatury zaczęto definiować na podstawie tezy angielskiego uczonego,

nazywając ją stopniem Kelwina ($^{\circ}\text{K}$). Po niespełna 10 latach XIII. Generalna Konferencja Miar w 1967 r. uchwałą 3 i 4 wprowadziła drobną poprawkę do definicji jednostki temperatury, zmieniając jej nazwę na Kelwin (K). Od tego czasu kelwin, a nie, jak było poprzednio, stopień kelwina równy jest $1/273,16$ temperatury termodynamicznej punktu potrójnego wody [36].

Poza ustaleniem jednostki temperatury ważnym problemem jest odtworzenie skali termodynamicznej. Skalę termodynamiczną temperatury Kelwina można odtwarzać za pomocą termometru gazowego pracującego z użyciem gazu doskonałego. Termometr taki jest kłopotliwy podczas jego stosowania. Przeto dokładne odtworzenie skali termodynamicznej do celów praktycznych jest bardzo trudne. Powstała konieczność stworzenia praktycznej skali termometrycznej, która odpowiadałaby skali termodynamicznej, lecz bez stosowania termometrów gazowych.



Rys. 6. Schemat termometru gazowego
(źródło: <http://fatcat.ftj.agh.edu.pl/~kaprzyk/Fizyka/fizyka/node120.html>)

Fig. 6. Scheme of the gas thermometer
(source: <http://fatcat.ftj.agh.edu.pl/~kaprzyk/Fizyka/fizyka/node120.html>)

W 1989 roku zaadaptowano Międzynarodową Skalę Temperatury 1990 (ITS-90) [4], którą oparto na wielu tzw. stałych punktach termometrycznych, tj. temperaturze punktów potrójnych (współistnienia w równowadze trzech faz: stałej, ciekłej i gazowej dla określonej substancji) lub temperatury wrzenia różnych ciał. Temperatury tych punktów ustalono za pomocą termometru gazowego (rys. 6), np. w punkcie potrójnym wody, tlenu lub punkcie topnienia miedzi dla ciśnienia $101\,325\text{ Pa}$.

Temperaturę w punkcie potrójnym wyznacza się na podstawie równania (21), dokonując pomiaru ciśnienia badanego gazu

$$T = 273,16 \frac{P}{P_p} \quad (21)$$

P_p – ciśnienie w punkcie potrójnym badanej substancji, np. wody (611,73 Pa).

Aktualnie od 1 stycznia 1990 roku obowiązuje nowa skala The International Temperature Scale of 1990 (ITS.90). Została wprowadzona przez Międzynarodowy Komitet Miar w 1989 r. Aktualna skala ITS190 w 2000 r. została rozszerzona w dół, przez adaptację skali The Provisional Low Temperature Scale (PLTS12000), która obejmuje zakres od 0,9 mK do 1 K i jest uznawana za lepsze przybliżenie temperatury termodynamicznej w tym zakresie.

Skalę bezwzględną lub absolutną temperatury określa się na podstawie równania stanu gazu idealnego (równania Clapeyrona)

$$pV = nRT \quad (22)$$

Temperaturę określa się na podstawie pomiaru ciśnienia i objętości gazu idealnego. W skali bezwzględnej zero zdefiniowane jest jako temperatura gazu idealnego, w której będzie on miał zerową objętość dla dowolnego ciśnienia.

W ujęciu mikroskopowym odpowiada to sytuacji, gdy wszelki ruch cząsteczek gazu ustaje. W rzeczywistości jednak objętości cząsteczek gazu są niezerowe, oznacza to, że rzeczywistym termometrem gazowym nie można mierzyć dowolnie niskiej temperatury [36].

Temperatura zera bezwzględnego jest najniższą temperaturą, jaką mogą uzyskać ciała. W temperaturze tej wszystkie cząsteczki mają najmniejszą możliwą energię, według mechaniki klasycznej ruch cząsteczek całkowicie ustaje, a według mechaniki kwantowej wszystkie cząsteczki znajdują się w stanie podstawowym, wykonując tzw. drgania zerowe (ang. *zero-point energy*). W temperaturze zera bezwzględnego wszystkie ciała są w stanie stałym, wyjątkiem jest hel, który w pewnym zakresie ciśnienia pozostaje cieczą w dowolnie niskiej temperaturze [35, 36].

6.3. PRAKTYCZNA REALIZACJA DEFINICJI KELWINA OPARTEJ NA PUNKCIE POTRÓJNYM WODY

Skala Kelvina (skala bezwzględna) jest skalą absolutną, tzn. zero w tej skali oznacza najniższą teoretycznie możliwą temperaturę, jaką może mieć kryształ doskonały, w którym ustały wszelkie drgania cząsteczek. Temperatura ta nie została nigdy zarejestrowana, gdyż praktycznie nie da się jej osiągnąć, lecz obliczono ją na podstawie funkcji uzależniającej temperaturę od energii kinetycznej drgań cząsteczek w kryształach doskonałych. Funkcję tę opracował lord Kelvin; od niego pochodzi też nazwa skali. Temperatura 0 K jest równa temperaturze $-273,15$ °C.

Ponieważ skala Kelvina oparta jest na skali Celsjusza i odległość jednego stopnia jest w obu przypadkach ta sama, temperaturę w kelwinach otrzymujemy przez dodanie

do liczby wyrażonej w stopniach Celsjusza stałej 273,16, zgodnie z następującym równaniem [38, 39]:

$$T = (t + 273,16) \text{ K} \quad (23)$$

gdzie t – temperatura w °C.

W rezultacie praktyczna skala ITS 90 może być uważana za przybliżoną realizację temperatury termodynamicznej. Kelwin i stopień Celsjusza są jednostkami Międzynarodowej Skali Temperatury 1990 (ITS-90), przyjętej przez Komitet CIPM w 1989 r. Rekomendacją 5 (CI-1989) [39, 40].

W 2005 roku istotne wymagania do definicji temperatury, opartej na punkcie potrójnym wody, wprowadził Międzynarodowy Komitet Miar (CIPM) swą Rekomendacją 2, według której woda stosowana do napełniania komórek, powinna mieć skład izotopowy zgodny z wzorcem Vienna Standard Mean Ocean Water (V-SMOW), zalecanej przez IUPAC (ang. *International Union of Pure and Applied Chemistry*):

0,000 155 76 mol ^2H na mol ^1H , 0,000 379 9 mol ^{17}O na mol ^{16}O ,
0,002 005 2 mol z ^{18}O na mol ^{16}O

Rys. 7. Skład izotopowy wody stosowany w realizacji termodynamicznej definicji Kelwina
Fig. 7. Water isotopic composition used in realization of the Kelvin thermodynamic definition

Zapis, mówiący, iż wartość 273, 16 K jest przypisana punktowi potrójnemu wody o określonym składzie izotopowym, pojawił się już znacznie wcześniej w Międzynarodowej Praktycznej Skali Temperatury z 1968 r. (MPST-68) [38]. Był to skład średniej wartości izotopów wody oceanicznej, rekomendowany przez Międzynarodową Agencję Energii Atomowej, oznaczony symbolem SMOW (ang. *Standard Mean Ocean Water*) [39].

Na wartość odtwarzanej temperatury mają wpływ zarówno skład izotopowy, jak i czystość wody oceanicznej. W praktyce stosuje się wodę destylowaną, dejonizowaną o czystości nie mniejszej niż 99,9995%. Punkt potrójny wody do celów wzorcowań realizowany jest w komórkach o konstrukcji pochodzącej jeszcze z lat 30. ubiegłego stulecia (rys. 8), wykonanych z wysokiej jakości szkła borokrzemowego lub kwarcowego.

Składniki komórki również mogą zanieczyszczać wodę, a tym samym wpływać na wartość odtwarzanej przez nią temperatury.

Komórki punktu potrójnego wody zapewniają odpowiednią realizację definicji termodynamicznej temperatury. Niepewność punktu potrójnego wody wynosi w przybliżeniu 50 mK, natomiast w wartościach względnych – około $2 \cdot 10^{-7}$ [39, 40]. Pomiar bezpośrednie temperatur innych niż punkt potrójny wody wymagają zastosowania termometrów podstawowych (pierwotnych) opartych na rozwiązaniach pomiaru innej

wielkości, według której wyprowadzona jest temperatura układu fizycznego. W praktyce termometry podstawowe (najwyższego rzędu) są nader skomplikowane, a pomiar realizowany za ich pomocą jest bardzo czasochłonny i niewystarczający do przekazywania jednostki kelwina. Praktyczną realizację temperatury umożliwiają skale temperatury.



Rys. 8. Komórka punktu potrójnego wody
(źródło” <http://www.alitech.com.pl/index/showproduct/id/180>)

Fig. 8. Cell of the triple point of water
(source: <http://www.alitech.com.pl/index/showproduct/id/180>)

6.4. DEFINICJA KELWINA OPARTA NA STAŁEJ BOLTZMANNNA

Uchwałą nr 1 XXIV Generalnej Konferencji Miar (XXIV. CGPM, Paryż, październik 2011 r.) zarekomendowano poddanie redefinicji podstawowej jednostki temperatury. Wybór stałej Boltzmannna – stałej energii termicznej ($E = k_B T$) jako najefektywniejszego sposobu definiowania temperatury termodynamicznej jest rezultatem wieloletnich prac prowadzonych przez utworzoną w 2005 r. (do realizacji zadań CIPM w zakresie prac nad redefinicją Kelwina) grupę zadaniową TG–SI w ramach Komitetu Doradczego Temperatury oraz grupy roboczej WG4.

Obowiązującą obecnie termodynamiczną definicję Kelwina rekomendacją Komitetu CIPM postanowiono oprzeć na wyznaczeniu stałej Boltzmannna, k_B . Aby zapewnić kontynuację z obowiązującą definicją Kelwina, środowisko ekspertów zajmujących się termometrią postanowiło wyznaczyć stałą Boltzmannna, k_B za pomocą kilku różnych metod pomiarowych [41–43]. Stała k_B zostanie wyznaczona eksperymentalnie po

zastosowaniu: podstawowej termometrii, włączając gazową termometrię akustyczną oraz gazową termometrię stałej dielektrycznej. Niezależnie uzyskane rezultaty mogą być ponadto potwierdzone między innymi przez realizację pomiarów metodą termometrii radiacyjnej oraz termometrii poszerzonej o zjawisko według Dopplera [42, 43] i in. Istotne jest zastosowanie przynajmniej dwóch diametralnie różnych metod pomiaru stałej Boltzmanna.

Po przeprowadzeniu wszystkich zaprojektowanych eksperymentów Komitet Danych dla Nauki i Techniki – CODATA (ang. *Committee on Data for Science and Technology*) – międzynarodowa organizacja z siedzibą w Paryżu, zajmująca się wszystkimi typami danych ilościowych uzyskiwanych w wyniku eksperymentalnych pomiarów i obserwacji w naukach fizycznych, chemicznych, biologicznych, geologicznych i astronomicznych, porówna oraz uśredni wszystkie wyniki pomiarów stałej Boltzmanna za pomocą ich niepewności, podając w rezultacie jej najbardziej prawdopodobną wartość do opinii publicznej. Zgodnie z rekomendacjami Komitetu Doradczego ds. Temperatury (CCT) potrzeba czasu, aby potwierdzić założoną niepewność względną wyznaczenia stałej k_B nie większą niż w przybliżeniu 2×10^{-6} przed ogłoszeniem jej wartości przez Komitet CODATA [38].

Po dokonaniu redefinicji jednostki temperatury wartość niepewności wyznaczenia stałej Boltzmanna k_B będzie wynosić zero, a wartość temperatury punktu potrójnego wody T_{TPW} pozostanie $-273,16$ K. Jednakże niepewność temperatury punktu potrójnego $-T_{TPW}$ będzie zawierać w sobie częściowo średnią ważoną niepewność, ustaloną na podstawie pomiarów k_B przez Komitet CODATA [47].

Stale aktualizowany dokument pt. *Mise and pratique for the definition of the kelvin* już wkrótce będzie zawierał definicję Kelwina opartą na stałej Boltzmanna, k_B , jej dokładną wartość oraz niepewność, z jaką została wyznaczona.

Stała fizyczna o symbolu k_B została zdefiniowana w XIX wieku przez austriackiego fizyka Ludwiga E. Boltzmanna w ramach badań nad kinetyczną teorią gazów, występując w równaniach parametrów określających rozkłady energii cząsteczek, opisana jest wzorem:

$$k_B = \frac{R}{N_A} \quad (24)$$

gdzie: R – stała gazowa, N_A – liczba Avogadra.

Stała Boltzmanna k_B powiązana jest z temperaturą termodynamiczną i energią termiczną. Zgodnie z danymi opublikowanymi przez Komitet CODATA w raporcie z 2010 r. wartość stałej k_B wyniosła $11,380\ 6488(13) \cdot 10^{-23} \text{ J K}^{-1}$, a standardowa niepewność jej wyznaczenia $-0,000\ 0013 \cdot 10^{-23} \text{ J K}^{-1}$.

W ramach fizyki klasycznej (tj. bez uwzględnienia efektów kwantowych) stała Boltzmanna jest czynnikiem proporcjonalności między średnią energią kinetyczną, E_k , przypadającą na pojedynczą cząsteczkę o f stopniach swobody a temperaturą bezwzględną T .

$$E_K = \frac{f}{2} k_B T \quad (25)$$

gdzie: k_B – stała Boltzmana od nazwiska austriackiego fizyka Ludwiga E. Boltzmana, który wprowadził ją w XIX w. w ramach badań nad kinetyczną teorią gazów, T – temperatura bezwzględna, f – stopnie swobody układu.

Po zastosowaniu poprzedniego równania można wyznaczyć temperaturę układu pomiarowego, stanowiącego wzorzec pierwotny na podstawie pomiaru stałej Boltzmana.

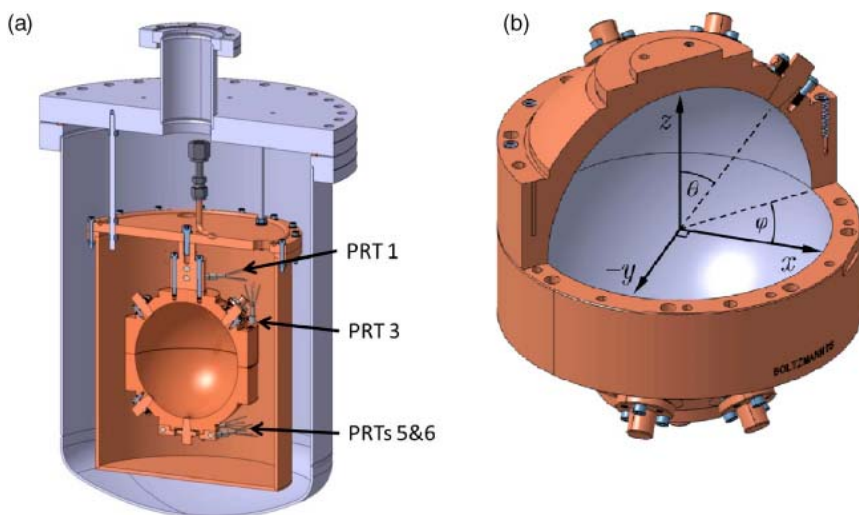
Badania nad wyznaczeniem stałej Boltzmana prowadzone są aktywnie od 2005 roku, czyli od momentu podjęcia pierwszej decyzji dotyczącej redefinicji Kelwina. Prace w tym zakresie prowadzone są dość intensywnie w kilku wiodących instytutach metrologicznych na świecie, czyli w NIST (USA), PTB (Niemcy), NPL (Wielka Brytania), LNE-CNAM (Francja) [43].

Metoda wyboru dokładnego i precyzyjnego wyznaczania stałej k_B jest znana od lat 70. ubiegłego stulecia gazowa termometria akustyczna, dokładnie rozpoznana przy okazji badań nad termodynamiczną temperaturą oraz termofizycznymi właściwościami gazów. Metoda akustyczna opiera się na możliwości dokładnego ustalenia prędkości dźwięku w wypełnionej gazem celi pomiarowej (w kształcie kuli) w stałej temperaturze. Stałą Boltzmana wyznacza się tutaj z równań, kolejno na energię kinetycznej (E_k) oraz na energię cieplną gazu ($E = 1/2 m v^2 = 3/2 k_B T$), w którym m jest masą pojedynczego atomu, v oznacza średnią prędkość proporcjonalną do prędkości dźwięku w gazie, a T oznacza temperaturę). Prędkość dźwięku może być mierzona przez analizowanie częstotliwości fal dźwiękowych rezonujących w sferycznej celi pomiarowej. Pomiar częstotliwości realizowany jest z dużą precyzją. De Podesta i współpracownicy mają w swoim laboratorium zegar rubidowy, który mierzy częstotliwość z dokładnością do 1 części na 10¹³.

Wadą metody jest konieczność dokładnego i precyzyjnego zwymiarowania celi pomiarowej, tak jak dokonali tego Moldover i współpracownicy w swoim przełomowym eksperymencie w 1988 r. w laboratorium NIST, wypełniając celę pomiarową (kulę) 3 litrami rtęci, a następnie ją zważyli. W ten sposób dokonali pomiaru stałej Boltzmana z niepewnością równą 1,7 · 10⁻⁶ [44].

W jednej z ostatnich swoich prac badawczych nad stałą Boltzmana de Podesta i współpracownicy opracowali precyzyjny termometr służący do realizacji definicji temperatury opartej na stałej Boltzmana, k_B to rezonator akustyczny zaprojektowany w brytyjskim instytucie metrologicznym NPL (ang. *National Physical Laboratory* w Teddington). De Podesta i współpracownicy skonstruowali rezonator z miedzi (wypełniony gazem obojętnym – argonem, wyposażony w wiele anten i mikrofonów, rys. 9), za pomocą którego wyznaczana jest wartość stałej Boltzmana. Celem badań i pomiarów realizowanych na rezonatorze akustycznym jest powiązanie temperatury z energią za pomocą stałej k_B . Nowy termometr to pod względem technicznym „rezonator akustyczny”, który zaczyna „dzwonić”, kiedy wzbudzamy w nim drgania

o pewnych częstotliwościach. Badając rezonans za pomocą mikrofonów, uczeni potrafią wyznaczyć prędkość dźwięku w wypełnionym gazem zbiorniku, a tym samym średnią prędkość atomów argonu, czyli ich energię kinetyczną. Dotychczasowe rezultaty pomiarów stałej Boltzmann, k_B , de Podesta i współpracownicy opisali w pracy [45].



PRT – platynowy termometr rezystancyjny

Rys. 9. Przekrój opracowanego w NPL rezonatora akustycznego typu Cranfield NPLC-2, zaopatrzonego w izotermiczną obudowę przy zewnętrznym ciśnieniu 25 L (źródło: http://iopscience.iop.org/0026-1394/50/4/354/pdf/0026-1394_50_4_354.pdf, [45])

Fig. 9. Cross-section through the apparatus showing the NPL–Cranfield resonator NPLC-2 suspended within an isothermal enclosure within an outer 25 L pressure (source: http://iopscience.iop.org/0026-1394/50/4/354/pdf/0026-1394_50_4_354.pdf, [45])

Metoda ta polega na wyznaczeniu dwóch zależności matematycznych. Pierwsza wiąże średnią kwadratową prędkość RMS masy atomu z jego energią kinetyczną (prowadząc w rezultacie do temperatury termodynamicznej)

$$\left(\frac{1}{2}\right)mv^2_{RMS} = \left(\frac{3}{2}\right)k_B T \quad (26)$$

Druga zaś zależność matematyczna wiąże średnią prędkość RMS z szybkością częstotliwości zerowej dźwięku c_0 tego samego gazu oraz stosunku jego pojemności cieplnej $C_p^0/C_v^0 \equiv \gamma_0$. Dla jednoatomowego gazu γ_0 wynosi dokładnie 5/3.

$$\left(\frac{1}{2}\right)mv^2_{RMS} = \left(\frac{3}{\gamma_0}\right)c_0^2 \quad (27)$$

Metodą rezonatora akustycznego wyznaczana jest wartość C_0 przez połączenie częstotliwości uzyskanych w pomiarach cylindrycznego elementu z wnęką szczeliny wypełnionej gazem obojętnym: argonem lub helem przy różnych wartościach ciśnienia w temperaturze punktu potrójnego wody z zachowaniem odpowiednich wymiarów wnęki rezonatora.

W efekcie pośrednio wyznaczana jest stała Boltzmanna, k_B :

$$k_B = \frac{c_0^2 M}{(T \gamma_0 N_A)} \quad (28)$$

gdzie: M – średnia masa molowa gazu wyznaczona na podstawie chemicznej analizy izotopowej, N_A – stała Avogadra.

Na podstawie wykonanych pomiarów z zastosowaniem quasi-cylindrycznych (jako najefektywniejszych) uzyskano dwa najlepsze wyniki z niepewnością wyznaczenia stałej k_B wynoszącą $1,2 \cdot 10^{-6}$ oraz $0,71 \cdot 10^{-6}$, potwierdzając w ten sposób poprzednio deklarowane rezultaty [42]. Zastosowanie delikatnie zdeformowanych cylindrycznych wnęk rezonatora, elipsoidalnych, eliminuje zjawisko degeneracji trybów rezonatora.

Jako alternatywę dla akustycznych pomiarów naukowcy PTB zaproponowali całkowicie metodę eliminującą źródła błędów systematycznych – termometrię gazową opartą na stałej dielektrycznej (ang. *Dielectric Gas Thermometry* DCGT) [47, 49]. Metoda ta polega na określeniu gęstości gazu obojętnego – helu za pomocą pomiaru pojemności kondensatora. Z pomiarów przeprowadzonych w punkcie potrójnym wody o różnych ciśnieniach na kondensatorze pomiarowym stałą Boltzmanna można określić za pomocą podstawowych zależności matematycznych.

Dokonanie dokładnych pomiarów ciśnienia rzędu 7 MPa wymagało nie lada wysiłku i było możliwe dzięki zastosowaniu czujników tłokowych z dokładnością do jednej milionowej, a pomiarów pojemności – z dokładnością do jednej części na miliard. Wymagana stabilność temperatury została zrealizowana dzięki zastosowaniu dużej łaźni termostaticznej, która została specjalnie zaprojektowana i zoptymalizowana we współpracy z Narodowym Instytutem Metrologii Włoch (INRIM). Opracowana w PTB w ten sposób konfiguracja umożliwia wykonywanie pomiarów metodą DCGT w punkcie potrójnym wody, dostarczając wartość stałej Boltzmanna k_B równą $1,380655 \cdot 10^{-23}$ J/K, z niepewnością wynoszącą 8 ppm. Dowodząc jednocześnie, że metoda DGCT może być wykorzystywana do określania stałej k_B . Uzyskanie niepewności rzędu 2 ppm stanowi nadal nie lada wyzwanie. Prace w tym kierunku są nadal prowadzone [49].

Zgodnie z rekomendacjami CIPM niepewność standardowa wyznaczenia stałej Boltzmanna k_B nie może przekroczyć wartości $1 \cdot 10^{-6}$. Jak uprzednio wspomniano, pomiaru stałej k_B można dokonać metodą podstawowej termometrii gazowej (z wykorzystaniem gazów obojętnych, takich jak azot, argon czy hel) lub różnych jej wariantów oraz termometrii radiacyjnej, hałasu lub wykorzystującej efekt Dopplera. Dalej zaprezentowano wartości względne niepewności wyznaczenia stałej Boltzmanna [41]

za pomocą dostępnych metod termometrii podstawowej i narzędzi pomiarowych w nich stosowanych.

Tabela 4. Wartości niepewności względnej wyznaczenia stałej Boltzmanna [8]
Table 4. Relative uncertainty values of the Boltzmann constant determination [8]

Metody termometrii podstawowej	Niepewność względna wyznaczenia stałej k_B
Akustyczny termometr gazowy	$1 \cdot 10^{-6}$
Termometr do pomiaru dielektrycznej stałej gazowej	$2 \cdot 10^{-6}$
Termometr gazowy indeksu refraktywności	$10 \cdot 10^{-6}$
Termometr hałasu Johnsona	$5 \cdot 10^{-6}$
Termometr z poszerzonym efektem Dopplera	$10 \cdot 10^{-6}$

6.5. REDEFINICJA KELWINA A MIĘDZYNARODOWA SKALA TEMPERATURY

Nowa definicja jednostki temperatury będzie niezależna od właściwości substancji, realizacji technicznej i zakresu temperatury. Ułatwi to rozwój termodynamicznych metod pierwotnego pomiaru temperatury (ang. *primary thermometry*), wpłynie na podniesienie dokładności skali temperaturowej i pomiarów temperatury, a w tym pomiarów radiacyjnych. Nowa definicja umożliwi pomiary temperatury ze standardową niepewnością względną na poziomie 10^{-7} . Ponadto implementacja kwantowej definicji Kelwina nie wprowadzi zmian do Międzynarodowej Skali Temperatury (ITS-90) [41, 46].

7. AMPER

Elektryczność i magnetyzm znalazły się w obszarze zainteresowań naukowców stosunkowo niedawno. Nieświadomie ich przejawy w otoczeniu człowiek obserwuje od zarania dziejów w postaci wyładowań atmosferycznych, zorzy polarnej, elektryczności statycznej itp. Znano zjawisko przyciągania naelektryzowanych ciał (np. bursztynu, od którego greckiej nazwy powstało pojęcie elektryczności), ryby broniące się za pomocą wyładowań elektrycznych, iskrzenie wyczesywanej przędzy, szklane kule magów, w których tliła się delikatna poświata. Jednakże dopiero w XVII w. zainteresowali się tymi zjawiskami naukowcy, szczególnie skupieni w brytyjskim Towarzystwie Królewskim, pod patronatem I. Newtona. Odkrywali nowe zjawiska, starali się nad nimi zapanować, odtwarzać je. Prezentowali szklane kule, w środku których tliła się delikatna poświata, przeskakujące iskry, przyciąganie materii naładowanej elektrostatycznie. Eksperymenty te balansowały na granicy świata nauki i alchemii, kuglarstwa. Fascynowały publiczność, ale ich istota stanowiła tajemnicę, ukrytą za kurtyną świata bogów – wczesnogreckich, pogańskich, chrześcijaństwa. Jednak dzięki nim powstała butelka lejdejska, odkryto właściwości materii umożliwiające podzielenie jej na przewodniki i izolatory,

powstał piorunochron Franklina, w końcu – pojawiło się pojęcie prądu elektrycznego. Przełom wieków XVIII i XIX stanowił złoty okres rozwoju tej nauki, gdy odkryto i sformułowano jej najważniejsze prawa – Coulomba, Kirchhoffa, Ohma; swoich odkryć dokonali A. Volta, A. Ampere, M. Faraday. Pojawiły się pierwsze wynalazki umożliwiające praktyczne wykorzystanie elektryczności, jej wytwarzanie, magazynowanie i wykorzystanie energii w niej zgromadzonej. Sformułowano, w postaci znanej do dziś, prawa rządzące nią oraz uzyskano możliwość jej opisu ilościowego i jakościowego. Ostatecznie od prostej obserwacji nieznanych zjawisk, uznawanych za przejaw absolutu, boskości, do teorii elektromagnetyzmu droga nie była daleka.

Zrodziło się pytanie, jakimi wielkościami fizycznymi charakteryzować prąd i jego przepływ, które z nich są wielkościami pierwotnymi, które z kolei najbardziej nadają się do praktycznego zastosowania, obliczeń inżynierskich. Stosując skrót myślowy, można stwierdzić, że źródłem prądu jest ładunek elektryczny q . Przepływ prądu jest charakteryzowany wartością natężenia elektrycznego. Napięcie elektryczne, stanowiące pochodną pojęcia siły elektromotorycznej źródła, wiąże ze sobą ładunek i pracę potrzebną do jego przeniesienia.

Tabela 5. Wielkości elektryczne – jednostki i ich definicje
Table 5. Electrical quantities – units and their definitions

Wielkość fizyczna	Jednostka	Definicja	
Natężenie prądu elektrycznego, I	amper, A	$I = \frac{dq}{dt}$	q – przepływający przez przekrój poprzeczny przewodnika ładunek elektryczny t – czas
Napięcie elektryczne, U	volt, V	$U_{AB} = \frac{W_{A \rightarrow B}}{q} = \int_A^B \vec{E} d\vec{l}$	W – praca konieczna do przesunięcia ładunku E – natężenie pola elektrycznego
Rezystancja, R	ohm, Ω	$R = \rho \frac{l}{S}$	ρ – rezystywność materiału przewodzącego l – długość przewodnika S – pole przekroju przewodnika

Odkryto również prawo Ohma, charakteryzujące środowisko przewodzące prąd elektryczny, równocześnie wiążąc ze sobą natężenia prądu z napięciem elektrycznym.

Wszystkim wielkościom fizycznym, umożliwiającym stworzenie podstawowego, matematycznego modelu przepływu prądu elektrycznego w obwodzie, przypisano jednostki. Ich nazwy pochodzą od nazwisk ojców elektryczności, którzy przyczynili się do sformułowania wymiarowanych nimi wielkości fizycznych. W tabeli 5 przedstawiono związek między tymi wielkościami fizycznymi oraz ich jednostki wraz z definicjami.

Pojawiło się zaraz pytanie, które z tych wielkości są pierwotne, podstawowe, charakteryzujące naturę zjawiska, a które wtórne, zależne od pozostałych, opisujące jedynie środowisko? Pierwsza myśl jest oczywista – wszystko wywodzi się od ładunku

elektrycznego. Przecież prąd elektryczny nie jest niczym innym niż przemieszczającym się ładunkiem. Stwierdza to wprost prawo Ampera.

Z kolei siła elektromotoryczna źródła, określająca możliwej do wygenerowania energii elektrycznej, również zależy od wielkości zgromadzonego w nim ładunku, powodującego przepływ prądu. Jednostka miary charakteryzująca ładunek elektryczny, kulomb, okazała się z wielu przyczyn niepraktyczna jako podstawowa jednostka układu SI. Trzeba pamiętać, że układ SI nie jest układem miar stworzonym na potrzeby nauki, do opisu ilościowego podstawowych zjawisk fizycznych. Powstał w celach bardziej przyziemnych, aby zaspokoić potrzebę ujednoczenia jednostek miar w handlu i życiu codziennym, ułatwić komunikację międzynarodową w technice. Jest też ściśle związany z potrzebą stworzenia uniwersalnych wzorców jednostek miar, umożliwiających ich odtworzenie. Odnosi się to również do wzorców niższych rzędów, stosowanych w życiu codziennym. To nauka służy potrzebom układu SI i codziennych potrzeb biznesu i techniki, nie odwrotnie. Dlatego za jednostkę podstawową układu została wybrana jednostka natężenia prądu elektrycznego – amper. Znajduje ona praktyczne zastosowanie w opisie ilościowym przepływu prądu, ponadto określa wprost ilość przepływającego przez obwód ładunku.

Stan wiedzy i techniki nie daje możliwości stworzenia wzorca Kulomba. Konieczny do tego jest obiekt, przyrząd wytwarzający ładunek o ściśle określonej wielkości. Pierwotnym wzorcem mógłby stać się przyrząd odwołujący się wprost do ładunku elementarnego, ładunku elektronu, wynoszącego $1\,602\,176\,53\,(14) \cdot 10^{-19}\text{ C}$ [50].

Uzyskanie ściśle określonego ładunku jest, delikatnie mówiąc, kłopotliwe do dziś. Toczą się obecnie prace eksperymentalne nad stworzeniem tzw. pompy elektronowej, SET (ang. *single electron pump*) [51, 52], umożliwiającej uzyskanie ściśle określonej ilości ładunku. Możliwe jest to dopiero dziś, po wykorzystaniu najnowszych odkryć z dziedziny fizyki ciała stałego, fizyki kwantowej, budowy materii, a także osiągnięć techniki.

Pompa elektronowa nie spowoduje jednak rewolucji w układzie SI i zmiany jednostki podstawowej z ampera na kulomba. Prowadzi jedynie do opracowania nowego wzorca ampera, odnoszącego się bezpośrednio do bezwzględnej stałej fizycznej, jaką jest ładunek pojedynczego elektronu. Wiążąc przepływ ściśle określonej prawami fizyki ilości ładunku z czasem, odtwarzanym za pomocą wzorca atomowego, stworzony zostanie nowy wzorec jednostki natężenia prądu elektrycznego. Wzorec ten oparty będzie na podstawowych stałych fizycznych – ładunek elektronu $e = 1,602\,176\,53\,(14) \cdot 10^{-19}\text{ C}$ wyznaczony z niepewnością względną $2,2 \cdot 10^{-8}$, pomiar czasu możliwy dzisiaj z niepewnością sięgającą nawet 10^{-21} [50].

Dotychczas układ SI definiował ampera jako:

Stały prąd elektryczny o natężeniu 1 A jest prądem, który płynąc w dwóch równoległych, prostoliniowych, nieskończenie długich przewodach o znikomym przekroju kołowym, umieszczonych w próżni w odległości 1 m od siebie, spowodowałby wzajemne oddziaływanie przewodów na siebie z siłą równą $2 \cdot 10^{-7}\text{ N}$ na każdy metr długości przewodu [39].

Definicja taka nastęrczała dużo kłopotów w praktycznej realizacji wzorca jednostki miary. Realizacja wzorca opartego wprost na definicji wymagałaby zastosowania wielu założeń upraszczających, a co za tym idzie, stosowania współczynników przeliczeniowych umożliwiających wyznaczenie wartości ampera odniesionego do ograniczeń rzeczywistego wzorca.

Klasycznym tradycyjnym wzorcem prądu jest waga prądowa. Jej zasada działania polega na równoważeniu siły pochodzącej od odważnika materialnego przez siłę wytwarzaną przez układ cewek.

Dokładność takiego wzorca, opartego na koncepcji lorda Kelvina z drugiej połowy XIX wieku, nie była wystarczająca dla potrzeb techniki. Źródłem niedokładności pomiaru były:

- wartość współczynnika c wyznaczanego na podstawie geometrii cewek,
- przyjętej wartości przyspieszenia ziemskiego g , zależnej od położenia wzorca na kuli ziemskiej,
- niemożliwości zniwelowania, praktycznego bądź przez zastosowanie odpowiednich współczynników korekcyjnych, wpływu zewnętrznych pól magnetycznych.

Oczywiście istnieje teoretyczna możliwość wyznaczenia współczynników korygujących, umożliwiających uniezależnienie odtwarzanej wartości jednostki natężenia prądu elektrycznego od parametrów układu pomiarowego. Jednakże zależności te są bardzo złożone i w praktycznych warunkach trudne do zastosowania. Idea zbudowania wzorca ampera opartego na klasycznej wadze prądowej upadła, szczególnie w sytuacji, w której kalibratory prądowe umożliwiały uzyskanie znacznie lepszej dokładności jednostki [54].

Zastosowanie wagi prądowej czyniło wzorzec jednostki natężenia prądu wzorcem wtórnym, odniesionym do jednostki masy – kilograma. Było to sprzeczne z ideą uniezależnienia od siebie wzorców podstawowych jednostek miar SI, które powinny być oparte wprost na zjawiskach fizycznych, do których odnosi się ich definicja bądź do niezależnych podstawowych stałych fizycznych.

W związku z opisanymi problemami w zbudowaniu wzorca ampera przez długie lata wartość jego wyznaczana była metodami pośrednimi z wykorzystaniem prawa Ohma, w odniesieniu do znacznie prostszych w odtworzeniu jednostek Ohma i Volta.

Rezystancja jest właściwością materii, przewodnika zastosowanego w obwodzie prądu elektrycznego. Zależy od właściwości elektrycznych zastosowanego przewodnika – jego geometrii i rezystywności. Wzorzec rezystancji był wzorcem materialnym, grupowym, do którego przez system wzajemnych wzorcowań i porównań odnosiły się wzorce niższych rzędów.

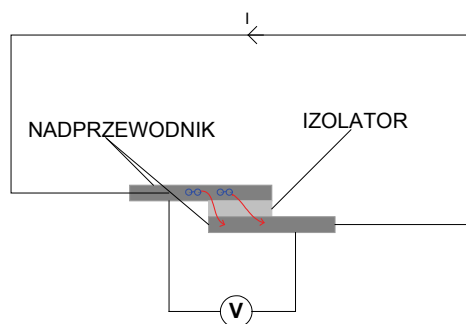
Napięcie elektryczne odtwarzano również za pomocą grupowego wzorca materialnego, które stanowiła grupa ogniów Westona, opracowane w 1880 r. Zachodzące w ogniwie kadmowym zjawisko elektrolizy wytwarza siłę elektromotoryczną o wartości względnie stałej 1,018 V. Oprócz ograniczonej dokładności, główną wadą była

duża niestabilność podczas transportu i wrażliwość na wstrząsy. Powodowało to, że praktycznie ogniwa Westona mogły być wykorzystywane jedynie jako wzorce stacjonarne. W przypadku transportu czas ich stabilizacji nie mógł być krótszy niż 7 dni. Oprócz tego ogniwa Westona są bardzo wrażliwe na zmiany temperatury. Częściowo trudności te wyeliminowało pojawienie się wzorców elektronicznych, opartych na diodach Zenera. Wzorce te w istocie są półprzewodnikowymi stabilizatorami napięcia, które również nie przybliżają do odtwarzania jednostki w odniesieniu do podstawowych stałych fizycznych bądź wykorzystując definicję jednostki. Mogą być wykorzystywane jedynie jako wzorce wtórne o wysokich parametrach metrologicznych.

Wymienione wzorce są wzorcami materialnymi. Cały czas w środowisku metrologów krążyło zagadnienie możliwości uniezależnienia się od wzorców materialnych, odtwarzanie wartości ampera w odniesieniu do stałych fizycznych. Stało się to możliwe pod koniec ubiegłego stulecia. Odkrycia w zakresie fizyki kwantowej, nadprzewodnictwa (zjawisko Josephsona) i fizyki półprzewodników (kwantowy efekt Halla) ułatwiły odniesienie wartości ampera do podstawowych stałych fizycznych.

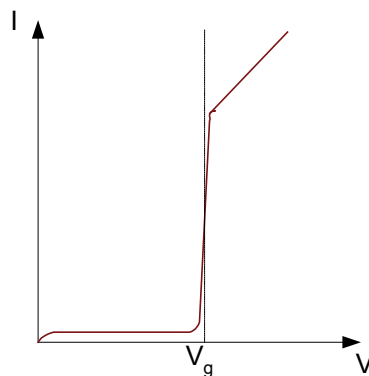
Współcześnie stosowany wzorec jednostki miary napięcia elektrycznego, volta, działa według zjawiska Josephsona [55, 56] występującego w materiałach nadprzewodzących. W 1987 r. Generalna Konferencja Miar zaleciła stosowanie układów do odtwarzania jednostki napięcia opartych na złączach Josephsona, a Międzynarodowy Komitet Miar ustanowił dokładne wartości stałych fizycznych, według których wyznaczana jest jednostka volta [57].

Efekt Josephsona, odkryty w 1962 roku, polegający na tunelowaniu par Coopera przez ciekłą warstwę nienadprzewodzącą, rozdzielającą warstwy nadprzewodnika (rys. 10) [59].



Rys. 10. Układ elektryczny zawierający złącze Josephsona

Fig. 10. Electric circuit with Josephson junction



Rys. 11. Charakterystyka I–V złącza Josephsona

Fig. 11. I–V characteristic of Josephson junction

W tego typu złączu tunelowym w temperaturze zera bezwzględnej prąd tunelowy popłynie dopiero po przekroczeniu wynikającego z charakterystyki materiału napięcia

granicznego złącza V_g , odpowiadającego energii potrzebnej do rozerwania par Coopera. W praktycznej realizacji, w temperaturze wyższej od 0 K, nawet przy niskiej wartości przyłożonego z zewnątrz napięcia następuje przepływ przez złącze prądu o małej wartości.

Po przekroczeniu wartości V_g następuje widoczny skok wartości prądu tunelowego. Charakterystyki prądowo-napięciowe złącza tunelowego wykonanego z materiałów przewodzących i nadprzewodzących pokazano na rys. 11.

Jeżeli do tego złącza Josephsona zostanie przyłożone napięcie stałe o wartości V , można zaobserwować oscylacje par Coopera pomiędzy warstwami nadprzewodzącymi. Częstotliwość tych oscylacji jest określona jako

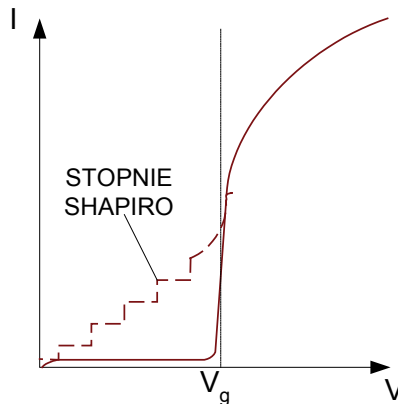
$$f = \frac{2e}{h}V \quad (29)$$

Stosunek $K_{J-90} = 2e/h$, określony nazwą stałej Josephsona, jest stały i wynosi $483\,597,925 \cdot 10^9$, Hz/V. Skrajnie duża częstotliwość i mała amplituda oscylacji czyni zjawisko trudne do zaobserwowania.

Poddając złącze Josephsona działaniu promieniowania mikrofalowego na charakterystyce prądowo-napięciowej złącza, obserwuje się charakterystyczne stopnie napięciowe – stopnie Shapiro (rys. 12) [50]. Wysokość każdego ze stopni można przedstawić zależnością (30)

$$V = \frac{f}{K_{J-90}} \quad (30)$$

będącą przekształceniem równania (1).



Rys. 12. Charakterystyka I–V złącza Josephsona poddanego wpływowi promieniowania mikrofalowego
Fig. 12. I–V characteristic of Josephson junction with microwave radiation

Dla określonego zasilania złącza poddanego promieniowaniu mikrofalowemu możliwe jest uzyskanie napięcia o ściśle określonych wartościach, zależnych jedynie od wartości stałych fizycznych oraz częstotliwości promieniowania mikrofalowego, jakiemu poddane jest złącze.

System pracujący na podstawie złącza Josephsona nie jest fizyczną realizacją jednostki napięcia elektrycznego, a jedynie bardzo stabilną metodą odtwarzania dokładnej wartości napięcia. Pojedyncze złącza Josephsona, co łatwo zauważyć podstawiając wartość stałej Josephsona oraz potrzebne do zaobserwowania promieniowanie o częstotliwości mikrofalowej, wytwarzają napięcie wyjściowe rzędu 1–10 mV. Współcześnie stosowane są matryce zawierające 2000–20 000 złącz Josephsona, zawierające do kilkudziesięciu tysięcy złącz na jednej płytce, co umożliwia generowanie napięcia o dowolnej wartości. W praktyce w laboratoriach wzorcujących odtwarzane są napięcia 1 V, 1,018 V i 10 V [60–67].

Zastosowanie nadprzewodnikowych złącz tunelowych umożliwiło odtworzenie jednostki napięcia elektrycznego w bezpośrednim odniesieniu do podstawowych stałych fizycznych – stałej Plancka i wartości ładunku elementarnego. Spełnione zostało zatem założenie metrologów o możliwości bezpośredniego odniesienia wzorców jednostek miar wielkości elektrycznych do podstawowych stałych fizycznych. Jednak wciąż nie jest to fizyczna realizacja jednostki miary ampera, stanowiącego podstawową jednostkę układu SI, a jedynie sposób odtworzenia jednostki pochodnej – volta. Korzystając z prawa Ohma, w celu umożliwienia odtworzenia jednostki miary natężenia prądu elektrycznego niezbędne stało się znalezienie sposobu na uniezależnienie wartości jednostki miary rezystancji od wzorców materialnych. W sukurs temu przyszło odkrycie przez Klitzinga kwantowego efektu Halla [68–70].

Kwantowy efekt Halla występuje w przypowierzchniowym obszarze złącz półprzewodnikowych w temperaturze kriogenicznej, np. 0,3 K dla heterostruktury GaAs – AlGaAs, i obecności silnego pola magnetycznego, o indukcji do 10 T. Warunkiem do wystąpienia zjawiska jest położenie dna pasma przewodnictwa jednego z materiałów półprzewodnikowych heterostruktury poniżej pasma przewodnictwa drugiego półprzewodnika. Wówczas domieszkowane nośniki z donorów przechodzą do pierwszego z materiałów. Z drugiej strony siły kulombowskie działające między elektronami przewodnictwa a zjonizowanymi domieszkami drugiego półprzewodnika pułapują nośniki w warstwie międzypowierzchniowej złącza. Należy przy tym pamiętać, że dla wystąpienia kwantowego efektu Halla niezbędny jest dobór właściwych, odpowiednio domieszkowanych, materiałów półprzewodnikowych złącza. Związane jest to z koniecznością wytworzenia odpowiednich położeń stanów przewodnictwa, a więc – lokalizacji trójkątnej studni potencjałów. W temperaturach kriogenicznych nośniki nie mają energii umożliwiającej im wydostanie się ze studni potencjału do pasma przewodnictwa któregoś z półprzewodników, natomiast nośniki donorów zostają zamrożone. Nośniki znajdujące się w studni potencjału są jedynymi nośnikami swobodnymi układu, przewodnictwo heterostruktury więc jest związane tylko z nimi. Jeżeli

wartość energii Fermiego E_F znajduje się w studni potencjału, konduktancja G_{xy} w kierunku prostopadłym do przepływu zewnętrznego prądu zasilającego wynosi

$$G_{xy} = ec \frac{\partial n_s(E_F)}{\partial B} \quad (31)$$

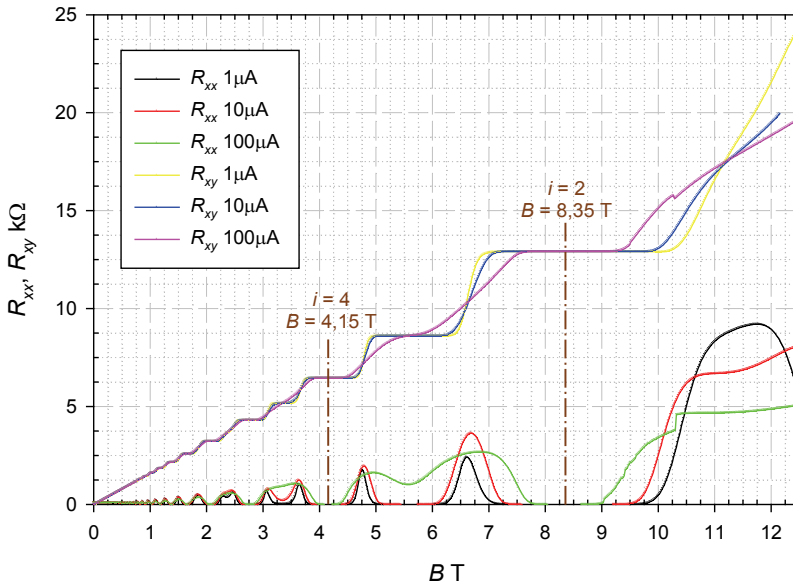
gdzie: e – ładunek elementarny, c – prędkość światła, B – wartość przyłożonego pola magnetycznego, n_s – dwuwymiarowa gęstość ładunku w przypadku pełnego obsadzenia i poziomów energetycznych, z równoczesnym pozostawieniem pustych innych pasm energetycznych

$$n_s(E_F) = \frac{ieB}{hc} \quad (32)$$

Po podstawieniu równania (32) do (31) otrzymujemy zależność określającą rezystancję Halla w kierunku prostopadłym do przyłożonego do niej zewnętrznego prądu I_x''

$$R_{xy} = \frac{h}{qei} = \pm \frac{h}{e^2 i} = \pm \frac{R_K}{i} \approx \pm \frac{25812,807}{i} \Omega \quad (33)$$

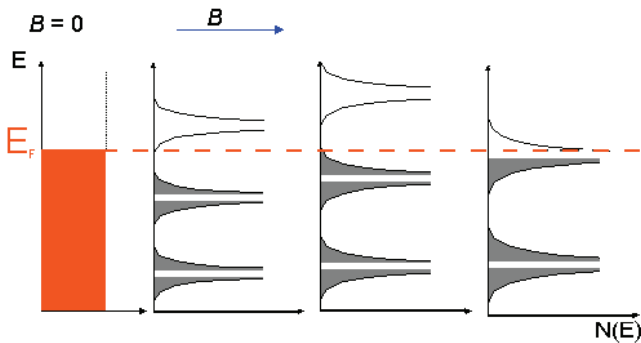
gdzie: h – stała Plancka, e – ładunek elektronu, $q = \pm e$ dla dziur lub elektronów, i – liczba całkowita, $R_K = h/e^2$ – stała von Klitzinga.



Rys. 13. Charakterystyka R - B próbki Halla
Fig. 13. R - B characteristic of quantum Hall sample

Rezystancja R_{xy} przyjmuje wartości skwantowane, zgodnie z wykresem (rys. 13). W kierunku równoległym do prądu transportu rezystancja R_{xx} jest zależna od geometrii próbki i gęstości powierzchniowej nośników prądu oraz staje się niemierzalnie mała, gdy R_{xy} osiąga wartości określone kolejnymi stopniami kwantyzacji.

Występowanie lokalizacji stanów energetycznych wyjaśnia kwantyzację rezystancji R_{xy} próbki Halla przy zmianach pola magnetycznego. Nośniki obsadzające stany zlokalizowane poruszają się w ograniczonym obszarze, nie wnosząc wkładu do prądu transportu. Jeżeli energia Fermiego znajduje się na poziomie stanów zlokalizowanych, nośniki obsadzają stany zlokalizowane bez wpływu na wartość R_{xy} . Stany zdelokalizowane poniżej energii Fermiego są zajęte i nie zachodzi rozpraszanie elastyczne, a więc rezystancja R_{xx} zanika.



Rys. 14. Rozkład gęstości stanów $N(E)$ dwuwymiarowego gazu elektronowego w polu magnetycznym o indukcji B – obszar ciemniejszy pokazuje stany zlokalizowane

Fig. 14. Density of $N(E)$ states of 2D electron gas in magnetic field B
– darker area shows localised states

Po osiągnięciu przez energię Fermiego poziomu odpowiadającego stanom zdelokalizowanym nośniki kierują się ku brzegom próbki i powodują zmianę napięcia V_{xy} , odpowiadającego rezystancji R_{xy} . Spowodowane jest to obsadzeniem przez nośniki wolnych stanów zdelokalizowanych. Liczba takich nośników jest równa liczbie stanów zdelokalizowanych poniżej poziomu Fermiego. Wówczas następuje przejście pomiędzy kolejnymi poziomami R_{xy} , oraz wzrost rezystancji R_{xx} , spowodowany rozpraszaniem nośników swobodnych [71, 72].

Dzięki odkryciu i zastosowaniu kwantowego efektu Halla możliwe stało się odniesienie wartości jednostki Ohma do podstawowych stałych fizycznych przez stałą Klitzinga. Jej wartość, jak pokazano wcześniej, zależy wyłącznie od wielkości stałej Plancka i ładunku elementarnego. Dzięki najnowszym odkryciom fizyki, możliwe stało się odtworzenie pochodnych jednostek układu SI, charakteryzujących przepływ prądu elektrycznego w materii w bezpośrednim odniesieniu do podstawowych stałych fizycznych, dlatego możliwe stało się stworzenie zespołu układów pomiarowych,

ułatwiających odniesienie wzorców prądowych – kalibratorów prądu – i źródeł prądowych do podstawowych stałych fizycznych.

Sytuacja taka musiała zrodzić konieczność rewizji definicji ampera. Okazało się, że tradycyjna definicja ampera, przyjęta wraz z utworzeniem układu SI, nadal nie może zostać wykorzystana do budowy fizycznego wzorca tej jednostki. Możliwość odniesienia jednostki do podstawowych stałych fizycznych spowodowała dyskusję nad redefinicją ampera, a także wzmożenie prac nad możliwością stworzenia bezpośredniego wzorca tej jednostki miary, opartego na zjawiskach kwantowych.

Korzystając z prawa Ohma, możemy więc zapisać, że:

$$I = \frac{U}{R} = \frac{f}{K_J} \cdot \frac{i}{R_K} = \frac{fh}{2e} \cdot \frac{ie^2}{h} = fi \frac{e}{2} \quad (34)$$

Z przekształcenia (34) wynika, że teoretycznie możliwe jest wyznaczenie wartości natężenia prądu elektrycznego jedynie w odniesieniu do ładunku elektronu. Pozostałe parametry są zależne od nastaw układu pomiarowego, przy czym częstotliwość f można odnieść do stałych fizycznych przez atomowy zegar czasu. Oczywiście stworzenie układu pomiarowego opartego na podanej zależności jest bardzo skomplikowane. Teoretycznie, po wyznaczeniu rezystancji próbki wzorcowej za pomocą układu odwrotzowania jednostki rezystancji według kwantowego efektu Halla (ang. *quantum Hall resistor* – QHR), możliwe jest podanie na nią napięcia elektrycznego uzyskanego za pomocą matrycy złącz Josephsona. W takiej sytuacji przez przewodnik o określonej rezystancji powinien popłynąć prąd określony zależnością (34), którego wartość można odnieść do podstawowych stałych fizycznych. Praktyka jednak wykazuje, że stworzenie układu pomiarowego opartego na opisanej zasadzie przekracza możliwości współczesnej techniki.

8. PRAWNOADMINISTRACYJNA PROBLEMATYKA SYSTEMÓW MIAR¹

– *Wielmożny Panie, kupcy nasi wrzawę na rynku wzniecili!*
– *A z jakiejże to przyczyny?*

¹ W zamierzeniu autora (WP) niniejszej części jest dążenie do ukazania problematyki miar w ujęciu systemowym, tj. takim, w którym dające się zidentyfikować elementy związane z pomiarami są powiązane różnorodnymi relacjami oraz umiejscowione są w ustrukturalizowanej całości. Podejście takie powinno zatem ukazać miary w perspektywie ogólniejszej niż postrzegane przez pojęcia takie, jak „układ wielkości” (por. VIM: PKN-ISO/IEC Guide 99 – p. 1.3) czy „układ jednostek miar” (por. VIM: PKN-ISO/IEC Guide 99 – p. 1.13).

– Powiedziano nam, że przybysze z Lewantu domagają się używania ich własnych miar. Gdyby im na to pozwolić, to nasi kupcy duże straty poniosą.

– A cóż to może znaczyć?

– Jak już zapewne wiesz, Wielmożny Panie, nasi kupcy mają prawo odkupić od każdego, kto przybył z karawaną do naszego grodu, przynajmniej 12 łokci jedwabiu i 24 łokci innych tkanin. Ale ich miary są krótsze od naszych, na Ratuszu wymierzonych, aż o 1 ćwierć, czyli 6 cali. Nie godzi się zatem, aby takim oto sposobem, nasi kupcy, a także nasz gród, i Twoja Wielmożny Panie szkatuła, uszczerbku doznały!

– Jeśli to prawda, nakazuję zatem, aby wszystkich tych, którzy miary nieprawomysłne stosować zamierzali, w kuny do pręgierza w rynku przykuć i wychłostać, miary ich spalić na stosie, a nadto – 1/3 ich towarów na rzecz grodu przekazać², a samych tych szubrawców z grodu wygnać!...

[zmyślone: WP.]

8.1. KSZTAŁTOWANIE SIĘ DAWNYCH SYSTEMÓW POMIAROWYCH

Współczesny rozwój układów jednostek miar ma swoją genezę w różnorodnych przejawach aktywności Człowieka, od początku kształtowania się Cywilizacji.

Jak powszechnie wiadomo, rozwój społeczeństw dokonywał się we wszystkich niemalże warunkach geograficznych i klimatycznych. Warunki te w niektórych miejscach łagodniejsze dla ludzi, a przez to dogodniejsze dla ich życia i rozwoju, z kolei w innych – surowsze, a przez to zubożające kulturę materialną i spowalniającą postęp społeczno-gospodarczy, w oczywisty sposób wpływały na narzędzia i technologie pracy. Wszędzie tam, gdzie dane społeczności dość wcześnie wykształciły wyspecjalizowane techniki wytwarzania i obrotu towarami, pojawiły się systemy pomiarowe odpowiednio dostosowane do zróżnicowanych obszarów zastosowań.

Warto przy tym dodać, że najstarsze sposoby mierzenia i wyrażania wartości mierzonych wielkości związane były z właściwościami zmysłowymi (zdolnościami postrzegania) oraz budową anatomiczną ludzi.

Na pierwszy plan wybija się rolnictwo, gdzie informacjami istotnymi dla ówczesnych „systemów zarządzania terenami” były m.in. niemetryczne miary powierzchni upraw:

- Ilości lub czasu pracy, związane ze stosowanymi technologiami, a co za tym idzie wydajnością pracy, mierzone takimi geometrycznymi jednostkami miar, jak np.:
 - **arpent** (*are-pennis*), powierzchnia ziemi zaoranej przez jeden pług, dwa woły albo jednego konia i jednego człowieka, jednostka miary stosowana na terenie Galii; 1 **arpent** \approx [180 stóp francuskich kwadratowych] \approx 3400 m²;

² Częściowe prawo składu [przyp. aut.].

- **morga** (*morgen*), powierzchnia ziemi zaoranej albo skoszonej przez jeden zaprzęg i jednego człowieka w ciągu jednego dnia, jednostka miary stosowana głównie na terenach ziem niemieckich i polskich; 1 **morga** (polska) $\approx 0,56$ ha, tj. około 5600 m².
- Ilości wysiewanego ziarna lepiej charakteryzującą wartość i efektywność gospodarczą obszaru upraw, zwłaszcza na terenach górzystych i słabiej zaludnionych, mierzone takimi jednostkami miar, jak np.: **setier** (przedrewolucyjna Francja), który, „reprezentując” jednakową pojemność ziarna, ze względu na jakość i konfigurację pola uprawnego, wysiewany był do gruntu o powierzchni od 1 **arpent** (gleba dobra) do 1,5 **arpent** (teren górzisty, nieurodzajny, piaszczysty) lub polski **korzec** (ok. 120 litrów), lub polski **garniec** (ok. 3,8 litra)³.

O pozostającej nadal jako aktualnej złożoności problematyki pomiarów w ocenie ekonomicznej efektywności gospodarowania gruntami rolnymi może świadczyć na przykład współczesny system podatku rolnego w Polsce. System ten oparty obecnie na ustawie z dnia 15 lutego 1984 r.⁴ przewiduje, że podstawę opodatkowania podatkiem rolnym stanowi dla gruntów gospodarstw rolnych – liczba **hektarów przeliczeniowych**⁵, ustalana na podstawie powierzchni, 2 rodzajów i 10 klas użytków rolnych wynikających z ewidencji gruntów i budynków oraz zaliczenia do 1 z 4 okręgów podatkowych (art. 4 ust. 1 pkt 1 ww. ustawy).

Specjalny udział i znaczenie miały systemy miar regulujące handel produktami, których dominujące i trwałe właściwości fizyczne determinowały tworzenie stosownej dla danego towaru „miary”. Najczęściej wykorzystywanymi cechami, a zarazem najbardziej optymalnymi, z perspektywy ekonomii i technologii pomiaru, były masa (waga), objętość i długość.

W pomiarach masy kluczowe znaczenie miały jej wzorce (odważniki), których nominal, kształt, materiał, z którego były wykonane oraz oznaczenia – były przedmiotem regulacji władzy państwowej⁶. Szczególną właściwością funkcjonalną antycznych odważników było powiązanie względnej stałości ich masy z systemem monetarnym. Często nazwy jednostek masy były tożsame z nazwą monet, na przykład szekel

³ „...ugodzili się, iż Wojciechowi Stecowi w części swojej pozwalają zasiewać gruntu pod korzec jeden” [1].

⁴ Por. zob.: ustawa z dnia 15 listopada 1984 r. o podatku rolnym (Dz.U. z 2013 r. poz. 1381 i z 2014 r. poz. 40), zastępująca wcześniejsze: ustawa z dnia 26 października 1971 r. o podatku gruntowym, dekret z dnia 30 czerwca 1951 r. o podatku gruntowym, ustawa z dnia 28 czerwca 1950 r. o podatku gruntowym, ustawa z dnia 3 czerwca 1947 r. o nadzorze nad wymiarem i poborem podatku gruntowego [przyp. aut.].

⁵ Należy w tym miejscu zauważyć, że ww. ustawa przewiduje także „hektary fizyczne” wynikające z ewidencji gruntów i budynków, które mają zastosowanie w gruntach rolnych niespełniających normy obszarowej gospodarstwa rolnego [przyp. aut.].

⁶ W do dziś zachowanych zabytkach kultury materialnej starożytnej Grecji niektóre odważniki mają napisy, poświadczające tytuł urzędnika, który dany odważnik sprawdził i gwarantował jego dokładność [2].

(8,36 g) albo drachma (4,5–6 g), funt (327 g)⁷, albo mina (60 szekeli, tj. ok. 500 g albo 100 drachm, tj. 450–600 g) czy talent (60 min, tj. 27–36 kg albo 100 funtów, tj. 32,7 kg)⁸. Wyznaczanie masy towarów odbywało się na wagach prostodźwigniowych, których pierwowzorem były wyciągnięte ramiona z porównywanymi przedmiotami i nosidła wsparte na barkach [3]. Zjawiskiem charakterystycznym dla ukształtowania się w tamtych czasach systemów miar (funkcjonujących do chwili obecnej w niektórych krajach anglosaskich) jest różnorodność rozmiaru (wymiaru) tak samo nazywającej się jednostki miary, np. „lżejszy” funt aptekarski lub jubilerski (ang. *apothecary or troy*), odpowiadający masie 373, 24172 g przeznaczony do pomiaru masy przedmiotów cenniejszych czy „cięższy” funt handlowy (ang. *avoirdupois*), odpowiadający masie 453, 59237 g, przeznaczony do użytku w handlu pozostałymi towarami.

O ile oczywiste było zastosowanie miar objętości do towarów płynnych, to miary takie były także stosowane do rozliczania obrotu towarów o właściwości ciała stałego, ale przez swoje bardzo małe rozmiary geometryczne – o właściwości materiałów sypkich (np. ziarna zbóż). Zamierzone oszustwa lub przypadkowe błędy w jednostajności i powtarzalności napełniania wzorców nasypu⁹ spowodowały w ciągu XIX w. zmianę tej metody pomiarowej na wyznaczanie masy ciał sypkich z użyciem wag i odważników [73].

8.2. WPŁYW ROZWOJU NAUKI I TECHNIKI NA EWOLUCJĘ SYSTEMÓW MIAR

Zasady fizyczne i rozwiązania konstrukcyjne systemów miar mających zastosowanie w gospodarce i handlu pozostawały od zarania dziejów praktycznie niezienne aż do przełomu wieków XVIII i XIX, kiedy to rozpoczął się żywiołowy rozwój nauki i techniki.

Za podwaliny tego rozwoju należy uznać dorobek wieku XVII, zwanego „wiekiem przyrządów naukowych”, takich jak: termometr, barometr, teleskop soczewkowy, teleskop zwierciadlany, mikroskop, higrometr, pompa próżniowa, zegar wahadłowy, maszyna elektrostatyczna, elektroskop itd. [76]. Nowe teorie naukowe odnoszące się do wyjaśniania otaczającej nas rzeczywistości czy koncepcje fizykalne świata antycznego uzyskały, szczególnie w epoce oświecenia, możliwości coraz to doskonalszej weryfikacji za pomocą doświadczeń i eksperymentów, a wynikająca z nich potrzeba

⁷ W okresie Cesarstwa Rzymskiego jednostką mniejszą od funta (*librea*) była uncja (*unciae*), stanowiąca 1/12 funta, tj. ok. 27 g, nb. współczesna brytyjska 1 uncja handlowa odpowiada masie 28,35 g [przyp. aut.].

⁸ Najstarsze wzorce masy oparte były na wielokrotności masy ziaren pszenicy lub jęczmienia, późniejsze także z wykorzystaniem wody – np. w niektórych rejonach świata antycznego masa 1 talentu była równoważna masie wody o objętości 1 stopy, tj. ok. 27 kg [przyp. aut.].

⁹ Staropolskimi miarami objętości dla ciał sypkich były: łaszt (tj. ok. 3840 litrów), korzec (30 korcy = 1 łaszt), garniec (960 garncy = 1 łaszt), kwarta (3 840 kwarty = 1 łaszt), kwaterka (7 680 kwatek = 1 łaszt), natomiast dla cieczy były: beczka (tj. ok. 288 litrów), garniec (jw.), kwarta (jw.) i kwaterka (jw.) [75].

wykonywania coraz to dokładniejszych pomiarów inspirowała projektowanie i budowę nowych przyrządów pomiarowych. Niejednokrotnie uczeni tamtych czasów byli jednocześnie twórcami nowych teorii naukowych oraz twórcami nowych rozwiązań technicznych, umożliwiających bądź weryfikację hipotez już sformułowanych, bądź pobudzających do tworzenia nowych koncepcji. Nadal aktualną wydaje się zatem maksyma przypisywana Galileuszowi: „Zmierz to, co da się zmierzyć i spraw, aby to, czego nie da się zmierzyć, dało się zmierzyć”.

Można więc powiedzieć, że w wieku XVIII nastąpiło trwale ukształtowanie się podstaw naukowych powstawania i rozwoju systemu metrycznego, a także zdeterminowania tego rozwoju w kierunku oparcia systemu miar na stałych uniwersalnych. Świadczyć o tym mogą formułowane wówczas teorie określane obecnie jako prawa lub zasady fizyczne, jak np. prawo powszechnego ciężenia Newtona (1687), zasady dynamiki Newtona (1687), równanie Bernoulliego (1738), prawo Coulomba (1785), prawo pojemności elektrycznej Volty (1784), prawo Ampera (1820), prawo Avogadra (1811), prawo Ohma (1825), prawa elektrolizy Faradaya (1832–1834), prawo Joule’a (1840), równania Maxwella (1864) itd. Wymowny jest fakt, że nazwiska autorów tych teorii posłużyły jako nazwy ważnych dla systemu metrycznego jednostek miar.

W 1832 r. C.F. Gauss (ur. 30 kwietnia 1777 r. w Brunzshwiku, zm. 23 lutego 1855 r. w Getyndze) zdecydowanie wsparł zastosowanie systemu metrycznego jako spójnego systemu jednostek miar w dziedzinie fizyki. Był pierwszym, który wykonał pomiary bezwzględne ziemskiego pola magnetycznego na podstawie dziesiętnego systemu 3 jednostek miar mechanicznych: milimetra, grama i sekundy. Wykorzystanie systemu metrycznego w obszarze elektryczności i magnetyzmu zostało rozwinięte przez J.C. Maxwella (ur. 13 czerwca 1831 r. w Edynburgu, zm. 5 listopada 1879 r. w Cambridge) i J.J. Thomsona (ur. 18 grudnia 1856 w Manchesterze, zm. 30 sierpnia 1940 r. w Cambridge; nagroda Nobla w dziedzinie fizyki w 1906 r.), którzy w ramach swych prac w Brytyjskim Stowarzyszeniu dla Rozwoju Nauk (ang. *British Association for the Advancement of Science* – BAAS, obecnie BA) opracowali kryteria spójnego systemu jednostek miar, podstawowych i pochodnych. W 1874 r. BAAS wprowadził system jednostek miar CGS (centymetr, gram, sekunda), wraz z ich wielokrotnościami (od mikro- do mega-). Ze względu na istotną trudność wykorzystania systemu SGC w obszarze elektryczności i magnetyzmu, w 1980 r. BAAS i Międzynarodowy Kongres Elektryczny (ang. *International Electrical Congress*)¹⁰ zaaprobowali wzajemną spójność zestawu praktycznych jednostek miar, włączając w to om (rezystancja), wolt (siła elektromotoryczna) i amper (prąd elektryczny).

Nowy rozdział aktywności naukowej, ukierunkowanej na bezpośredni rozwój systemu metrycznego, rozpoczął się niemalże nazajutrz po proklamowaniu w dniu 20 maja 1875 r. Konwencji Metrycznej, kiedy to powołane zostały: Międzynarodowy Komitet Miar (CIPM) i Międzynarodowe Biuro Miar (BIPM).

¹⁰ Poprzednik Międzynarodowej Komisji Elektrotechnicznej (IEC) [przyp. aut.].

Na przełomie wieków XIX i XX nastąpiło wyraźne zróżnicowanie dziedzin naukowych fizyki. Świadectwem tej swoistej dyferencjacji może być struktura programowa I Międzynarodowego Kongresu Fizyków w Paryżu (6–12 sierpnia 1900 r.), która obejmowała prace w następujących sekcjach [76]:

- Zagadnienia ogólne i metrologia,
- Mechanika i fizyka cząsteczkowa,
- Optyka i termodynamika,
- Elektryczność i magnetyzm,
- Magnetoptyka, promienie katodowe, uranu itd.,
- Fizyka kosmiczna,
- Fizyka biologiczna.

W ramach I sekcji omówiono, m.in. takie zagadnienia, jak:

- Dokładność pomiarów w metrologii (Benoit),
- Laboratoria narodowe (Pellat),
- Przegląd propozycji układów jednostek (Guillame),
- Pomiar interferometryczny w metrologii (Macé de Lépinay),
- Prędkość dźwięku (Violle),
- Skale termometryczne (Chappuis),
- Postęp w pirometrii (Barus),
- Stała grawitacji (Boys).

Wśród innych zagadnień dyskutowanych w pozostałych sekcjach na uwagę zasługują:

- Wyznaczanie stałych krytycznych (Mathias, Sekcja 2),
- Prędkość światła (Cornu, Sekcja 3),
- Ciśnienie światła (Lebediew, Sekcja 3),
- Prędkość fal elektrycznych (Blondlot i Gutton, Sekcja 4),
- Teoria zjawisk magnetoptycznych (Lorentz, Sekcja 5),
- Nowe pierwiastki promieniotwórcze i ich promieniowanie (M. i P. Curie, Sekcja 5),
- Stała słoneczna (Crova, Sekcja 6),
- Analiza widmowa w fizyce biologicznej (Hénocquee, Sekcja 7).

Kardynalną w swej istocie cezurą w rozwoju nauki, a co za tym idzie – w uformowaniu obecnej koncepcji oparcia jednostek miar na stałych fizycznych, należałoby uznać dzień 14 grudnia 1900 r. [76], kiedy to Max Karl Ernst Ludwig Planck ukończył artykuł¹¹ dot. teorii promieniowania emitowanego przez ciało doskonale czarne i zmodyfikował zapis prawa Wiena przez wprowadzenie do jego wzoru nowej stałej fizycznej (stałej Plancka). Dzień ten uznaje się za początek mechaniki kwantowej.

¹¹ *Annalen der Physik*, IV. Folge 4, (1901), 564–566 [76].

Innymi faktami oddziałującymi na obecnie kształtujący się system miar są rezultaty prac, m.in., B.D. Josephsona (zjawisko Josephsona, 1962, nagroda Nobla w 1973 r.) czy K. von Klitzinga (kwantowy efekt Halla, 1980, nagroda Nobla 1985 r.). Warty przypomnienia jest także fakt przyznania w 1907 r. amerykańskiemu fizykowi Albertowi A. Michelsonowi (ur. 19 grudnia 1852 r. w Strzelnie na Kujawach) nagrody Nobla w dziedzinie fizyki. Nagrodę tę otrzymał on za konstrukcję precyzyjnych instrumentów optycznych (m.in. interferometru) i pomiary w dziedzinie spektroskopii.

Obecnie mamy do czynienia z globalnym już, a także ogromnym w swej skali, zaangażowaniem wielu ośrodków naukowych i instytucji¹² zarówno w sferze badań podstawowych, jak też aplikacyjnych, mogących mieć zastosowanie w rozwoju układu wielkości i jednostek miar.

Wśród wielu publikacji powstałych w wyniku takich prac, są m.in. takie, jak:

- Tuninsky V.S., 1999, *Unit systems based on the fundamental constants*, *Metrologia*, 36 (1999), 9–14.
- Steiner R.L., Williams E.R., Newell D.B. and Liu R., 2005, *Towards an electronic kilogram: an improved measurement of the Planck constant and electron mass*, *Metrologia*, 42 (2005), 431–441.
- Becker P., De Bièvre P., Fujii K., Glaeser M., Inglis B., Luebbig H. and Mana G., 2006, *Considerations on future redefinitions of the kilogram, the mole and of other units*, *Metrologia*, 44 (2007), 1–14.
- Mohr P.J., 2007, *Defining units in the quantum based SI*, *Metrologia* 45 (2008), 129–133.
- Milton M.J.T., Williams J.M. and Forbes A.B., 2010, *The quantum metrology triangle and the redefinition of the SI ampere and kilogram; analysis of a reduced set of observational equations*, *Metrologia*, 47 (2010), 279–286.
- Mana G., Massa E. and Predescu M., 2012, *Model selection in the average of inconsistent data: an analysis of the measured Planck-constant values*, *Metrologia*, 49 (2012), 492–500.
- de Podesta M., Underwood R., Sutton G., Morantz P., Harris P., Mark D.F., Stuart F.M., Vargha G. and Machin G., 2013, *A low-uncertainty measurement of the Boltzmann constant*, *Metrologia*, 50 (2013), 354–376.

¹² Na przykład europejskich instytucji metrologicznych (NMI), współpracujących w ramach kolejnych programów ramowych Wspólnoty Europejskiej (*European Research Area in Metrology – MERA, Implementing the Metrology European Research Area – iMera / iMERA-PLUS, European Metrology Research and development Programme – EMRP, European Metrology Programme for Innovation and Research – EMPIR*), które realizują, m.in. takie projekty badawcze, jak: *Avogadro and molar Planck constants for the redefinition of the kilogram, Foundations for a redefinition of the Ampere, Determination of the Boltzmann constant for the redefinition Kelvin, Optical clocks for the redefinition second, Towards quantum based photon standards, The Watt balance route towards a new definitions of the kilogram.* etc.; vide: <http://www.euramet.org/index.php?id=research-empir> [przyp. aut.].

- Steiner R., 2012, *History and progress on accurate measurements of the Planck constant*, Rep. Prog. Phys., 76 (2013), 016101, 46.
- Schurr J., Ahlers F. and Kibble B.P., 2012, *The ac quantum Hall resistance as an electrical impedance standard and its role in the SI*, Meas. Sci. Technol., 23 (2012), 124009, 20.
- Nawrocki W., 2006, *Revising the SI: the joule to replace the kelvin as a base unit*, Metrology & Measurement System, Vol. XV (2006), No. 2, 171–183.
- Lipiński L., Szmyrka-Grzebyk A., 2008, *Proposals for the new definition of the kelvin*, Metrology & Measurement System, Vol. XV (2008), No. 2, 227–235.
- Nawrocki W., 2010, *The quantum SI – Towards the New System of Units*, Metrology & Measurement System, Vol. XVII (2010), No. 2, 139–150.

8.3. UNIFIKACJA I GLOBALIZACJA SYSTEMÓW JEDNOSTEK MIAR

Pierwsze postulaty uniwersalizacji (globalizacji) układu wielkości i jednostek miar znaleźć można w pracach Johna Wilkinsa, członka Royal Society, który w rozdziale VII swego dzieła pt. *An Essay towards a Real Character And a Philosophical Language* (Londyn, 1668) pisze on, że oczywiście nie obyłyby się bez trudności sprowadzenie miar stosowanych przez różne narody do jednej powszechnie znanej.

Około 1641 roku Stanisław Pułdowski (1597–1645), profesor prawa Akademii Krakowskiej, uczeń Galileusza, prowadząc badania nad łokciem krakowskim, w celu ustalenia jego stosunku do stopy rzymskiej i do wzorców innych miast polskich¹³, sformułował koncepcję oparcia miary długości na zjawisku fizycznym, powtarzalnym i odtwarzalnym. Koncepcja ta miałaby mieć odniesienie do wahadła matematycznego odmierzającego sekundy, wynalezionej przez żyjącego jeszcze Galileusza. Pomysł Pułdowskiego, wyprzedzający podobne badania o przeszło 30 lat, zwłaszcza w odniesieniu do prac J. Pickarda (1671) Ch. Huygensa (1673), opublikował w 1675 r. w Wilnie fizyk włoskiego pochodzenia Tytus Liwiusz Burattini¹⁴ w dziele pt. *Miara powszechna*. W podtytule do tego dzieła zawarte zostały następujące słowa: „Traktat, w którym się wykazuje jako na każdym miejscu Ziemi daje się znaleźć jedna miara i jeden ciężar powszechny, niewymagające odnoszenia się do żadnej innej miary lub

¹³ Burattini napisał na str. 22 *Miary Powszechnej* [6] „Odwiedzałem często w Krakowie Księdza Stanisława Pułdowskiego, znakomitego matematyka i zastałem go porównywającego Miarę krakowską ze Stopą Rzymską Kapitolińską, która to miara została mu powierzona przez Magistrat owego miasta, przechowujący ją w Urzędzie miejskim wraz z Ciężarem.”

¹⁴ Urodzony około 1615 r. w Agordo (płn. Włochy), po licznych podróżach, szczególnie do Egiptu, gdzie razem ze znanym orientalistą Johnem Geraves badał wnętrza piramid, osiadł na stałe od 1641 r. w Rzeczypospolitej Obojga Narodów. W latach 60. XX wieku urządził w Ujazdowie pod Warszawą warsztaty mechaniki precyzyjnej i szlifiernię soczewek, w których zbudował liczne instrumenty pomiarowe. Prowadził także bogatą korespondencję naukową m. in. z takimi uczonymi, jak: Marino Mersenne, Ismael Boulliau, Atanazy Kircher, Jan Hevelius. Zmarł w 1682 r. [78].

ciężaru, które wszędzie będą te same, nie podlegające zmianom i wiekuiście tak długo świat istnieć będzie”.

Burrattini we wstępie pisze, m.in.: „Zbudowawszy wspomnianą wagę, napisałem następnie komentarz do traktatu p. Galileusza i pokazałem rzeczonemu Księdzu Pułdowskiemu, który go w swej uprzejmości bardzo chwalił, a kiedy doszedł do czytania, jak za pomocą wody znajduje się stosunek pomiędzy Kulą a Sześcianem, zastanowił się chwilę, i rzekł mi te wyraźne słowa: «Doszedłeś Pan bardzo blisko do wynalezienia rzeczy tak poszukiwanej przez cały świat, mianowicie do Powszechnej Miary i Ciężaru» (...).” Nie zdążył on już dokonać zaplanowanych na 1672 r. pomiarów „odległości dwóch lub trzech stopni, zdjętych na powierzchni Ziemi i mierzoną Metrem Powszechnym od południa ku północy na wielkich i płaskich równinach Polski”¹⁵ [77, 78].

Z kolei w 1670 roku Gabriel Mouton, ksiądz z Lugdunu (Francja), zaproponował system miar oparty na układzie dziesiętnym oraz jednostkę miary, stanowiącej długość łuku odpowiadającej 1' południka ziemskiego, a którego $1 \cdot 10^{-6}$ (milionowa) część tego łuku miała odpowiadać francuskiej jednostce długości: *toise*.

W innych propozycjach (Picard – 1671, Huyghens – 1673) wskazywano na odniesienie jednostki miary długości do długości wahadła sekundowego dla czasu średniego [79].

W okresie Wielkiej Rewolucji Francuskiej (1789–1799), Zgromadzenie Narodowe przyjęło w marcu 1790 r. propozycję, wniesioną przez Charles-Maurice de Talleyrand-Périgorda, wyznaczenia za wzorzec miary długość wahadła sekundowego, zlokalizowanego na 45° szerokości geograficznej (m.in. okolice Bordeaux, Grenoble, Turynu itd.). Na posiedzeniu w dniu 8 maja 1790 r. Zgromadzenie Narodowe rozpatrzyło projekt Talleyranda i na mocy dekretu zobowiązało Paryską Akademię Nauk do dokonania wyboru i przygotowania nowego systemu miar. Akademia wybrała dwie komisje, z których pierwsza, w składzie: Borda, Lagrange, Lavoisier, Tillet i Condorcet, zajęła się projektem podziału dziesiętnego dla wszystkich głównych jednostek miar, druga zaś w składzie: Borda, Lagrange, Laplace, Monge i Condorcet zajęła się wyborem nowej jednostki miar. Ta ostatnia odrzuciła ostatecznie propozycję Talleyranda oparcia jednostki na wahadle sekundowym i zdecydowała (27 października 1790 r.) o przyjęciu odniesienia nowej jednostki miary długości do 1/4 południka paryskiego, a właściwie do jej dziesięciomilionowej części oraz jej nazwę *mètre* („miarą”) (26 marca 1791 r.). Niezbędne w tym celu pomiary wykonano metodą triangulacyjną¹⁶ w 1792 r. na części południka przechodzącego przez Francję pomiędzy Dunkierką a Barceloną, przy czym Jan Baptysta Delambre (1749–1822) dokonał po-

¹⁵ Pomiary takie dokonali, nieco ponad 100 lat później, wzmiankowani uczeni francuscy, a nazwa metr podana przez Burrattyniego została oficjalnie nadana podstawowej jednostce systemu metrycznego [5].

¹⁶ Początkową niedokładność pomiarów kątów w metodzie triangulacyjnej, która wynosiła ok. 1/4 minuty, po skonstruowaniu przez Bordę nowego przyrządu – zmniejszono do ok. 1 sekundy [5].

miaru pomiędzy Dunkierką a Rodez, a Piotr Franciszek Méchain – między Rodez a Dunkierką. 1 sierpnia 1793 r. Konwent Narodowy wydał dekret wprowadzający nowy, metryczny system miar. W wyniku dalszych prac, prowadzonych pod auspicjami nowego Instytutu Narodowego, na początku 1799 r. uczeni francuscy wraz z zaproszonymi uczonymi zagranicznymi utworzyli trzy komisje. Dwie pierwsze, po ostatecznym zweryfikowaniu wykonanych prac geodezyjnych i astronomicznych, orzekły, że długość ćwierci południka ziemskiego wynosi 5 150 740 *toises*, co oznacza, że długość metra (zdefiniowana uprzednio jako jedna dziesięciomilionowa część 1/4 południka) wynosi 3 stopy 11,296 linii, czyli jest o 0,146 linii krótsza od metra prowizorycznego, przyjętego w 1795 r. Komisja trzecia ustaliła jednostkę ciężaru (po uprzednim zmodyfikowaniu jej nazwy – gram na kilogram) jako ciężar 1 dm³ wody destylowanej, w Paryżu, w próżni i przy temperaturze 4 °C. Wykonano wzorce nowych jednostek miar w postaci dwóch platynowych wzorów końcowych metra i jednego platynowego walca – wzorca kilograma.

Instytut Narodowy w dniu 22 czerwca 1799 r. przedstawił je Radzie Starszych i Radzie Pięciuset¹⁷, po czym wzorce te złożono w Archiwum Republiki i umieszczono „w podwójnej szafie żelaznej, zamkniętej na 4 zamki”. Dekretem z 9 grudnia 1799 r. uznano podane wzorce „za ostateczne”.

W rzeczywistości system metryczny stał się jedynie obowiązującym systemem miar we Francji dopiero po przyjęciu dekretu z 4 lipca 1837 r. Natomiast już w 1820 r. system metryczny stał się obowiązujący w Belgii i Holandii [79].

Swoistą emanacją rozwiązań francuskich na ziemiach Królestwa Kongresowego stały się tzw. miary nowopolskie. Na bazie prac porównawczych dawnych miar polskich z nowymi francuskimi, realizowanymi przez profesora matematyki R. Hube i profesora fizyki K. Markiewicza z Akademii Krakowskiej we współpracy z J.B. Delambre (1749–1822)¹⁸, publikacji prac Jędrzeja Śniadeckiego, Aleksandra Sapięhy (1802) i Aleksandra Chodkiewicza (1811) – 2 czerwca 1816 r., z inicjatywy Stanisława Staszica, Ministra Stanu, Towarzystwo Królewskie Przyjaciół Nauki powołało – w celu uporządkowania systemu miar w ówczesnym kraju – Komisję, złożoną z profesorów Uniwersytetu Królewsko-Warszawskiego. Komisja ta¹⁹ po zapoznaniu się z jednostkami miar wielu miast polskich przyjęła miary warszawskie, jako najbliższe „miarom francuskim”. Nowy układ miar polskich został oparty „na samym pier-

¹⁷ Odpowiednio: izba wyższa i niższa w ustanowionym w 1794 r. organie władzy ustawodawczej, jakim było Ciała Prawodawcze [przyp. aut.].

¹⁸ Prace te miały na celu przygotowanie reformy metrologicznej w Wolnym Mieście Krakowie [73].

¹⁹ „Na zachęcenie i pod przewodnictwem Rady Stanu Staszycza, Towarzystwo Przyjaciół Nauk w roku 1816 przyjęło na siebie ułożenie nowych miar i wag dla Polski. Wyznaczona była w tym celu z członków Towarzystwa deputacja, która, pod czynną prezydencją kasztelana Hrabi Alexandra Chodkiewicza, przygotowała stosowny do tego projekt, a ten w r. 1818 Czerwca 13, od najwyższej władzy przyjętym został i od 1 stycznia 1819 roku, jako normalny, dla Królestwa polskiego jest przeznaczony” [80].

wiastku, czyli etalonie, który, wzięty z południka, jest zasadą nowych miar francuskich” oraz zachował dawną polską nomenklaturę metrologiczną, tj. łokcie, sąźnie, ćwierci, cale, linie, kamienie, funty, łuty, uncje, korce, miarki, garnce, kwarty, kwaterki itd.²⁰; ze względu na powszechny nawyk dzielenia „miar i wag” na połowy i ćwierci – wprowadzenie zasady podziału dziesiętnego odłożono na przyszłość. Formalnie nowy system miar, zwany później nowopolskim, został wprowadzony rozporządzeniem ówczesnego rządu w dniu 13 czerwca 1818 r. [73], [78]. Po około 30 latach, tj. w 1848 r. władze carskie zastąpiły dotychczasowe miary nowopolskie systemem miar rosyjskich. Na przykład „Zasadą miary długości, jako też tworzących się z niej miar powierzchni i objętości, jest sążeń, wyrównywający długości złożonego w Twierdzy St. Petersburgskiej sążenia platynowego. Zasadą miary do ciał sypkich jest czetwieryk, obejmujący w próżni przy temperaturze $13 \frac{1}{3}$ °Réaumura, wody destylowanej 64 funty, czyli 1601,22 cali sześciennych. (...) Jednostkę zasadniczą wagi stanowi funt zawierający 9216 doli, a którego podstawą jest: że cal sześcienny Rossyjski wody destylowanej w próżni przy temperaturze $13 \frac{1}{3}$ ° Réaumura, waży 368,361 doli, czyli że funt tejez wody obejmuje 25,019 cali sześciennych. Funt aptekarski równa się 8064 doli, czyli $\frac{7}{8}$ częściom powyższego funta” [21].

1 lipca 1863 roku brytyjska Izba Gmin, 110 głosami przeciwko 75, przyjęła ustawę o wprowadzeniu obowiązku stosowania systemu metrycznego, bez podania jednak daty wejścia w życie tego obowiązku. Przeciwnicy tego rozwiązania podnosili zarzut, że nie jest korzystny dla tradycji brytyjskiej przyjmowania wzorców francuskich.

Istotnym czynnikiem oddziałującym na globalne rozpowszechnianie się systemu metrycznego była II Powszechna Wystawa Światowa w Paryżu w 1867 r., której podstawową ideą było „lepsze zrozumienie pomiędzy narodami i niesienie pokoju poprzez wymianę handlową”. W rezultacie rozpowszechniania się wiedzy o systemie metrycznym, zwłaszcza w ujęciu praktycznym ukazanym na wzmiankowanej wcześniej Wystawie w Paryżu, wzrastało przekonanie środowisk naukowych o potrzebie zwoływania międzynarodowych konwencji w sprawie pomiarów. W 1867 r. na pierwszej Międzynarodowej Konferencji Geodezyjnej w Berlinie podkreślono znaczenie [ustanowienia] jednolitego systemu miar w Europie oraz zalecono opracowanie nowego europejskiego prototypu metra. Akademia Nauk w Petersburgu oraz angielski komitet wzorców (ang. *English Standards Commission*) wyraziły zainteresowanie udziałem w realizacji podanego zalecenia. 1 września 1869 r. cesarz Napoleon III zaakceptował propozycje zawarte w raporcie francuskiego Ministra Handlu i Rolnictwa o utworzeniu międzynarodowej komisji naukowej, która miałaby na celu rozpowszechnianie idei stosowania pomiarów metrycznych dla ułatwienia wymiany handlowej i porównywania wyników pomiarów pomiędzy państwami, a także

²⁰ Podstawą systemu była tzw. „linia polska” równoważna dokładnie 2 milimetrom; 1 kwarta = 1 litr, 1 granik = 8 miligramów. Wzorce jednostek miar nowopolskich oparte były na francuskich prototypach wzorców miar [77].

opracowałyby konstrukcję prototypu wzorca międzynarodowego. W dniu 16 listopada 1869 r. rząd francuski zaprosił inne państwa do delegowania swych przedstawicieli do prac tej komisji naukowej; pozytywne odpowiedzi nadeszły z następujących krajów: Austro-Węgry, kraje niemieckie²¹ (tj. Bawaria, Prusy i Związek Północnoniemiecki, Wirtembergia), Belgia, Chile, Kolumbia, Ekwador, Hiszpania, Włochy²², Państwo Rzymskie²³, Stany Zjednoczone, Wielka Brytania, Grecja, Nikaragua, Holandia, Peru, Portugalia, Rosja, Salwador, Szwecja i Norwegia, Szwajcaria, Turcja, Wenezuela.

Pierwsze zebranie Międzynarodowej Komisji Metra, złożone z przedstawicieli 24 państw, odbyło się w dniach 8–13 sierpnia 1870 r., przerwane jednak wybuchem wojny pomiędzy Francją a Prusami. Pracujący dalej uczeni francuscy, po zweryfikowaniu, że oba wzorce metra i kilograma, znajdują się w doskonałym stanie, zaproponowali na kolejnym posiedzeniu Komisji Międzynarodowej (2–14 kwietnia 1872 r.), że podstawą metra międzynarodowego będzie ów metr archiwalny wytworzony ponad 70 lat temu. Przyjęto jednak, że nowy wzorec międzynarodowy będzie wzorcem kreskowym (a nie końcowym, jak te z 1799 r.) wykonanym z pręta platynowo-irydowego (90% Pt+10% Ir) długości 102 cm i przekroju w kształcie litery X. Wszystkie państwa reprezentowane na tej konferencji miały otrzymać identyczne kopie takiego metra. Postanowiono także skopiować wzorec archiwalny kilograma, stosując analogiczny do wymienionego stopu platynowo-irydowego. W celu realizacji przyjętych postanowień w zakresie wykonania nowych wzorców metra i kilograma Komisja Międzynarodowa powołała stały komitet. Ostateczne ustalenia, podjęte w dniach 24 września do 12 października 1872 r., rekomendowały kontynuowanie podjętych prac z udziałem już około 30 krajów (w tym – 10 krajów kontynentu amerykańskiego).

Aby rozwiązać problematykę podziału kosztów nowych badań i pomiarów pomiędzy poszczególne państwa – zorganizowano Konferencję Dyplomatyczną metra, na której w dniu 20 maja 1875 r. 20 państw podpisało tzw. Konwencję metryczną [82].

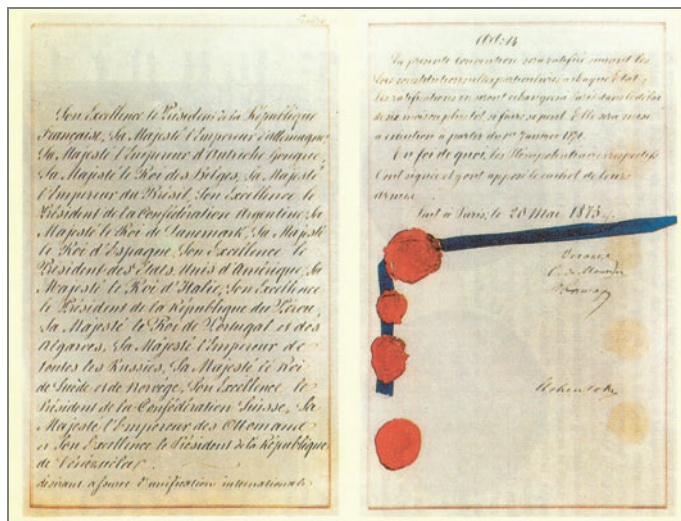
W preambule tej Konwencji wysokie układające się strony deklarują, że: „pragnąc zapewnić ujednostajnienie międzynarodowe i udoskonalenie systemu metrycznego, postanowili zawrzeć Konwencję w tym celu (...)”.

Kluczowe w tej Konwencji jest zobowiązanie zawarte w artykule 1: „Wysokie Układające się Strony zobowiązują się założyć i utrzymywać na koszt wspólny międzynarodowe Biuro Miar (ang. *Bureau International des Poids et Mesures* – BIPM), naukowe i stałe, z siedzibą w Paryżu.”

²¹ Zjednoczone w 1871 r. po zwycięstwie Prus (i innych zaangażowanych krajów niemieckich) nad Francją [przyp. aut.].

²² Królestwo Italii ze stolicą w Florencji [przyp. aut.].

²³ Państwo to przetrwało jedynie do 1870 r., kiedy to upadł (po klęsce wojennej z Prusami) protektorat francuski nad Rzymem [przyp. aut.].



Rys. 15. Pierwsza i ostatnia strona oryginalnego tekstu Konwencji Metrycznej²⁴
 Fig. 15. The first and last page of the original text of the Metre Convention

Zgodnie z art. 3 tej Konwencji – BIPM pozostawać miało pod wyłącznym kierownictwem i nadzorem Międzynarodowego Komitetu Miar (ang. *Comité international des Poids et Mesures – CIPM*), podlegającego Generalnej Konferencji Miar (ang. *Conférence générale des Poids et Mesures – CGPM*) zbierającej się co najmniej raz na 6 lat (zgodnie z art. 7 Regulaminu do Konwencji). „Miejsce przechowywania prototypów można będzie otwierać jedynie za pomocą trzech kluczy, z których jeden będzie w posiadaniu dyrektora Archiwów Francji, drugi – przewodniczącego Komitetu, a trzeci – dyrektora Biura”²⁵ [82].

Wkrótce po podpisaniu Konwencji, państwo francuskie oddało do dyspozycji Komitetu Międzynarodowego pawilon Breteuil w Parku Saint-Cloud. Komitet wybudował Instytut i zaopatrzył go w odpowiednie przyrządy [79].

8.4. MERYTORYCZNE I FORMALNE PODSTAWY POWSTAWANIA REGULACJI PRAWNYCH DOTYCZĄCYCH JEDNOSTEK MIAR

Najważniejszym, i jak dotąd najstarszym znanym, aktem prawa międzynarodowego publicznego, regulującym globalną współpracę w sferze metrologii jest wzmiankowana powyżej Konwencja Metryczna, podpisana w dniu 20 maja 1875 r. w Paryżu przez przedstawicieli 20 ówczesnych państw [82].

²⁴ Źródło: zbiory GUM.

²⁵ Art. 18 *in fine* Regulaminu Konwencji.



Rys. 16. Międzynarodowy Komitet Miar (CIPM) w 1894 r.

Źródło: Quinn T.J., *Metrology and the International System of Units: A Time for Change*, w wykładzie wygłoszonym w GUM w dniu 28 kwietnia 2009 r.

Fig. 16. International Committee for Weights and Measures (CIPM) in 1894

Source: Quinn T.J., *Metrology and the International System of Units: A Time for Change*, in the lecture delivered at GUM on April 28, 2009

Merytoryczne treści przyszłych regulacji prawnych, dotyczących naukowych aspektów poszczególnych jednostek miar oraz ich układu (obecnie – układu SI) są następstwem rezolucji Generalnej Konferencji Miar (CGPM), przygotowywane do dyskusji i aprobaty tegoż CGPM przez Międzynarodowy Komitet Miar (CIPM). CIPM kieruje także wszystkimi niezbędnymi pracami metrologicznymi, a w szczególności ustanawia współdziałanie specjalistów w zagadnieniach metrologii, skupionych w komitetach doradczych (Consultative Committee – CC)²⁶ oraz koordynuje wyniki ich prac. Podjęte rezolucje o charakterze metrologicznym wdrażane są do praktycznej realizacji, przede wszystkim przez Międzynarodowe Biuro Miar (BIPM) oraz naro-

²⁶ Obecnie CIPM obejmuje następujące CC: elektryczności i magnetyzmu (CEEM), fotometrii i radiometrii (CCPR), termometrii (CCT), długości (CCL), czasu i częstotliwości (CCTF), promieniowania jonizacyjnego (CCRI), jednostek miar (CCU), masy i wielkości związanych (CCQM) oraz akustyki, ultradźwięków i drgań (CCAUV) [przyp. aut.].

dowe instytucje metrologiczne (NMI) i jednostki desygnowane, a także inne organizacje i instytucje.

Już na I Generalnej Konferencji Miar, w dniu 26 września 1889 r., postanowiono zatwierdzić – jako wybrane przez CIPM wzorce międzynarodowe jednostek miary długości i masy – prototypy metra i kilograma. Notabene definicja metra oparta na wymienionym prototypie oraz definicji sekundy zostały zmienione dopiero po blisko 70 latach, odpowiednio – postanowieniami 6. i 9. Rezolucji XI CGPM (1960), natomiast definicja kilograma – w dalszym ciągu oparta jest na dotychczasowym jego prototypie.



Rys. 17. Źródło: Quinn T.J., *Metrology and the International System of Units: A Time for Change*, w wykładzie wygłoszonym w GUM w dniu 28 kwietnia 2009 r.

Fig. 17. Source: Quinn T.J., *Metrology and the International System of Units: A Time for Change*, in the lecture delivered at GUM on April 28, 2009

Ustanowiona odrębnie (poza Konwencją Metryczną) jednostka miary czasu – sekunda astronomiczna, zdefiniowana wówczas jako $1/86\,400$ część średniej doby słonecznej, tworzyła z ww. prototypami swoistą triadę – układ jednostek mechanicznych: **Metr – Kilogram – Sekunda**.

Wśród znaczących dokonań kolejnych CGPM należałoby wymienić:

- IX. Generalna Konferencja Miar w 1948 r. rozpisała ankietę dotyczącą wprowadzenia nowego układu miar po wielu krytycznych studiach na temat dotychczasowego (niedoprecyzowanego) układu (w szczególności miano wybrać czwartą jednostkę podstawową, dla pomiaru wielkości elektrycznych, spośród sześciu zaproponowanych). Wprowadziła (nadal obowiązującą) definicję ampera, jako jednej z proponowanych w ankiecie jednostek. Ustaliła (obowiązująca do 1979 r., z drobną poprawką) definicję kandeli.
- X. Generalna Konferencja Miar w 1954 r. ustanowiła sześć jednostek podstawowych (metr, kilogram, sekunda, amper, kandela, stopień Kelvina) i przyjęła

zasadę spójności układu jednostek podstawowych. Wybrała amper za elektryczną jednostkę podstawową. Ustaliła (nadal obowiązującą, z późniejszą poprawką) definicję stopnia Kelvina. Wprowadziła definicję jednostki ciśnienia – atmosfery fizycznej.

MÈTRES.	
<i>Prototype international M.</i>	
$M = 1^m + 8^{\mu} 651 T + 0^{\mu} 00100 T^2$	
<i>Étalon provisoire I₁.</i>	
$I_1 = 1^m + 6^{\mu} 6 + 8^{\mu} 644 T + 0^{\mu} 00100 T^2$	
<i>Prototypes de l'alliage de MM. Johnson, Matthey et C^{ie}.</i>	
N ^o 1	$= 1^m - 1,1 + 8,657 T + 0,00100 T^2$
N ^o 2	$= 1 - 1,5 + 8,665 T + 0,00100 T^2$
N ^o 3	$= 1 + 0,5 + 8,642 T + 0,00100 T^2$

K I L O G R A M M E S .							
Volume.		Masse.		Masse.			
ml		kg	mg	ml			
N ^o 1	46,426	1 +	0,002	N ^o 6	46,416	1 +	0,169
N ^o 2	412	—	953	N ^o 7	406	—	530
N ^o 3	422	+	021	N ^o 8	419	+	260
N ^o 4	418	—	075	N ^o 9	420	+	282
N ^o 5	46,413	1 +	0,018	N ^o 10	46,420	1 +	0,228

Rys. 18. Charakterystyki metrologiczne prototypów metra i kilograma, opublikowane w dokumentach I Generalnej Konferencji Miar w 1889 r.²⁷

Źródło: Quinn T.J., *Metrology and the International System of Units: A Time for Change*, w wykładzie wygłoszonym w GUM w dniu 28 kwietnia 2009 r.

Fig. 18. Metrological characteristics of the meter and kilogram prototypes, published in the documents of the General Conference on Weights and Measures in 1889.

Source: Quinn T.J., *Metrology and the International System of Units: A Time for Change*, in the lecture delivered at GUM on April 28, 2009

- XI Generalna Konferencja Miar w Paryżu w październiku 1960 r. ustanowiła ostatecznie międzynarodowy układ jednostek, wprowadzając dlań skrót SI (ang. *Système International d'Unités*). Zatwierdziła (obowiązującą do 1967 r.) definicję sekundy, używaną obecnie tylko do celów astronomicznych jako 1/31 556 925,9747 część roku zwrotnikowego 1900 stycznia 0 godzina 12:00, czasu efe-

²⁷ Źródło: <http://www.bipm.org/en/convention/cgpm/resolutions.html>

meryd. Przyjęła nową (obowiązującą do 1983 r.) definicję metra jako długości równej 1 659 763,73 długości fali promieniowania w próżni odpowiadającego przejściu między poziomami 2p a 5d atomu Kr (kryptonu 86). Oficjalnie wprowadzono przedrostki.

- XII Generalna Konferencja Miar w 1964 r. rozbudowała tabelę nazw krotności dziesiętnych.
- XIII Generalna Konferencja Miar w październiku 1967 r. wprowadziła nową nazwę jednostki temperatury – kelwin i drobną poprawkę w jej definicji. Przyjęła nową (obecnie obowiązującą) definicję sekundy, opartą na wzorcu odtwarzalnym w warunkach laboratoryjnych (zegar atomowy). Wprowadziła przeliczenie jednostki ciśnienia atmosfery fizycznej na jednostki SI. Wprowadziła drobną poprawkę w definicji kandeli.
- XIV Generalna Konferencja Miar w październiku 1971 r. wprowadziła do układu SI, jako jego siódmą jednostkę podstawową, jednostkę ilości (liczności) substancji mol i przyjęła jej (obecnie obowiązującą) definicję. Zatwierdziła nową nazwę jednostki ciśnienia – paskal.
- XV Generalna Konferencja Miar w maju/czerwcu 1975 r. rozbudowała tabelę nazw krotności dziesiętnych. Zdefiniowała grej i bekerel.
- XVI Generalna Konferencja Miar w październiku 1979 r. przyjęła nową (obecnie obowiązującą) definicję kandeli.
- XVII Generalna Konferencja Miar w październiku 1983 r. przyjęła nową (obecnie obowiązującą) definicję metra.
- XVIII Generalna Konferencja Miar w październiku 1987 r. rozpoczęła prace przygotowawcze do opracowania nowych definicji i wzorców ampera i kilograma.
- XIX Generalna Konferencja Miar we wrześniu/październiku 1991 r. rozbudowała tabelę nazw krotności dziesiętnych.
- XX Generalna Konferencja Miar w październiku 1995 r. zlikwidowała odrębną kategorię jednostek uzupełniających i włączyła je (tzn. radian i steradian) do kategorii jednostek pochodnych.
- XXI Generalna Konferencja Miar w październiku 1999 r. wprowadziła jako jednostkę pochodną katal.
- XXII Generalna Konferencja Miar w październiku 2003 r. dokonała rewizji *mise en pratique*²⁸ dla definicji metra.
- XXIII Generalna Konferencja Miar w październiku 2007 r. dokonała interpretacji definicji Kelwina (Rezolucja 10.), a w Rezolucji 12. omówiła możliwości redefinicji niektórych jednostek podstawowych układu SI (kilogram, amper, kelwin i mol) przez ich odniesienie do stałych fizycznych.

²⁸ *Mise en pratique* dotyczy praktycznych zasad postępowania i sposobu realizacji na najwyższym poziomie definicji danej jednostki miary.

- XXIV Generalna Konferencja Miar w październiku 2011 r. postanowiła w Rezolucji 1 o przygotowaniu w przyszłości redefinicji jednostek miar dla nowego układu SI. Nowy układ jednostek miar SI będzie oparty na 7 fundamentalnych stałych fizycznych:
 - częstotliwości stanu podstawowego struktury subtelnej atomu cezu ^{133}Cs – $(\Delta\nu^{133}\text{Cs})_{\text{hfs}}$, wynoszącej dokładnie 9 192 631 770 Hz,
 - prędkości światła w próżni c wynoszącej dokładnie 299 792 458 m/s,
 - stałej Plancka h , wynoszącej dokładnie $6.62606\text{X}\cdot 10^{-34}$ J·s,
 - ładunku elementarnego e , wynoszącego dokładnie $1,60217\text{X}\cdot 10^{-19}$ C,
 - stałej Boltzmanna k wynoszącej dokładnie $1,3806\text{X}\cdot 10^{-23}$,
 - stałej Avogadra N_{A} , wynoszącej dokładnie $6,02214\text{X}\cdot 10^{23}$,
 - skuteczności świetlnej K_{cd} dla promieniowania monochromatycznego o częstotliwości $540\cdot 10^{12}$ Hz,

gdzie symbol X określa dodatkowe cyfry wyznaczone i zatwierdzone przez Komitet CODATA przed przyjęciem nowych definicji.

Jakkolwiek treść rezolucji kolejnych Generalnych Konferencji Miar ma – bez żadnych wątpliwości – charakter źródłowy dla regulacji pochodnych, tym niemniej podkreślenia wymaga fakt, że rozwinięciem i uzupełnieniem tych rezolucji, a tym samym swoistym źródłem technicznego prawa metrologicznego, są oficjalne publikacje CIPM, CC i BIPM, a wśród nich, m.in.:

- Broszura SI, wyd. 8²⁹,
- Raport dot. działalności metrologicznej BIPM³⁰,
- Raport z 102 posiedzenia CIPM (VI i X 2013 r.)³¹,
- Raport z 21 posiedzenia CCU (11–12.06.2013 r.)³²,
- JCGM 106:2012, *Evaluation of measurement data – The role of measurement uncertainty in conformity assessment*³³,
- JCGM 200:2012, *International vocabulary of metrology – Basic and general concepts and associated terms (VIM)*³⁴,
- JCGM 104:2009, *Evaluation of measurement data — An introduction to the “Guide to the expression of uncertainty in measurement” and related documents*³⁵,

²⁹ http://www.bipm.org/utls/common/pdf/si_brochure_8.pdf

³⁰ <http://www.bipm.org/utls/common/pdf/DIR2013/2013supplementary.pdf>

³¹ <http://www.bipm.org/utls/en/pdf/CIPM2013-EN.pdf>

³² <http://www.bipm.org/utls/common/pdf/CCU21.pdf>

³³ <http://www.bipm.org/en/publications/guides/>

³⁴ Ibid.

³⁵ Ibid.

- JCGM 102:2011, *Evaluation of measurement data – Supplement 2 to the “Guide to the expression of uncertainty in measurement” – Extension to any number of output quantities*³⁶,
- JCGM 101: 2008, *Evaluation of measurement data — Supplement 1 to the “Guide to the expression of uncertainty in measurement” — Propagation of distributions using a Monte Carlo method*³⁷,
- JCGM 101:2008, *Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM)*³⁸.

Szczególną rolę w kształtowaniu treści regulacji prawnych dotyczących jednostek miar stanowi prawodawstwo Unii Europejskiej. Godzi się w tym miejscu przypomnieć, że jedynie rozporządzenia unijne wiążą w całości i są bezpośrednio stosowane we wszystkich Państwach Członkowskich. Natomiast dyrektywy unijne wiążą każde Państwo Członkowskie, do którego są kierowane w odniesieniu do rezultatu, który ma być osiągnięty, pozostawiając jednak organom krajowym swobodę wyboru formy i środków transpozycji (wdrożenia, implementacji). W odniesieniu do problematyki jednostek miar wydana została następująca dyrektywa:

- Dyrektywa Rady [Wspólnot Europejskich] z dnia 20 grudnia 1979 r. w sprawie zbliżenia ustawodawstw Państw Członkowskich odnoszących się do jednostek miar i uchylająca dyrektywę 71/354/EWG³⁹.

Pobieżna choćby analiza treści tej dyrektywy wskazuje na pochodny charakter warstwy merytorycznej tego aktu. Kolejne zmiany tej dyrektywy są w istocie „echem” rezolucji kolejnych Generalnych Konferencji Miar. Dominujące miejsce w tej dyrektywie zajmuje zagadnienie pojęcia i zakresu legalności stosowania poszczególnych jednostek miar, zwłaszcza gdy kwestia ta dotyczy obszaru Zjednoczonego Królestwa i Irlandii.

W preambule ww. dyrektywy⁴⁰ stwierdza się, m.in.:

- „(...) jednostki miar są niezbędne do używania wszystkich przyrządów pomiarowych do wyrażania wartości wielkości mierzonej lub wskazania wielkości; jednostki miar są używane w większości dziedzin działalności ludzkiej; konieczne jest zapewnienie możliwie najbardziej jasnych zasad ich stosowania; niezbędne więc jest uregulowanie ich stosowania wewnątrz Wspólnoty w dzie-

³⁶ Ibid.

³⁷ Ibid.

³⁸ Ibid.

³⁹ Dz.Urz. UE z 15.2.1980 Nr L 39/40; zmiana 1: Dz.Urz. UE z 3.1.1985 Nr L2/11, sprostowanie: Dz.Urz. UE z 23.8.2007 Nr L 218/18; zmiana 2: Dz.Urz. UE z 7.12.1989 Nr L 357/28, sprostowanie: Dz.Urz. UE z 23.8.2007 Nr L 218/16; zmiana 3: Dz.Urz. UE z 9.2.2000 Nr L 34/17, sprostowanie 1: Dz.Urz. UE z 29.4.2000 Nr L 104/89, tylko w wersji ang.; sprostowanie 2: Dz.Urz. UE z 12.12.2000 Nr 311/50, tylko w wersji ang.; zmiana 4: Dz.Urz. UE z 7.5.2009 Nr L 114/10; Vide: <http://eur-lex.europa.eu/oj/direct-access.html>

⁴⁰ W wersji pierwotnej z 1979 r. [przyp. aut.]

dzinie gospodarki, zdrowia publicznego, bezpieczeństwa publicznego, jak również dla celów administracyjnych;

- ustawodawstwo Państw Członkowskich regulujące używanie jednostek miar różni się w poszczególnych Państwach Członkowskich, co prowadzi do utrudnień w handlu; w tych warunkach niezbędne jest przeprowadzenie harmonizacji przepisów ustawowych, wykonawczych i administracyjnych w celu przezwyciężenia tych przeszkód;
- jednostki miar stanowią przedmiot międzynarodowych rezolucji przyjętych przez Generalną Konferencję Miar (CGPM) ustanowioną Konwencją Metryczną podpisaną w Paryżu dnia 20 maja 1875 r., do której należą wszystkie Państwa Członkowskie; wynikiem tych rezolucji było sporządzenie „Międzynarodowego Układu Jednostek Miar” (SI);
- Międzynarodowa Organizacja Normalizacyjna (ISO) przyjęła dnia 1 marca 1974 r. międzynarodową normę dotyczącą przedstawienia jednostek SI oraz innych jednostek przeznaczonych do używania w systemach zawierających ograniczone zbiory znaków; wskazane jest, aby Wspólnota przyjęła rozwiązania, które zostały już zaakceptowane na szerszym poziomie międzynarodowym w ramach normy ISO 2955 z dnia 1 marca 1974 r.;⁴¹
- przepisy prawa wspólnotowego dotyczące jednostek miar znajdują się w licznych tekstach wspólnotowych; zagadnienie jednostek miar jest tak ważne, że istotne jest, by można się było odnosić do jednego tekstu wspólnotowego; dlatego niniejsza dyrektywa łączy wszystkie przepisy prawa wspólnotowego w tym zakresie i uchyla dyrektywę 71/354/EWG (...).”

Natomiast w nowelizacji dokonanej w 2009 r., stwierdzono, m. in.:

- Dyrektywa Rady 80/181/EWG (3) wymaga od Zjednoczonego Królestwa i Irlandii ustalenia końcowej daty obowiązywania zwolnień, które nadal są stosowane w odniesieniu do jednostek miar znanych jako „pół kwarty” w przypadku mleka w butelkach zwrotnych oraz beczkowego piwa i cydru, „mili” w przypadku tablic sygnalizacji drogowej i oznaczeń prędkości oraz „uncji jubilerskiej” w przypadku handlu metalami szlachetnymi. Doświadczenie wykazało jednak, że ze względu na lokalny charakter tych zwolnień i ograniczoną liczbę produktów, których zwolnienia te dotyczą, utrzymanie ich nie spowodowałoby ograniczenia handlu przez utworzenie bariery pozataryfowej, w związku z czym nie ma już potrzeby zniesienia tych zwolnień.
- Wskazane jest, aby Komisja, w kontekście jej stosunków handlowych z państwami trzecimi, w tym z Transatlantycką Radą Gospodarczą, nadal zdecydowanie dążyła do zaakceptowania na rynkach państw trzecich produktów, na których etykietach widnieją wyłącznie jednostki Międzynarodowego Układu Jednostek Miar (SI).

⁴¹ Obecnie, tj. od 2007 r., jest to norma EN ISO 80000 (ICS 01.060) [przyp. aut.].

- W roku 1995 Generalna Konferencja Miar postanowiła znieść odrębną dotąd klasę jednostek uzupełniających SI i sklasyfikować jednostki „radian” i „steradian” jako bezwymiarowe pochodne jednostki SI, których nazwy i symbole mogą, lecz nie muszą, być stosowane w wyrażeniach dotyczących innych jednostek pochodnych SI.
- W roku 1999 Generalna Konferencja Miar przyjęła w ramach SI jednostkę „katal” o oznaczeniu „kat”, jako jednostkę SI określającą aktywność katalityczną. Ta nowa zharmonizowana jednostka SI miała zapewnić spójne i jednolite oznaczanie jednostek miar w dziedzinie medycyny i biochemii, a w konsekwencji wyeliminować wszelkie ryzyko nieporozumień wynikających ze stosowania niezharmonizowanych jednostek.
- W roku 2007 Generalna Konferencja Miar przyjęła uwagę do definicji „kelwina”, aby wyeliminować jedno z głównych źródeł obserwowanej zmienności realizacji punktu potrójnego wody. „Kelwin” zdefiniowany jest jako ułamekowa część temperatury termodynamicznej punktu potrójnego wody. Uwaga ta dotyczy wody o określonym składzie izotopowym.

Zauważyć można, że w podanej dyrektywie wskazano na źródłowy charakter normy europejskiej, jaką jest obecnie EN ISO 80000 (1–8):YYYY. **Wielkości fizyczne i jednostki miar** opracowano w ramach prac ISO/REMCO i ISO/TC 12.

Należy jednak ponownie zauważyć, że także i ten dokument ma charakter wtórny w odniesieniu do prac Konwencji Metrycznej. Świadczyć o tym może treść fragmentów wstępu do ww. normy:

„Nazwy jednostek miar odpowiednich wielkości fizycznych podano łącznie z oznaczeniami międzynarodowymi i definicjami. (...) Dalsze informacje – patrz Broszura BIPM dotycząca Układu SI. (...) Jednostki miar SI zostały przyjęte przez Generalną Konferencję Miar. (...) Następnie podano jednostki miar nie należące do układu SI, które zostały dopuszczone do stosowania łącznie z jednostkami miar układu SI przez Międzynarodowy Komitet Miar (...) lub Międzynarodową Organizację Metrologii Prawnej (...) albo przez ISO i IEC.”

Za swoistą ciekawostkę uznać można dokument innej organizacji międzynarodowej, jakim jest Zalecenie Nr 20, **Kody dla jednostek miar stosowanych w handlu międzynarodowym**, wydanym przez ONZ Komisję Ekonomiczną dla Europy⁴², a w którym dla ponad wielu setek rodzajów jednostek miar spoza układu SI, „nieobecnych” w ww. Broszurze BIPM, podane zostały współczynniki przeliczeniowe do tego układu, np. dla jednostki miary energii: *brake horse power (BHP)*, podano następującą relację: 1 BHP = 245,7 W.

W praktyce prawodawstwa polskiego recepcja systemu metrycznego realizowana była w różnorodny sposób. Na 24 artykuły zawarte w dekrete z dnia 8 lutego 1919 r. o miarach (Dz. Praw Nr 15, poz. 211), 8 z nich dotyczyło zagadnień legalnych jedno-

⁴² Źródło: <http://tfig.unece.org/contents/recommendation-20.htm>

stek miar, wskazując w niektórych z nich wprost na jednostki miar systemu metrycznego⁴³ (art. 2 – metr, art. 3 – metr kwadratowy, art. 4 – metr sześcienny, art. 6 – kilogram). Charakterystyczny w tej regulacji jest zakres ingerencji Państwa w przedmiocie jednostek miar – obowiązek stosowania legalnych jednostek miar dotyczył wyłącznie obrotu publicznego.

Analogiczne ramy prawodawcze, a w konsekwencji podobny zakres jurysdykcji, były stosowane w kolejnych rozwiązaniach prawnych:

- w dekreście z dnia 19 kwietnia 1951 r. o organach administracji miar oraz o miarach i narzędziach pomiarowych (Dz.U. Nr 26, poz. 195) – w art. 7 postanowiono, że: „1. We wszystkich aktach i czynnościach urzędowych oraz gospodarczych należy wyrażać miarę w legalnych jednostkach miar. 2. Rada Ministrów ustali w drodze rozporządzenia legalne jednostki miar.”;
- w ustawie z dnia 17 czerwca 1966 r. o miarach i narzędziach pomiarowych (Dz.U. Nr 23, poz. 148) – w art. 3 postanowiono, że: „1. W aktach i czynnościach urzędowych i zawodowych, czynnościach wchodzących w zakres gospodarki narodowej, mających znaczenie dla bezpieczeństwa i higieny pracy, ochrony zdrowia lub dla interesu społecznego, miarę wyraża się w legalnych jednostkach miar. 2. Rada Ministrów ustala w drodze rozporządzenia legalne jednostki miar oraz określa zasady i sposób wprowadzania ich w życie.”;
- w ustawie z dnia 3 kwietnia 1993 r. Prawo o miarach (Dz.U. Nr 55, poz. 248) – w art. 2 postanowiono, że: „We wszystkich dziedzinach życia publicznego do wyrażania wartości wielkości fizycznych należy stosować jednostki miar określone w ustawie.”, w art. 4 postanowiono, że: „1. Legalnymi jednostkami miar są: 1) jednostki Międzynarodowego Układu Jednostek Miar (SI) oraz ich dziesiętne podwielokrotności i wielokrotności, 2) jednostki nie należące do układu SI, lecz dopuszczone do stosowania, wraz z jednostkami SI, w drodze rozporządzenia Rady Ministrów. 2. Prezes Urzędu [GUM – przyp. aut.] ustala nazwy, definicje i oznaczenia legalnych jednostek miar.”, zaś w art. 5 postanowiono, że: „Jednostki miar inne niż określone w art. 4 mogą być stosowane: 1) na mocy porozumień międzynarodowych w dziedzinie transportu wodnego, lotniczego, kolejowego i drogowego, 2) w pracach badawczych, 3) we współpracy i wymianie międzynarodowej.”⁴⁴;
- w ustawie z dnia 11 maja 2001 r. Prawo o miarach (Dz.U. Nr 63, poz. 636) – w art. 5 postanowiono, że: „1. Legalnymi jednostkami miar są: 1) jednostki Międzynarodowego Układu Jednostek Miar (SI), 2) jednostki nienależące do Międzynarodowego Układu Jednostek Miar (SI), dopuszczone do stosowania na terytorium Rzeczypospolitej Polskiej, 3) dziesiętne podwielokrotności i wielokrotności jednostek, o których mowa w pkt 1 i 2. 2. Rada Ministrów określi,

⁴³ Polska przystąpiła do Konwencji Metrycznej dopiero 12 maja 1925 r. [przyp. aut.].

⁴⁴ Można w tym miejscu zauważyć swoistą transpozycję do prawa polskiego ww. dyrektywy Wspólnot Europejskich i to jeszcze 10 lat przed wstąpieniem Polski do UE. [przyp. aut.].

w drodze rozporządzenia: 1) legalne jednostki miar nienależące do Międzynarodowego Układu Jednostek Miar (SI), które mogą być stosowane na terytorium Rzeczypospolitej Polskiej, 2) nazwy, definicje i oznaczenia legalnych jednostek miar, 3) przedrostki i ich oznaczenia przeznaczone do tworzenia dziesiętnych podwielokrotności i wielokrotności legalnych jednostek miar, 4) zasady pisowni oznaczeń legalnych jednostek miar – przy uwzględnieniu zobowiązań wynikających z wiążących Rzeczpospolitą Polską umów międzynarodowych.”, w art. 6 zaś postanowiono, że: „1. Obowiązek stosowania legalnych jednostek miar dotyczy użytkowania przyrządów pomiarowych, wykonywania pomiarów i wyrażania wartości wielkości fizycznych w gospodarce, ochronie zdrowia i bezpieczeństwa publicznego oraz przy czynnościach o charakterze administracyjnym. 2. Jednostki miar inne niż określone w art. 5 mogą być stosowane w gospodarce w dziedzinach: transportu morskiego, lotniczego i kolejowego, na mocy porozumień międzynarodowych.”

Ostatnie aktualnie obowiązujące regulacje ustawowe dopełnione są przez następujące akty prawne:

- Rozporządzenie Rady Ministrów z dnia 30 listopada 2006 r. w sprawie legalnych jednostek miar (Dz.U. Nr 225, poz. 1638 oraz z 2010 r. Nr 9, poz. 61);
- Rozporządzenie Ministra Gospodarki, Pracy i Polityki Społecznej z dnia 30 stycznia 2003 r. w sprawie uznawania wzorców jednostek miar za państwowe wzorce jednostek miar (Dz.U. Nr 31, poz. 257 oraz z 2007 r. Nr 44, poz. 280);
- Obwieszczenie Prezesa Głównego Urzędu Miar z dnia 23 czerwca 2003 r. w sprawie wzorców jednostek miar spełniających warunki określone dla państwowych wzorców jednostek miar (M.P. Nr 35, poz. 492); także: M.P. z 2003 r. Nr 44, poz. 660; M.P. z 2004 r. Nr 1, poz. 20; M.P. z 2004 r. Nr 22, poz. 394.

Kształtowanie się systemów miar jest bez wątpienia istotnym elementem rozwoju cywilizacyjnego od początków dziejów. Stosując współczesne narzędzia opisu i interpretacji należałoby dorobek intelektualny i materialny ewolucji miar postrzegać – analogicznie do informatyki – jako naukę o metodach uzyskiwania i przetwarzania informacji. Po takim założeniu, wiarygodnym intuicyjnie i po części zweryfikowanym empirycznie, obiektywny stan rzeczywistości i coraz rozleglejsza oraz doskonalsza o niej wiedza skłaniają do postawienia następującej tezy:

nie jest rzeczą, ani słuszną ani konieczną, aby zjawiska i przejawy naszej obiektywnie funkcjonującej rzeczywistości, a zarazem wiedza o jej stanach ilościowych i jakościowych były przedmiotem jakichkolwiek regulacji prawa stanowionego.

Pobieżna choćby analiza historii metrologii dowodzi, że jakakolwiek czy kiedykolwiek ustanowiona norma prawna dotycząca miar⁴⁵ była lub jest efektem pochod-

⁴⁵ Zarówno w rozumieniu ogólnym, jak i w przypadkach rozwiązań konkretnych, szczegółowych, jak np. sporządzenie artefaktu kilograma i uznania go w 1889 r. za międzynarodową jednostkę miary masy [przyp. aut.].

nym względem określonego zjawiska przyrodniczego⁴⁶ albo wykorzystywała dogodną materializację istniejącego już zjawiska do desygnowania wprowadzanej do porządku prawnego określonej regulacji metrologicznej.

Podane rozważania oznaczają niewątpliwy prymat dorobku naukowego w sferze jednostek miar nad zastosowaniem jego w dziedzinie prawa; co dowodnie świadczy o tym, że treść technicznych regulacji metrologicznych oparta jest na efektach i dorobku pracy metrologów, relacja odwrotna nie zachodzi. Mimo to, nadal utrzymuje się wiara w omnipotentną funkcję prawa *sensu largo*, a co za tym idzie – narasta tendencja do obejmowania coraz to liczniejszych zagadnień życia codziennego ramami aktów normatywnych⁴⁷. Za zabiegi szczególnie absurdalne należy postrzegać działania takich „prawodawców”, którzy ustanowione już jednostki miar⁴⁸ czynią ponownym obiektem (elementem treści) innych przepisów prawnych, regulujących m. in. zasady ich odniesienia do postanowień już przyjętych⁴⁹. Poczucie absurdalności takich kreacji może być szczególnie silne, gdy uznamy, że w procesie stanowienia prawa dominuje – przede wszystkim – aspekt polityczny.

Jakie więc racjonalne argumenty (inne niż będące dorobkiem Nauki), procedowane w parlamencie, mogłyby zadecydować za uznaniem, że „ $2 \times 2 = 4$ ”, a może znalazłyby się jakieś inne powody (np. społeczno-ekonomiczne⁵⁰, ideologiczne, geopolityczne, itd.), aby uchwalić ustawę zawierającą zdanie „ $2 \times 2 = 4,123$ ”?

Jest oczywiste, że przedmiotem krajowych regulacji ustawowych w sferze systemu jednostek miar może być jedynie zagadnienie legalności określonego zestawu jednostek miar, a nie ich aspekt fizyczny lub matematyczny. Jeżeli więc wolą prawodawcy byłoby uznanie za legalną jednostkę miary objętości **kwarty**, równoważnej objętości 1 dm^3 , to, pomijając różnorodne skutki takiego aktu na przykład społeczno-gospodarcze, działanie takie można by uznać za uprawnioną decyzję suwerennego państwa.

Kończąc ten wywód, należałoby jeszcze przytoczyć (tytułem *de lege ferenda*) treść przesłanek normatywnych stanowienia określonych regulacji prawnych⁵¹. Zgodnie z po-

⁴⁶ W taki sposób ukształtuje się ostatecznie układ SI, oparty w całości na stałych fizycznych [przyp. aut.].

⁴⁷ Tzw. **jurydyzacja życia**; termin rozpowszechniony przez Janusza Kochanowskiego – Rzecznika Praw Obywatelskich, zm. trag. 10 kwietnia 2010 r. [przyp. aut.].

⁴⁸ W trybie rezolucji Generalnej Konferencji Miar lub stanowiące dorobek naukowy Konwencji Metrologicznej [przyp. aut.].

⁴⁹ Por. ww. Dyrektywa Rady [Wspólnot Europejskich] z dnia 20 grudnia 1979 r. w sprawie zbliżenia ustawodawstw Państw Członkowskich odnoszących się do jednostek miar.

⁵⁰ W okresie międzywojennym funkcjonowało w obiegu potocznym zdanie: „w rozumieniu ustawy rak jest rybą”, ponieważ nie wydano odrębnego aktu prawnego dotyczącego gospodarki i ochrony środowiskowej dla raków, uznając, że regulacja taka, dedykowana dla ryb, będzie miała zastosowanie także dla raków [przyp. aut.].

⁵¹ Por. rozporządzenie Prezesa Rady Ministrów z dnia 20 czerwca 2002 r. w sprawie Zasad techniki prawodawczej (M.P. Nr 100, poz. 908).

stanowieniami § 1 odnośnie do rozporządzenia Prezesa Rady Ministrów z dnia 20 czerwca 2002 r. – podjęcie decyzji o przygotowaniu projektu ustawy, należy poprzedzić:

- wyznaczeniem i opisaniem stanu stosunków społecznych w dziedzinie wymagającej interwencji organów władzy publicznej oraz wskazaniem pożądanych kierunków ich zmiany;
- ustaleniem potencjalnych, prawnych i innych niż prawne środki oddziaływania umożliwiające osiągnięcie zamierzonych celów;
- określeniem przewidywanych skutków społecznych, gospodarczych, organizacyjnych, prawnych i finansowych każdego z rozważanych rozwiązań;
- zasięgnięciem opinii podmiotów zainteresowanych rozstrzygnięciem sprawy;
- dokonaniem wyboru sposobu interwencji organów władzy publicznej.

Wydaje się zatem uzasadniona taka koncepcja, że w najbliższej przyszłości będą niezbędne nowe rozwiązania o charakterze normatywnym, harmonizujące relacje między dorobkiem naukowym instytucji Konwencji Metrycznej a zakresem regulacji prawa metrologicznego, ustanawianego w poszczególnych państwach.

Rewizji podlegać będzie zakres merytoryczny wymienionych wcześniej regulacji unijnej dotyczącej jednostek miar, ponieważ „prawem pierwotnym” w zakresie takich regulacji są postanowienia (rezolucje) Generalnych Konferencji Miar, a nie wspomniane przepisy unijne. Optymalnym rozwiązaniem może być na przykład akt unijny w randze rozporządzenia Rady, mającym walor bezpośredniego zastosowania w krajach członkowskich UE.

9. PODSUMOWANIE

Reasumując należy zatem zauważyć, że ustawiczne dążenia do osiągnięcia najniższych możliwych niepewności pomiaru oraz fakt zużywania się wzorców materialnych spowodowały opisane działania prowadzące do zmian definicji podstawowych jednostek Międzynarodowego Układu Jednostek Miar (SI). Jednak w wielu publikacjach [3, 33, 84, 85] pojawia się krytyka nowego podejścia definicyjnego jednostek. Zgłaszane są postulaty, aby przed wprowadzeniem zmian w dotychczasowych definicjach dokładnie przemyśleć wiele aspektów tego zagadnienia [3]. Stwierdzono między innymi, że zmiany takie spowodować mogą nieporządek, ponieważ w definicjach opartych na fundamentalnych stałych fizycznych nie ma odniesienia do ilości. Ponadto weryfikacja Międzynarodowego Układu Jednostek Miar (SI) może zaburzyć funkcjonowanie nauki, a nowe definicje podstawowych jednostek miar uczynią niemożliwym do wykrycia wszelkie przyszłe zmiany w stałych fundamentalnych. Dalejszymi skutkami mogą być ponadto straty ekonomiczne z powodu wzrostu kosztów transakcji oraz barier w handlu międzynarodowym. Pomimo pojawiania się takich krytycznych sygnałów, jest niemal pewne, że proces redefinicji jednostek miar będzie kontynuowany.

LITERATURA

- [1] OBALSKI J., *Zasady międzynarodowego układu jednostek miar SI*. Wydawnictwa Naukowo-Techniczne, Warszawa 1970.
- [2] Guide for the Use of the International System of Units (SI) (ang.), National Institute of Standards and Technology, dostęp: <http://physics.nist.gov/cuu/pdf/sp811.pdf>
- [3] LEONARD B.P., *Comments on recent proposals for redefining the mole and kilogram*, 2010, *Metrologia*, 47.
- [4] MILLS M. et al., *The new Si*, *Phil. Trans. R. Soc. A*, 2011, Vol. 369, 1953, 3910–3924.
- [5] *The International System of Units (SI)*, wyd. 8, Bureau International des Poids et Mesures (BIPM), Paris 2006.
- [6] BORDÉ Ch.J., *Base units of the SI, fundamental constants and modern quantum physics*, *Phil. Trans. R. Soc. A*, 2005, 363, 1834, 2177–2201.
- [7] KIND D., LÜBBIG H., *Metrology – the present meaning of a historical term*, *Metrologia*, 2003, 40, 255.
- [8] KIND D., QUINN T.J., *Metrology: Quo vadis?*, *Physics Today*, 1999, 52, 8.
- [9] Broszura BIPM, *The International System of Units (SI)*, Wyd. 8. Paris, Bureau International des Poids et Mesures, 2006.
- [10] KLUZA J., GRZANKA A., PLEBAN A., *Ku nowej definicji kilograma*, *Postępy fizyki*, Tom 58, z. 3, 2007, Praca zbiorowa, *The New SI based on fundamental Constants*, *Phil. Trans. R. Soc. A*, Londyn, Wielka Brytania, 2011.
- [11] DAVIS R., *The SI unit of mass*, *Metrologia*, 2003, 40, 299–305.
- [12] Raport Przewodniczącego Komitetu Doradczego ds. Masy przy Komitecie CIPM, wygłoszony w formie prezentacji w trakcie XXIV. Generalnej Konferencji Miar (CGPM), Paryż, 17–21.10.2011 r., dostępny na stronie internetowej BIPM.
- [13] ANDREAS B., AZUMA Y., BARTL G., BECKER P., BETTIN H. et al., *Counting the atoms in a ²⁸Si crystal for a new kilogram definition*, *Metrologia*, 2011, 48, S1. doi:10.1088/0026-1394/48/2/S01; Bettin H., *Determination of the Avogadro constant by counting ²⁸Si atoms*, Workshop, Turyn, wrzesień 2013.
- [14] KIBBLE B.P., *Atomic Masses and Fundamental Constants*, t. 5, red. J.H. Sanders, A.H. Wapstra, Plenum, New York 1975, 545.
- [15] PRAMANN A., RIENITZ O., SCHIEL D., SCHLOTE J., GUTTLER B., VALKIERS S., *Molar mass of silicon highly enriched in ²⁸Si determined by IDMS*, *Metrologia*, 2011, 48, S20–25.
- [16] YANG L., MESTER Z., STURGEON R.E., MEIJA J., *Determination of the atomic weight of the ²⁸Si-enriched silicon for revised estimate of the Avogadro constant*, *Anal. Chem.*, 2012, 84, 2321–2327.
- [17] NATRUKAWA T., HIOKI A., KURAMOTO N., FUJI K., *Molar-mass measurement of a ²⁸Si-enriched silicon crystal for determination of the Avogadro constant*, *Metrologia*, 2014, 51, 161–168.
- [18] BECKER P., GLÄSER M., *Int. J. Mass Spectrom.*, 251, 220, (2006).
- [19] STEINER R., *History and progress on accurate measurements of the Planck constant*, *Rep. Prog. Phys.*, 2013, 76, 016101, 1–46.
- [20] STOCK M., *Watt balance experiments for the determination of the Planck constant and the redefinition of the kilogram*, *Metrologia*, 50, 2013, R1-R16.
- [21] Raport GUM, sporządzony w oparciu o materiały z posiedzenia Komitetu Technicznego EURAMET ds. Masy (TC-M), Dubrownik, Chorwacja, marzec 2013.
- [22] Consultative Committee for Mass and Related Quantities (CCM) Working Group on the Realization of the Kilogram (WGR-kg), *Mise en pratique*, Vol. 8.7, 2013.
- [23] ABBOTT P.J., KUBARYCH Z.J., *The New Kilogram Definition and its Implications for High-Precision Mass Tolerance Classes*, *Journal of Research of the National Institute of Standards and Technology*, 2013, 118, 353–358.

- [24] PERRIN J., *Mouvement brownien et réalité moléculaire*, Annales de Chimie et de Physique, 1909, 8e (18), 1–114.
- [25] SI brochure, podrozdz. 2.1.1.6., *Unit of amount of substance (mole)*, zał. “Realising the mole”, wyd. 8 BIPM, 2006, http://www.bipm.org/utis/common/pdf/si_brochure_8_en.pdf, „SI unit of amount of substance (mole)” 1971-10, <http://www.bipm.org/en/CGPM/db/14/3/>
- [26] *Avogadra liczba*, [w:] *Encyklopedia techniki, CHEMIA*, WNT, Warszawa 1965, 58.
- [27] HUTTER H., *Dynamic SIMS*, [w:] *Ullmann's Encyclopedia of Chemical Technology*, Wiley, Weinheim 2005, 60, DOI: 10.1002/14356007.b06_023.
- [28] MARTIN J., MILTON T., *A new definition fo the mole based on the Avogadro constant: a journey from physics to chemistry*, Phil. Trans. R. Soc. A, 2011, 369, 1953, 3993–4003.
- [29] MILTON M.J.T., MILLS I.M., *Amount of substance and the proposed redefinition of the mole*, Metrologia, 2009, 46, 332–338.
- [30] MILTON M.J.T., QUINN T., *Primary methods for the measurement of amount of substance*, Metrologia, 2001, 38, 289–296.
- [31] WHEATLEY N., *On the dimesionality of the Avogadro constant and the definition of the mole*, Metrologia, 2011, 48, 71–82.
- [32] MILTON M.J.T., MILLS I.M., *Amount of substance and the proposed redefinition of the mole*, Metrologia, 2001, 38, 289–296.
- [33] PAVESE F., *Some reflections on the proposed redefinition of the unit for the amount of substance and of other SI units*, Accreditation and Quality Assurance: Journal for Quality, Comparability and Reliability in Chemical Measurement, 2011, 16 (3), 161–165.
- [34] RADOSZ A., *Cząstki i pola*, Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 1995.
- [35] *Termodynamika dla Studium Poddyplomowego Fizyki z Astronomią*, Kamińska M., Instytut Fizyki Doświadczalnej, Wydz. Fizyki, Uniwersytet Warszawski.
- [36] SI Brochure, wyd. 8 BIPM, rozdział. 2.1.1.5 – *Unit of thermodynamic temperature (kelwin)*, dostępne pod adresem internetowym: http://www.bipm.org/en/si/base_units/; „Mise en pratique for the definition of the Kelvin”, przyjęty w 2011 r. przez Komitet Doradczy CIPM ds. Temperatury (CCT).
- [37] PRESTON-THOMAS H., *The International Temperature Scale of 1990 (ITS-90)*, Metrologia, 1990, 27, 3–10 i 107;
- [38] *BIPM – The International Practical Temperature Scale of 1968 (IPTS-68)*, Metrologia, 1969, 15.
- [39] NICHOLAS J.V. et al., *Isotopic influence on the triple point of water and the definition of the k*, 6th International Symposium on Temperature and Thermal Measurements in Industry and Science TEMPMEKO '96, 1996 – Proceedings; Kowal A., Szmyrka A., Peruzzi A., PAK, 2013, Vol. 59, No. 5, 437–440.
- [40] Rekomendacja 5 (CI-1989) przyjęta przez Komitet Doradczy CIPM ds. Temperatury (CCT).
- [41] NEYEZHMAKOV P., NAZARENKO L., PAVLENKO Y., *About a new definition of Kelvin*, PAK, 2013, Vol. 59, No. 9, 954–957.
- [42] PELNUTH B., GAISER Ch., FISHER J., *Determination of the Boltzmann constant – status and prospects*, Meas. Sci. Technol., 2006, 17, R145–159.
- [43] Prezentacja Przewodniczącego Komitetu Doradczego CIPM ds. Temperatury (CCT), wygłoszona podczas 24. CGPM, Paryż, październik 2011 r.
- [44] MOLDOVER M.R., TRUSLER J.P.M., EDWARDS T.J., MEHLAND T.J., DAVIS R.S., *Measurement of the universal gas constant R using a spherical acoustic resonator*, J. Res. Natl. Bur. Stand., 1988, 93, 85–114.
- [45] de PODESTA M., UNDERWOOD R. et al., *A low-uncertainty measurement of the Boltzmann constant*, Metrologia, 2013, 50, 354–376, http://iopscience.iop.org/0026-1394/50/4/354/pdf/0026-1394_50_4_354.pdf
- [46] MERLONE A., MUSACHCIO C., SZMYRKA-GRZEBYK A., *New definition of the kelvin in terms of the Boltzmann's constant*, Elektronika: Konstrukcje, Technologie, Zastosowania, Wyd. SIGMA-NOT, 2011, Vol. 52, 6, 42–44.

- [47] FELLMUTH B., FISCHER J., GAISER Ch., JUSKO O., PRIRUENROM T., SABUGA W., ZANDT T., *Determination of the Boltzmann constant by dielectric-constant gas thermometry*, Metrologia, 2011, 48 (5), 382, DOI: 10.1088/0026-1394/48/5/020.
- [48] LIU H., FENG X.J. et al., *Improved determination of the Boltzmann constant using a single, fixed-length cylindrical cavity*, Metrologia, 2013, 50, 417–432.
- [49] GAISER Ch., ZANDT T. et al., *Improved determination of the Boltzmann constant by dielectric-constant gas thermometry*, Metrologia, 2013, 50, L7-L11.
- [50] http://www.bipm.org/en/si/si_brochure/chapter4/table7.html
- [51] GENEVES G., PIGUEMAL F., *Argument for direct realization of the quantum metrological triangle*, Metrologia, 2000, 37, 207–211.
- [52] KELLER M.W., *Current status of the quantum metrological triangle*, Metrologia, 2008, 45, 102–109.
- [53] http://www.bipm.org/en/si/si_brochure/chapter2/2-1/ampere.html
- [54] KAMPIK M., LAIZ H., KLONZ M., *Comparison of three accurate methods to measure AC voltage at low frequencies*, IEEE Trans. on Instr. And Meas., 2000, Vol. 49, 2, 429–433.
- [55] JOSEPHSON B.D., *Possible new effects in superconductive tunneling*, Physics Letters, 1962, 1, 251.
- [56] NARLIKAR A. et al., *Studies of Josephson Junction Arrays (II)*, Studies of High Temperature Semiconductors, 40, Nova Science Publishers, Inc., Nowy Jork 2002.
- [57] http://www.bipm.org/en/si/si_brochure/appendix2/electrical.html
- [58] ROSE-INNES A.C., RHODERICK E.H., *Nadprzewodnictwo*, Warszawa 1973.
- [59] HAMILTON C.A., BURROUGHS C., *Kao Chieh Operation of NIST Josephson Array Voltage Standards*, Journal of Research of the NIST, 1990, 4, 95.
- [60] SOCHOCKA D., STANIOCH W., *Odtwarzanie i przekazywanie jednostki napięcia elektrycznego w Głównym Urzędzie Miar*, XXXII Międzyuczelniana Konferencja Metrologów, Materiały Konferencyjne, Rzeszów 2000.
- [61] BENZ S.P., DRESSELHAU P.D., BURROUGHS Ch.J. Jr., *Nanotechnology for next generation Josephson voltage standards*, IEEE Trans. Instr. Meas., 2001, 6, 50.
- [62] BENZ S.P., HAMILTON C.A., *Application of the Josephson effect to voltage metrology*, Proc. IEEE, 2004, 10, 92.
- [63] LALLMANN K.P., *Josephson-voltage standard in a working calibration laboratory*, CSIR Test and Measurement Conference, Johannesburg 2002.
- [64] BURROUGHS Ch.P., BENZ S.P., HARVEY T.E., HAMILTON C.A., *1 Volt DC programmable Josephson voltage standard*, NIST, 1998.
- [65] HASSEL J., SEPPA H., GRONBERG L., SUNI I., *Optimization of a Josephson voltage array based on frequency dependently damped superconductor-insulator-superconductor junctions*, Rev. Sci. Instr., 2003, 7, 74.
- [66] Final report – key comparison – BIPM.EM-K11.b.
- [67] HETLAND P.O., LIND K., SLINDE H., *Characterisation of SINIS programmable Josephson arrays at Justervesenet*, CPEM 2002, Conference Digest, 2002, 506–507.
- [68] WYSOKIŃSKI K.I., *Quantum Hall Effect: the fundamentals*. Metrology and Measurement Systems, 2006, Vol. 13, No. 2, 113–124.
- [69] von KLITZING K., *25 Years of Quantum Hall Effect (QHE) – A Personal View on the Discovery*, Physics and Applications of this Quantum Effect, Seminaire Pincare, Paris, 2004, Vol. 2, 1–16.
- [70] *New Method for High-Accuracy Determination of the Fine-Structure Constant Based on Quantized Hall Resistance K_v von Klitzing*, G. Dorda, M. Pepper, Phys. Rev. Lett., 1980, 45, 494–497.
- [71] SIERAŃSKI K., KUBISA M., SZATKOWSKI J., MISIEWICZ J., *Półprzewodniki i struktury półprzewodnikowe*, Wydawnictwo Politechniki Wrocławskiej, Wrocław 2002.

- [72] GIRVIN S.M., *The Quantum Hall Effect: Novel Excitations and Broken Symmetries*, Springer Verlag, Heidelberg 1999.
- [73] KULA W., *Miary i ludzie*, wyd. 2, Książka i Wiedza, Warszawa 2004.
- [74] WIPSYZKA E. (red.), *Vademecum historyka starożytnej Grecji i Rzymu*, Źródłoznawstwo starożytności klasycznej, wyd. 2, Wydawnictwo Naukowe PWN, Warszawa 2001, 573–586.
- [75] JAWORSKI J.M., *Człowiek a pomiar czyli krótka historia mierzenia*, Oficyna Wydawnicza Politechniki Opolskiej, Opole 2002.
- [76] WRÓBLEWSKI A.K., *Elementy historii nauki i rozwoju fizyki – 2009*, źródło: <http://info.fuw.edu.pl/~akw/historia.html>
- [77] JANISZEK A., *Miary staropolskie, Miary nowopolskie i organizacja polskiej służby miar*, [w:] *Dzieje miar w Polsce*, Polski Komitet Normalizacji i Miar, Warszawa 1979.
- [78] BURATTINI T.L., *Miara powszechna*, Akademia Umiejętności w Krakowie, 1879, Źródło: http://rcin.org.pl/Content/34412/WA35_14702_7395_Miara.pdf
- [79] KOWALCZEWSKA Z., *Historia systemu metrycznego (1791–1921)*, Przegląd Techniczny, nr 13 i 14, 5 kwietnia 1921 r., 85–89.
- [80] MASSALSKI E.T., *Tablice porównawcze wszystkich wiadomych monet, wag i miar, wyrachowanych na monety, wagi i miary nowe francuskie, nowe polskie i rosyjskie*, Drukarnia Karola Kraya, Petersburg 1834, 13, źródło: <http://rcin.org.pl/dlibra/docmetadata?id=7895&from=pubindex&dirids=9&lp=4841>
- [81] Tablice zamiany miar i wag rosyjskich na polskie i nawzajem, w Komitecie Miar i Wag ułożone, a z mocy artykułu 7go Postanowienia Rady Administracyjnej Królestwa z dnia 2/14 marca 1848 roku, Komisję Rządową Spraw Wewnętrznych i Duchownych dla powszechnego użytku wydane, Warszawa 1849, 2, źródło http://rcin.org.pl/Content/24523/WA35_3595_7401tabl_Tablice-zamiany-miar.pdf
- [82] Konwencja dotycząca utworzenia i utrzymania Międzynarodowego Biura Miar, wraz z regulaminem, podpisana w Paryżu dnia 20 maja 1875 r. (Dz.U. R.P. z 29 sierpnia 1925 r., nr 87, poz. 608) – tzw. KONWENCJA METRYCZNA⁵².
- [83] Konwencja ustanawiająca Międzynarodową Organizację Metrologii Prawnej, podpisana w Paryżu dnia 12 października 1955 r. (Dz.U. z 1959 r, nr 4, poz. 24 oraz z 1969 r. nr 15, poz. 112) – tzw. KONWENCJA OIML.
- [84] PRICE G., *A skeptic's review of the New SI*, Accreditation and Quality Assurance: Journal for Quality, Comparability and Reliability in Chemical Measurement, 2011, 16 (3), 121–132.
- [85] MOHR P.J., *Defining units in the quantum based SI*, Metrologia, 2008, 45, 129–133.

THE QUANTUM SYSTEM OF MEASUREMENT UNITS (SI) – EVOLUTION OR REVOLUTION?

In conclusion, therefore it should be noted that long-term continuous striving for attainment of the lowest possible measurement uncertainty and the fact that primary material standards consumption, described above, have resulted in actions leading to changes of the basic units definitions of the International System of Units (*SI*). However, a critical voice of the new units definitional approach has been appeared in many publications [3, 33, 84, 85]. There are some postulates expressing that before putting the idea of the redefinition of units being currently in force into practice it is necessary to consider carefully many aspects of this issue [3]. It has been stated, that such changes may result in disorder, because in the unit definitions based on the fundamental physical constants there is no reference to quantity. Moreover,

⁵² Uwaga: teksty opublikowane w dziennikach ustaw (także, w tych najstarszych) można uzyskać z witryny RCL: <http://www.dziennikustaw.gov.pl/>

verification of the International System of Units (SI) can disturb science functioning and the new definitions of the basic measurement units will make it impossible to detect any future changes in the fundamental constants. Besides further consequences may cause economic losses due to the rise in transaction costs and barriers in the international trade. Despite the emergence of such critical signals, it is almost certain that the process of redefinition of units will continue.

Adam BIENKOWSKI*, Roman SZEWCZYK*

MAGNETOSPŘĘŻYSTE WŁAŚCIWOŚCI MATERIAŁÓW MAGNETYCZNIE MIĘKKICH JAKO SENSORÓW NAPRĘŻEŃ I SIŁ

Niniejsze opracowanie zawiera zagadnienie magnetośprężystych właściwości materiałów magnetycznie miękkich i możliwości wykorzystania ich w budowie sensorów naprężeń i sił. Opisano zjawisko magnetostrykcji i wyjaśniono rolę magnetostrykcji jako głównego czynnika mającego wpływ na magnetośprężyste właściwości materiałów. Omówiono magnetośprężyste zjawisko Villariego w magnetykach magnetycznie miękkich. Przedstawiono wyniki własnych badań doświadczalnych magnetostrykcyjnych i magnetośprężystych właściwości wybranych magnetyków ceramicznych – ferrytów oraz magnetyków amorficznych. Sformułowano uogólnione ujęcie w całym zakresie naprężeń (w obszarze sprężystym) od ściskających $-\sigma$ do rozciągających $+\sigma$ charakteru zależności $B(\pm\sigma)_H$ dotyczących zjawiska Villariego w odniesieniu do magnetyków magnetycznie miękkich w ogóle. W dalszej części omówiono model Jilesa–Athertona–Sablaka opisujący procesy magnesowania magnetyka, następnie dokonano rozwinięcia tego modelu w odniesieniu do procesu magnesowania magnetyka w obecności naprężeń od sił zewnętrznych. Ostatnia część opracowania poświęcona jest nowym możliwościom wykorzystania magnetośprężystego efektu Villariego w budowie sensorów naprężeń i sił ściskających.

1. WSTĘP

1.1. WPROWADZENIE

Zjawisko magnetostrykcji odkryte przez Joule'a [1] polega na zmianie wymiarów magnetyka pod wpływem przyłożonego pola magnesującego. Zjawiskiem odwrotnym jest zjawisko magnetośprężyste, odkryte przez Villariego [2, 3], polegające na zmianie stanu magnetycznego materiału magnetycznie miękkiego pod wpływem naprężeń wywołanych działaniem sił zewnętrznych w granicach sprężystości. Efekt ten może

* Instytut Metrologii i Inżynierii Biomedycznej, Politechnika Warszawska, ul. św. A. Boboli 8, 02-525 Warszawa.

być opisany m.in. jako zmiana indukcji magnetycznej B przy danym natężeniu pola magnetycznego H w wyniku zmiany naprężeń σ – charakterystyki $B(\pm\sigma)_H$. Znane są również zjawiska pochodne od zjawiska magnetostrykcji, jak np. efekty Guillemina [4] oraz Wiedemanna [5, 6].

Badania magnetosprężystego zjawiska Villariego w magnetykach mają dwa aspekty: fizyczny i techniczny. Fizyczny (podstawowy) – ze względu na poznanie natury fizycznej procesów magnesowania magnetyków w warunkach działania naprężeń wywołanych działaniem sił zewnętrznych; techniczny (użyteczny) – ze względu na możliwości zastosowania magnetyków do budowy sensorów naprężeń i sił.

Zależność indukcji magnetycznej B i względnych zmian wymiarów $\lambda = \Delta l/l$ magnetyka odpowiednio od naprężeń σ i natężenia pola magnesującego H są podstawowymi jego właściwościami magnetosprężystymi. Właściwości te mogą być wykorzystywane w praktyce m.in. do budowy sensorów magnetosprężystych, ale muszą też być często brane pod uwagę jako zjawiska niepożądane podczas projektowania niektórych podzespołów indukcyjnych. W drugim przypadku chodzi głównie o podzespoły magnetyczne mające rdzenie z małą szczeliną, np. głowice magnetyczne i/lub podzespoły narażone na działanie obciążeń mechanicznych.

Zależność indukcji magnetycznej B magnetyka od naprężeń σ wywołanych działaniem sił zewnętrznych stwarza duże możliwości praktycznych zastosowań, wykorzystujących tę właściwość m. in. w budowie sensorów naprężeń i sił. Możliwości te nie są dotąd w pełni wykorzystywane ze względu na brak jednoznacznej oceny zjawisk i magnetosprężystych cech magnetyków. Hamuje to w dużym stopniu zastosowanie i rozwój magnetyków ceramicznych – ferrytów oraz magnetyków amorficznych jako sensorów naprężeń i sił, mimo ich wielu zalet. Przykładowo w odniesieniu do ferrytów duża, rzędu 10^4 , wartość współczynnika czułości odkształceniowej $k_\mu = (1/\mu)(d\mu/d\varepsilon)$, (ok. 100 razy większa od analogicznej wartości współczynnika $k_r = (1/R)(dR/d\varepsilon)$ dla tensometrów półprzewodnikowych), bardzo duża rezystywność (10^0 – $10^4 \Omega\text{m}$), a więc małe straty związane z przewodnictwem elektrycznym, duża (rzędu 800 MPa) wytrzymałość mechaniczna na ściskanie, obojętność chemiczna i odporność na korozję, relatywnie niska w stosunku do magnetyków metalowych cena oraz możliwość wykonania rdzeni o różnych kształtach.

Prezentowany w niniejszym opracowaniu materiał zawiera rezultaty badań magnetosprężystych właściwości magnetyków ceramicznych – ferrytów oraz magnetyków amorficznych, ze względu na możliwości zastosowania tych materiałów jako magnetosprężystych rdzeni sensorów naprężeń i sił.

1.2. STAN WYJŚCIOWY ZAGADNIENIA

Ze względu na swą naturę fizyczną, magnetostrykcyjne zjawisko Joule'a oraz magnetosprężyste Villariego rozważane są zazwyczaj przez fizyków spójnie. Podejście takie prezentują w swoich pracach m.in. Bozorth [7], Lee [8], Cullity [9], Chikazumi [10].

Zarówno ze względu na zależność badanych właściwości magnetostrykcyjnych lub magnetosprężystych od fizykochemicznych cech materiałów, jak również ze względu na stosowane metody badań zjawiska te stanowią odrębne zagadnienia. Różne bowiem cechy fizykochemiczne materiału muszą być uwzględniane w badaniach magnetostrykcyjnych oraz magnetosprężystych. Różne są też problemy pomiarowe i kryteria oceny badań. Z tych względów zjawiska te mogą być rozpatrywane jako zagadnienia odrębne.

Naturę fizyczną zjawiska magnetostrykcji opisuje fenomenologiczna teoria makroskopowa opracowana przez Akułowa [11, 12] i Beckera [13], a rozwinięta następnie przez Browna [14]. Stan badań w zakresie zjawiska magnetostrykcji podsumowali w swoich monografiach dotyczących magnetyzmu i magnetyków Bozorth [7], Bielów [15, 16], Wonsowski [17, 18], Becker i Döring [19], Chen [20], Chikazumi [10], Cullity [9] i in. [21, 22]. Z polskich autorów znane są z tej dziedziny prace Smolińskiego [23], Kaczkowskiego np. [24–26], Wadasa [27], Szymczaka i in. [28, 29] oraz Lachowicza i in. [30].

Modele procesów magnesowania oraz magnetostrykcji i efektu magnetosprężystego są aktualnie bardzo intensywnie rozwijane ze względu na zastosowania praktyczne tych zjawisk. Na szczególne podkreślenie w tym zakresie zasługują prace Jilesa [31], Ramesha [32], Sablika [33], Chwastka [34], Augustyniaka [35] oraz Andrei [36].

W zakresie zjawiska magnetosprężystego badania wpływu naprężeń na właściwości magnetyczne materiałów metalowych omówiono w pracach [9, 37–48, 55–57], natomiast w ujęciu magnetyków amorficznych referowano w pracach [49–54, 58–62]. Badania doświadczalne wpływu naprężeń na strukturę domenową, umożliwiające zgłębienie natury fizycznej zjawiska magnetosprężystego, omówiono w pracy [63].

Charakteryzując stan wiedzy w zakresie magnetosprężystego zjawiska Villariego w magnetykach ceramicznych-ferrytach oraz magnetykach amorficznych, można stwierdzić, że wiele problemów teoretycznych i doświadczalnych nie zostało jeszcze dostatecznie zbadanych i wyjaśnionych. Można zauważyć brak opracowań dotyczących istoty występowania punktu Villariego. Nie są jeszcze dostatecznie znane charakterystyki $B(H)_\sigma$ tych magnetyków opisujące zmiany indukcji magnetycznej B w funkcji pola magnesującego H dla ustalonych naprężeń σ (obejmujące krzywe pierwotne i pętle histerezy magnetycznej) w zakresie występowania punktu Villariego. Odczuwa się potrzebę prac dotyczących zależności charakterystyk magnetosprężystych $B(\pm\sigma)_H$ (opisujących zależność indukcji magnetycznej od naprężeń dla ustalonego natężenia pola magnesującego) od magnetostrykcyjnych, anizotropowych i makrostrukturalnych właściwości tych magnetyków. Problem stanowi nie tylko istota zjawiska Villariego, ale i jego zależności od fizykochemicznych cech magnetyka. Wybór materiału odpowiedniego do danego zastosowania lub dostosowanie budowy sensora do magnetosprężystych i fizycznych cech magnetyka wymaga znajomości i właściwej oceny tych cech i ich wpływu na charakterystyki przetwarzania sensora.

Odczuwa się potrzebę dalszych opracowań dotyczących roli magnetostrykcji w zjawisku magnetosprężystym, a w tym z zakresu powiązania właściwości magnetostrykcyjnych $\lambda(\pm\sigma, H)$ (obejmujących zależność magnetostrykcji od pola i naprężeń),

z właściwościami magnetostryżystymi $B(\pm\sigma, H)$ (obejmującymi zależność indukcji od pola i naprężeń), szczególnie w zakresie naprężeń, w którym występuje punkt Villariego. I tak magnetyki charakteryzujące się zbliżoną do zera magnetostrykcją nasycenia λ_s , (np. ferryty Mn-Zn) uważane są jeszcze powszechnie jako magnetyki niewrażliwe na działanie sił zewnętrznych, pomimo że efekt magnetostryżysty dla tych materiałów, nie występujący dla $|\pm\sigma| \approx 0$, może być bardzo wyraźny, a czułość magnetostryżysta znaczna dla naprężeń (zarówno ściskających $-\sigma$, jak i rozciągających $+\sigma$) znacznie większych od zera.

W niniejszym opracowaniu omawiane jest wykorzystanie magnetostryżystych właściwości materiałów magnetycznie miękkich w budowie sensorów naprężeń i sił.

W rozdziale 2 przedstawiono fizyczne podstawy zjawiska magnetostrykcji, a w tym omówiono mechanizm atomowy magnetostrykcji. Wyjaśniono rolę magnetostrykcji jako głównego czynnika mającego wpływ na właściwości magnetostryżyste materiałów magnetycznie miękkich. Podano wyniki własnych badań doświadczalnych magnetostrykcyjnych właściwości wybranych magnetyków magnetycznie miękkich.

W rozdziale 3 przedstawiono fizyczne podstawy magnetostryżystego zjawiska Villariego w magnetykach magnetycznie miękkich. Sformułowano uogólnione ujęcie magnetostryżystego zjawiska Villariego dla magnetyków magnetycznie miękkich w odniesieniu do całego zakresu naprężeń, w obszarze sprężystym, od naprężeń ściskających $-\sigma$ do rozciągających $+\sigma$.

W rozdziale 4 omówiono, wykorzystując analizę energii swobodnej, model Jilesa–Athertona–Sabluka wraz z zaproponowanym rozszerzeniem uwzględniającym zmiany energii potrzebnej do pokonania zaczepu ściany domenowej w trakcie magnesowania. Model ten umożliwia nie tylko modelowanie pętli histerezy magnetycznej i magnetostrykcyjnej, lecz także ilościowy opis wpływu naprężeń na kształt pętli histerezy materiału magnetycznie miękkiego. Na szczególne podkreślenie zasługuje fakt, że parametry modelu Jilesa–Athertona–Sabluka są powiązane z właściwościami fizycznymi magnetyka.

W rozdziale 5 przedstawiono nowe możliwości wykorzystania efektu magnetostryżystego w budowie sensorów naprężeń i sił. Omówiono sposoby obciążania magnetostryżystego rdzenia sensora siłą zewnętrzną. Podano i omówiono wyniki badań doświadczalnych użytkowych właściwości badanych magnetyków jako sensorów naprężeń i sił.

W rozdziale 6 podsumowano niniejsze opracowanie.

2. WŁAŚCIWOŚCI MAGNETOSTRYKCYJNE MAGNETYKÓW

2.1. WPROWADZENIE

Zjawisko magnetostrykcji (magnetostrykcja wymuszona) polega na zmianie wymiarów magnetyka w procesie jego magnesowania. Rozróżnia się dwa rodzaje ma-

gnetostrykcji: liniową i objętościową. Magnetostrykcja liniowa jest zmianą wymiarów magnetyka (bez zmiany jego objętości) po umieszczeniu go w polu magnetycznym. Magnetostrykcja objętościowa (ω) w czystej postaci polega na jednoczesnej i jednakowej zmianie wszystkich wymiarów próbki, tzn. na skurczeniu bądź spęcznieniu. Zmienia się wówczas objętość z zachowaniem kształtu.

Jeżeli stan magnetyczny materiału zmienia się od tego, jaki występuje podczas rozmagnesowania elementu, do stanu nasycenia magnetycznego, to odpowiadające tym zmianom wydłużenie względne $\Delta l/l = \lambda$ nazywane jest współczynnikiem magnetostrykcji nasycenia λ_s .

Magnetostrykcja liniowa może być dodatnia ($\lambda > 0$), jeżeli element magnesowany wydłuża się lub ujemna ($\lambda < 0$) w przypadku, gdy się kurczy. Wartość λ_s np. większości ferrytów jest ujemna i zawiera się w granicach od wartości zbliżonej od zera do $-200 \mu\text{m/m}$ [64]; o kilka rzędów mniejsza zatem od odkształceń wywołanych działaniem sił zewnętrznych. Przykładowo przy module Younga ferrytu równym 150 GPa odkształcenie odpowiadające $\lambda_s = 10 \mu\text{m/m}$ uzyskuje się dla naprężeń mechanicznych równych zaledwie 1,5 MPa.

Interpretacja magnetosprężystych zależności $B(\pm\sigma, H)$ wymaga określenia roli magnetostrykcji w zjawisku Villariego. Ponieważ znajomość samej wartości λ_s jest w większości przypadków niewystarczająca do takiej interpretacji, zachodzi potrzeba określenia zmian magnetostrykcji od wartości naprężeń powodowanych działaniem sił zewnętrznych. Jest to zagadnienie określenia efektu magnetostrykcyjnego próbki, na którą działają jednocześnie pole magnesujące H oraz naprężenia $\pm\sigma$. Ma to znaczenie praktyczne w odniesieniu do magnetosprężystych sensorów naprężeń i sił. Do wyboru materiału magnetycznego na sensor określonego rodzaju lub dostosowania budowy sensora do fizycznych, a w tym magnetosprężystych, cech magnetyka konieczna jest znajomość magnetostrykcji materiału zastosowanego na rdzeń sensora.

Zjawisko magnetostrykcji, poza autorami wymienionymi we wstępie, omawiają opracowania syntetyczne Néela [65], Lee [8,66], Callena [67, 68], Cullity'ego [69], Rinaldiego i in.[70], Sablika i Jilesa [71, 72]. Ukazały się również prace teoretyczne i doświadczalne dotyczące magnetostrykcyjnych właściwości materiałów metalowych [73–81], w tym amorficznych [82–91] i stopów metali ziem rzadkich [92–95].

Publikowane dane w odniesieniu do magnetyków zarówno mono-, jak i polikryształów najczęściej dotyczą tylko współczynnika magnetostrykcji nasycenia λ_s i nie zawsze zawierają krzywe pierwotne i pętli histerezy magnetostrykcji.

2.2. FIZYCZNE PODSTAWY ZJAWISKA MAGNETOSTRYKCJI

2.2.1. NATURA FIZYCZNA ZJAWISK MAGNETYCZNYCH W MATERIALE

Magnetostrykcja spontaniczna pojawia się podczas schładzania materiału podczas przejścia przez temperaturę Curie T_C lub Néela (T_N) i jest wynikiem naruszenia wa-

runków równowagi sieci krystalicznej. Materiał paramagnetyczny w temperaturze $T \leq T_C$ lub $T \leq T_N$ staje się ferro-, ferri- lub antyferromagnetykiem.

Źródłem spontanicznego namagnesowania jest ruch elektronów dookoła jądra i własnej osi. Z pierwszymi związane są magnetyczne momenty orbitalne, z drugimi – spinowe (spiny); te ostatnie z reguły decydują o wypadkowym momencie powłok elektronowych i całego atomu. Spontaniczna magnetyzacja, czyli równoległe ustawienie spinów, wywołane jest przez kwantowo-mechaniczne oddziaływanie wymienne.

Uporządkowanie spinów sąsiednich atomów danego układu (sieci krystalicznej) jest wynikiem istnienia energii wymiany E_{ex} . Można ją wyrazić dla pary spinów S_i i S_j wzorem [9]

$$E_{ex} = 2J\bar{S}_i\bar{S}_j \quad (1)$$

w którym J jest tak zwaną całką wymiany, związaną z nakładaniem się ładunków i, j ; czyli energią umożliwiającą wymianę elektronów między atomami.

Po przyjęciu pewnych uproszczeń energię wymiany układu można wyrazić w formie matematycznej następująco [169]:

$$E_{ex} = -JNy^2z \quad (2)$$

przy czym J jest całką wymiany (jak wyżej), N jest liczbą atomów danego układu, y^2 jest prawdopodobieństwem przyjęcia tego samego kierunku przez spiny dwu sąsiadujących ze sobą atomów, z jest liczbą sąsiadujących bezpośrednio ze sobą atomów.

Ze względu na minimum energii układu powstaje struktura domenowa. Wektory magnetyzacji poszczególnych domen, czyli obszarów Weissa o równoległym uporządkowaniu wypadkowych magnetycznych momentów atomów (zgodnie z zasadą minimum energii swobodnej) przyjmują konfigurację umożliwiającą zamykanie się całego, lub prawie całego, strumienia magnetycznego wewnątrz materiału.

Struktura domenowa wynika z zasady minimum energii wewnętrznej (swobodnej) materiału, na którą składają się: energia anizotropii magnetokrystalicznej (wynikająca ze struktury krystalicznej materiału), energia magnetostrykcji związana z odkształceniami i naprężeniami powstającymi pod wpływem magnetyzacji materiału (magnetostrykcja spontaniczna i naprężenia własne), energia od magnesowania wynikająca z istnienia niemagnetycznych mikroprzestrzeni w materiale (pory, wtrącenia niemagnetyczne) oraz energia powstających domen.

O kierunku łatwej magnetyzacji i wielkości anizotropii magnetokrystalicznej kryształów decyduje ich skład chemiczny i rozkład atomów oraz jonów w sieci krystalicznej. Źródłem tej anizotropii są zależne od wymienionych czynników spinowo-orbitalne oddziaływania elektronów powłoki 3D pierwiastków przejściowych oraz wpływ na te oddziaływania pola krystalicznego. Oddziaływania typu spin-orbita są oddziaływaniami magnetycznego momentu spinowego z efektywnym polem magnetycznym wytworzonym przez elektron w jego ruchu orbitalnym dookoła jądra atomowego.

Uporządkowanie spinów, a z nim powstanie struktury domenowej, występuje, gdy istniejąca już, ze względu na oddziaływania kulombowskie, struktura krystaliczna jest ustalona („zamrożona”) i nie jest możliwe swobodne jej odkształcenie. Atomy w sieci krystalicznej w związku z anizotropią magnetokrystaliczną dążą do zajęcia innych położeń, w wyniku czego następują pewne odkształcenia sieci.

Magnetostrykcja spontaniczna jest zmianą geometryczną kryształu będącą następstwem zmiany warunków równowagi między węzłami sieci krystalicznej w wyniku pojawienia się magnetyzacji spontanicznej podczas schładzania materiału poniżej temperatury T_C lub T_N . Poniżej temperatury T_C stała sieci krystalicznej w żelazie zwiększa się w kierunku łatwej magnetyzacji, natomiast maleje w kierunku poprzecznym. W niklu poniżej temperatury T_C , a także w większości ferrytów o strukturze spinelowej poniżej temperatury T_N , zachodzi zjawisko odwrotne.

Energia wymiany jest izotropowa; zależy tylko od kąta między sąsiednimi spinami, a nie zależy od kierunku osi spinów w stosunku do osi kryształu. Tak więc sprzężenie typu spin–spin (oddziaływanie wymienne między dwoma sąsiednimi spinami wzajemnie równoległymi lub antyrównoległymi) nie ma udziału w anizotropii magnetokrystalicznej.

Przykładowo w odniesieniu do większości magnetyków ceramicznych – ferrytów o strukturze spinelowej orbitalny moment magnetyczny powłoki 3D kationów jest „zamrożony” przez pole krystaliczne. Oznacza to, że orientacje orbit są silnie unieruchomione przez sieć, z czego wynika, że sprzężenie typu orbita–sieć jest silne.

Jeżeli pole zewnętrzne zmienia orientację momentu spinowego (spinu) elektronu, to moment orbitalny elektronu ma również tendencję do zmiany swojej orientacji. Ponieważ moment orbitalny jest silnie sprzężony z siecią, to przeciwdziała próbom obrotu momentu spinowego. Energia liczona na jednostkę objętości lub masy, jaka jest wymagana do zmiany kierunku osi spinu z kierunku łatwego, jest energią anizotropii magnetokrystalicznej i jest właśnie energią wymaganą do pokonania sprzężenia typu spin–orbita. O wielkości anizotropii magnetokrystalicznej, a więc i magnetostrykcji spontanicznej, decyduje sprzężenie między momentem spinowym a orbitalnym każdego elektronu w polu krystalicznym.

Sprężenie to jest słabe, ponieważ pola już kilku tysięcy amperów na metr są praktycznie dostatecznie silne do obrotu spinów. W tej sytuacji o kierunku łatwej magnetyzacji decyduje orientacja spinów. Kierunkami tymi są zazwyczaj w większości ferrytów osie krystalograficzne [111]. Jedynie w przypadku jonów Co^{2+} i Fe^{2+} , ze względu na zwyrodnienie stanu podstawowego elektronów, moment orbitalny elektronów nie jest „zamrożony” przez pole krystaliczne w ogóle (Co^{2+}) lub przynajmniej w pewnych warunkach (Fe^{2+}) związanych z położeniem wspomnianych kationów w ferrospinelu. Tłumaczy się tym fakt, że ferryty kobaltu, a także niektóre ferryty zawierające jony Co^{2+} lub Fe^{2+} , charakteryzują się kierunkami łatwej magnetyzacji równoległymi do krawędzi kryształu. Ponieważ formalnie przyjęto, że dla kryształów

układu regularnego kierunkami odniesienia (będącymi zarazem kierunkami łatwej magnetyzacji) są osie [001], [010], [100], prawie wszystkie ferrospinele, poza wymienionymi ferrytami zawierającymi Co^{2+} lub Fe^{2+} , charakteryzują się tzw. ujemną anizotropią magnetokrystaliczną.

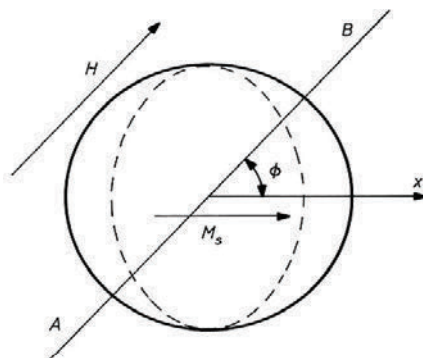
W metalach ziem rzadkich, z których większość stanowią ferromagnetyki w zakresie temperatur znacznie poniżej temperatury pokojowej, momenty orbitalne nie są zamrożone, co oznacza, że sprzężenie typu spin-orbita jest silne. Przyłożone pole, obracając momenty spinowe, powoduje również obroty momentów orbitalnych i związany z tymi obrotami znaczny efekt magnetostrykcyjny. Na przykład magnetostrykcja nasycenia λ_s dysprozu w temperaturze 22 K jest równa około $4,5 \cdot 10^{-3}$ [9, 96].

Skoro zarówno anizotropia magnetokrystaliczna, jak i magnetostrykcja spontaniczna mają to samo źródło, którym jest sprzężenie spin-orbita, należy oczekiwać pewnej korelacji pomiędzy tymi wielkościami. Istotnie, dużej bezwzględnej wartości stałej anizotropii K_1 odpowiada zwykle duża bezwzględna wartość magnetostrykcji nasycenia λ_s . Przykładowo ferryt Mn-Zn o małej anizotropii magnetokrystalicznej ma zbliżoną do zera magnetostrykcję λ_s . W stopach binarnych domieszka drugiego składnika w roztworze stałym często zmniejsza zarówno stałą anizotropii $|K_1|$, jak i $|\lambda_s|$. Są to tendencje główne, jakkolwiek występuje wiele wyjątków.

Zależności fizyczne pomiędzy anizotropią magnetokrystaliczną a magnetostrycją rozważał Kittel [97]. Rozszerzył on definicję energii anizotropii magnetokrystalicznej, określając ją nie tylko jako funkcję kierunku magnesowania odnoszonego do osi kryształu, ale i jako funkcję odległości międzyatomowych. Gdy anizotropia magnetokrystaliczna traktowana jest jako wielkość pierwotna, wówczas można uważać, że magnetostrykcja pojawia się wskutek zmniejszenia energii anizotropii magnetokrystalicznej dzięki spontanicznemu odkształceniu sieci, przy czym zmniejszenie to jest większe niż przyrost energii sprężystej. Ten punkt widzenia zaproponował właśnie w swojej pracy Kittel [97]. Tak więc np. sześcienna komórka elementarna żelaza, podczas schładzania po przejściu przez temperaturę Curie, odkształci się w tetragonalną, jeżeli energia anizotropii zmniejszy się przez to bardziej niż wzrosnie energia sprężysta.

W celu określenia zależności odkształcenia magnetostrykcyjnego od kierunku magnetyzacji spontanicznej rozważmy w kryształcie jednodomenową przestrzeń kulistą, która po schłodzeniu poniżej T_C odkształci się spontanicznie w elipsoidę. Średnica kuli w wyniku tego procesu odkształci się spontanicznie (w rozpatrywanym przypadku ulegnie skróceniu) wzdłuż kierunku magnetyzacji M_s o wartość $\Delta/l = e$ (rys. 1). Odkształcenie średnicy wzdłuż osi AB tworzącej kąt Φ z kierunkiem magnetyzacji M_s (oś x) określone jest zależnością

$$\frac{\Delta l}{l} = e \cos^2 \Phi \quad (3)$$

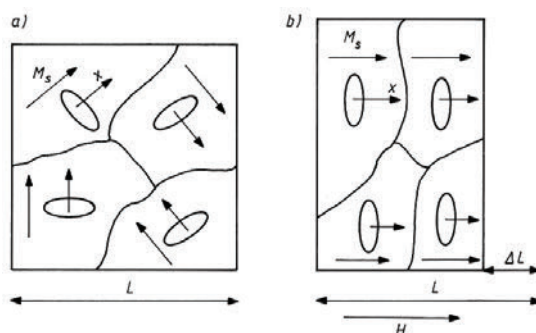


Rys. 1. Spontaniczne odkształcenie kuli wzdłuż kierunku AB tworzącego kąt Φ z kierunkiem x spontanicznej magnetyzacji M_s (schematycznie przypadek ujemnej magnetostrykcji) [98]

Fig. 1. Spontaneous of elongation in a direction which makes an angle Φ with the axis of spontaneous strain

Jeżeli w stanie rozmagnesowania rozkład wektorów magnetyzacji w próbce jest przypadkowy (rys. 2a), to wartość odkształcenia spontanicznego próbki otrzyma się przez uśrednienie zależności (3)

$$\left(\frac{\Delta l}{l}\right)_{\text{rozmag}} = \int_0^{\frac{\pi}{2}} e \cos^2 \Phi \sin \Phi d\Phi = \frac{e}{3} \quad (4)$$



Rys. 2. Rotacje spontanicznej magnetyzacji M_s i towarzyszące im rotacje osi spontanicznych odkształceń (schematycznie przypadek ujemnej magnetostrykcji):

a) – stan rozmagnesowania, b) – stan nasycenia [98]

Fig. 2. Rotation of domain magnetization and accompanying rotation of the axis of spontaneous strain

a) – demagnetization state, b) – saturation state

W procesie magnesowania jednocześnie z obrotami magnetyzacji występują obroty osi odkształceń spontanicznych, powodując w efekcie odkształcenie całej próbki.

Ponieważ w stanie nasycenia (rys. 2b)

$$\left(\frac{\Delta l}{l}\right)_{\text{nas}} = e \quad (5)$$

to magnetostrykcja nasycenia λ_s jest określona zależnością

$$\lambda_s = \frac{\Delta L}{L} = \left(\frac{\Delta l}{l}\right)_{\text{nas}} - \left(\frac{\Delta l}{l}\right)_{\text{rozmag}} = \frac{2}{3}e \quad (6)$$

Magnetostrykcja spontaniczna e powiązana jest z magnetostrycją nasycenia λ_s zależnością

$$e = \frac{3}{2}\lambda_s \quad (7)$$

Magnetostrycję spontaniczną trudno jest zaobserwować bezpośrednio, ale przejawia się ona ewidentnie przez istnienie lokalnego ekstremum, przy temperaturze T_c , temperaturowego współczynnika rozszerzalności liniowej. Efekt ten obserwowany jest na przykład w stopie Fe-Ni zawierającym 36% Ni w temperaturze bliskiej pokojowej. Punkt Curie tego stopu jest bliski temperatury pokojowej.

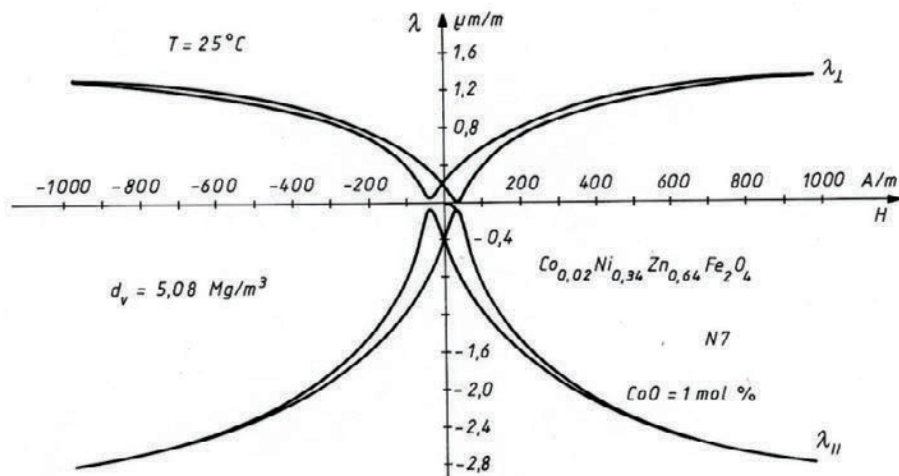
Przyłożenie pola magnesującego wzdłuż kierunku AB (rys. 1) spowoduje obrót wektora magnetyzacji spontanicznej, a wraz z nią obrót spontanicznej elipsoidy magnetostrykcji do tego kierunku. W wyniku tego obrotu odkształcenie względne $\Delta l/l$ wzdłuż kierunku AB zmieni się o wartość $\delta(\Delta l/l)$ równą

$$\delta\left(\frac{\Delta l}{l}\right) = e - e \cos^2 \Phi = \frac{3}{2}\lambda_s(1 - \cos^2 \Phi) \quad (8)$$

Z zależności (8) wynika, że pole magnesujące H przyłożone wzdłuż kierunku spontanicznej magnetyzacji ($\Phi = 0$ lub 180°) nie spowoduje zmiany odkształcenia próbki; $\delta(\Delta l/l) = 0$. Natomiast przyłożenie pola H prostopadle do kierunku M_s , ($\Phi = \pm 90^\circ$) spowoduje zmianę odkształcenia względnego o wartość $\delta(\Delta l/l) = 3/2 \lambda_s$. W procesie magnesowania, w którym zachodzą przesunięcia ścian 180 i 90° oraz obroty wektora magnetyzacji, tylko przesunięcia ścian 90° i obroty są źródłem odkształceń magnetostrycyjnych.

Jeżeli między dwoma stanami rozmagnesowania i nasycenia objętość próbki nie zmienia się, oznacza to, że zmianie wymiarów w kierunku równoległym (λ_{\parallel} – magnetostrykcja wzdłużna) do pola magnesującego towarzyszy przeciwna i mniejsza o połowę zmiana wymiarów w kierunku prostopadłym (λ_{\perp} – magnetostrykcja poprzeczna); $\lambda_{\perp} = -\lambda_{\parallel}/2$.

Na rysunku 3 podano doświadczalne krzywe pierwotne i pętle histerezy magnetostrykcji λ_{\parallel} i λ_{\perp} magnetyka ceramicznego – ferrytu $\text{Co}_{0,02}\text{Ni}_{0,34}\text{Zn}_{0,64}\text{Fe}_2\text{O}_4$.



Rys. 3. Krzywe pierwotne i pętle histerezy magnetostrykcji $\lambda_{\parallel}(H)$ i $\lambda_{\perp}(H)$ ferrytu $\text{Co}_{0,02}\text{Ni}_{0,34}\text{Zn}_{0,64}\text{Fe}_2\text{O}_4$ [98]
 Fig. 3. Initial curve and magnetostriction hysteresis loops, $\lambda_{\parallel}(H)$ and $\lambda_{\perp}(H)$ of the $\text{Co}_{0,02}\text{Ni}_{0,34}\text{Zn}_{0,64}\text{Fe}_2\text{O}_4$ ferrite [98]

Magnetostrykcja jest zjawiskiem parzystym; zmiana znaku polaryzacji magnetycznej (natężenie pola lub indukcji magnetycznej) nie powoduje zmiany znaku magnetostrykcji.

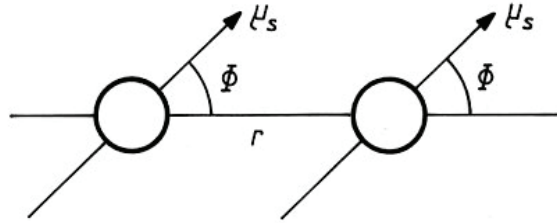
2.2.2. MECHANIZM ATOMOWY MAGNETOSTRYKCJI

Jak już wspomniano wcześniej, magnetostrykcja, podobnie jak i anizotropia magnetokrystaliczna, związana jest z oddziaływaniem atomowych momentów magnetycznych. Rozpatrując oddziaływanie pary atomowych momentów magnetycznych (rys. 4) z uwzględnieniem tylko ich spinów, otrzymuje się zależność (1) określającą wartość energii wymiany. Energia ta nie zależy od położenia spinów względem łączącej je osi wiązania, czyli od kąta (φ). Van Vleck [99] zaproponował model par uporządkowanych, który uwzględni również oddziaływania spin-orbita. W wyniku tego sprzężenia zmiana kierunku położenia spinów powoduje zmianę położenia orbitali elektronowych rozważanych atomów, przez co zmienia się ich elektrostatyczna część energii oddziaływania. W rezultacie energia (w) oddziaływania pary dipoli jest zależna od kąta (φ) położenia jonów.

Ogólnie energię oddziaływania pary dipoli można wyrazić za pomocą wielomianów Legendre'a w postaci [10]:

$$w(r, \cos \Phi) = g(r) + l(r) \left(\cos^2 \Phi - \frac{1}{3} \right) + q(r) \left(\cos^4 \Phi - \frac{6}{7} \cos^2 \Phi + \frac{3}{35} \right) + \dots \quad (9)$$

przy czym r jest odległością między jonami (rys. 4) a l współczynnikiem oddziaływania typu dipol–dipol.



Rys. 4. Para dipoli magnetycznych
Fig. 4. Magnetic dipole pair

Człon pierwszy $g(r)$ nie zależy od kierunku magnetyzacji, zawiera energię wymiany i jest związany z magnetostrykcją objętościową. Wyraz drugi reprezentuje oddziaływanie typu dipol–dipol i jest on uwzględniany jako czynnik wywołujący magnetostrykcję liniową (wzdłużną). Pomijając wyrazy wyższych rzędów, energię oddziaływania pary dipoli można zapisać w postaci

$$w(r, \Phi) = l(r) \left(\cos^2 \Phi - \frac{1}{3} \right) \quad (10)$$

Jeżeli przez $\alpha_1, \alpha_2, \alpha_3$ oznaczymy kosinusy kierunkowe magnetyzacji domeny, a przez $\gamma_1, \gamma_2, \gamma_3$ kosinusy kierunkowe osi wiązania, z zależności (10) otrzymuje się

$$w = l(r) \left[(\alpha_1 \gamma_1 + \alpha_2 \gamma_2 + \alpha_3 \gamma_3)^2 - \frac{1}{3} \right] \quad (11)$$

Rozważmy odkształconą regularną sieć prostą, której składowe tensora odkształceń są określone przez $e_{xx}, e_{yy}, e_{zz}, e_{xy}, e_{yz}$ oraz e_{zx} . Podczas odkształcania kryształu każda para dipoli zmienia zarówno kierunek osi wiązania, jak i jego odległość. Przykładowo dla stanu nieodkształconego energia oddziaływania pary spinów o kierunku wiązania równoległym do osi x jest określona zależnością (11) dla $\gamma_1 = 1, \gamma_2 = \gamma_3 = 0$, to jest

$$w_x = l(r_0) \left(\alpha_1^2 - \frac{1}{3} \right) \quad (12)$$

Jeżeli kryształ zostanie odkształcony, zmieni się długość jego wiązania z r_0 do $r_0(1 + e_{xx})$ oraz zmienią się kosinusy kierunkowe osi wiązania; $\gamma_1 = 1, \gamma_2 = (1/2)e_{xy}, \gamma_3 = (1/2)e_{zx}$.

Energia oddziaływania pary atomów zmieni się wtedy o wartość

$$\Delta w_x = \left(\frac{\partial l}{\partial r} \right) r_0 e_{xx} \left(\alpha_1^2 - \frac{1}{3} \right) + l \alpha_1 \alpha_2 e_{xy} + l \alpha_3 \alpha_1 e_{zx} \quad (13)$$

Podobnie dla par y i z otrzyma się:

$$\Delta w_y = \left(\frac{\partial l}{\partial r} \right) r_0 e_{yy} \left(\alpha_2^2 - \frac{1}{3} \right) + l \alpha_2 \alpha_3 e_{yz} + l \alpha_1 \alpha_2 e_{xy} \quad (14)$$

$$\Delta w_z = \left(\frac{\partial l}{\partial r} \right) r_0 e_{zz} \left(\alpha_3^2 - \frac{1}{3} \right) + l \alpha_3 \alpha_1 e_{zx} + l \alpha_2 \alpha_3 e_{yz} \quad (15)$$

Sumując wyrażenie (13), (14) i (15) dla wszystkich sąsiadujących bezpośrednio z sobą par w jednostce objętości, otrzyma się [10]

$$E_M = B_1 \left[e_{xx} \left(\alpha_1^2 - \frac{1}{3} \right) + e_{yy} \left(\alpha_2^2 - \frac{1}{3} \right) + e_{zz} \left(\alpha_3^2 - \frac{1}{3} \right) \right] + B_2 (e_{xy} \alpha_1 \alpha_2 + e_{yz} \alpha_2 \alpha_3 + e_{zx} \alpha_3 \alpha_1) \quad (16)$$

przy czym:

$$B_1 = N \left(\frac{\partial l}{\partial r} \right) r_0, \quad B_2 = 2Nl \quad (17)$$

są stałymi sprzężenia magnetosprężystego (stałe magnetosprężyste) [10, 8]), a N jest liczbą atomów w jednostce objętości.

Energia E_M , zależna od składowych tensora odkształceń (wywołanych porządkiem się spinów) oraz od kierunku magnetyzacji domeny, nazywana jest energią magnetostrykcji. Energia E_M reprezentuje oddziaływanie między energią anizotropii magnetokrystalicznej a energią sprężystą.

Zależność (16) jest również słuszna w przypadku sieci regularnej, przestrzennie i płasko centrowanej. Dla sieci przestrzennie centrowanej otrzymuje się

$$B_1 = \frac{8}{3} Nl, \quad B_2 = \frac{8}{9} N \left[l + \left(\frac{\partial l}{\partial r} \right) r_0 \right] \quad (18)$$

natomiast dla sieci płasko centrowanej

$$B_1 = \frac{1}{2} N \left[6l + \left(\frac{\partial l}{\partial r} \right) r_0 \right], \quad B_2 = N \left[2l + \left(\frac{\partial l}{\partial r} \right) r_0 \right] \quad (19)$$

Odształceniom magnetostrykcyjnym kryształu przeciwdziała energia sprężysta E_L , która jest kwadratową funkcją odkształcenia kryształu regularnego

$$E_L = \frac{1}{2}c_{11}(e_{xx}^2 + e_{yy}^2 + e_{zz}^2) + \frac{1}{2}c_{44}(e_{xy}^2 + e_{yz}^2 + e_{zx}^2) + c_{12}(e_{yy}e_{zz} + e_{yy}e_{xx} + e_{zz}e_{xx}) \quad (20)$$

przy czym c_{11} , c_{12} , c_{44} są stałymi sprężystymi.

Energia sprężysta E_L wzrasta aż do osiągnięcia stanu równowagi wynikającego z minimum energii całkowitej E względem składowych tensora odkształceń.

$$E = E_M + E_L \quad (21)$$

Prowadzi to do 6 następujących równań

$$\frac{\partial E}{\partial e_{xx}} = B_1 \left(\alpha_1^2 - \frac{1}{3} \right) + c_{11}e_{xx} + c_{12}(e_{yy} + e_{zz}) = 0 \quad (22)$$

$$\frac{\partial E}{\partial e_{yy}} = B_1 \left(\alpha_2^2 - \frac{1}{3} \right) + c_{11}e_{yy} + c_{12}(e_{zz} + e_{xx}) = 0 \quad (23)$$

$$\frac{\partial E}{\partial e_{zz}} = B_1 \left(\alpha_3^2 - \frac{1}{3} \right) + c_{11}e_{zz} + c_{12}(e_{xx} + e_{yy}) = 0 \quad (24)$$

$$\frac{\partial E}{\partial e_{xy}} = B_2\alpha_1\alpha_2 + c_{44}e_{xy} = 0 \quad (25)$$

$$\frac{\partial E}{\partial e_{yz}} = B_2\alpha_2\alpha_3 + c_{44}e_{yz} = 0 \quad (26)$$

$$\frac{\partial E}{\partial e_{zx}} = B_2\alpha_3\alpha_1 + c_{44}e_{zx} = 0 \quad (27)$$

Rozwiązaniem tego układu równań są odkształcenia odpowiadające stanowi równowagi

$$e_{xx} = \frac{-B_1}{c_{11} - c_{12}} \left(\alpha_1^2 - \frac{1}{3} \right), \quad e_{yy} = \frac{-B_1}{c_{11} - c_{12}} \left(\alpha_2^2 - \frac{1}{3} \right)$$

$$e_{zz} = \frac{-B_1}{c_{11} - c_{12}} \left(\alpha_3^2 - \frac{1}{3} \right), \quad e_{xy} = \frac{-B_2}{c_{44}} \alpha_1 \alpha_2$$

$$e_{yz} = \frac{-B_2}{c_{44}} \alpha_2 \alpha_3, \quad e_{zx} = \frac{-B_2}{c_{44}} \alpha_3 \alpha_1 \quad (28)$$

Odształcenie względne, obserwowane wzdłuż kierunku wyznaczonego kosinusami kierunkowymi $\beta_1, \beta_2, \beta_3$, określone jest wyrażeniem

$$\frac{\Delta l}{l} = e_{xx}\beta_1^2 + e_{yy}\beta_2^2 + e_{zz}\beta_3^2 + e_{xy}\beta_1\beta_2 + e_{yz}\beta_2\beta_3 + e_{zx}\beta_3\beta_1 \quad (29)$$

Po uwzględnieniu zależności (28) otrzymuje się

$$\begin{aligned} \frac{\Delta l}{l} = & \frac{-B_1}{c_{11} - c_{12}} \left(\alpha_1^2 \beta_1^2 + \alpha_2^2 \beta_2^2 + \alpha_3^2 \beta_3^2 - \frac{1}{3} \right) \\ & + \frac{-B_2}{c_{44}} (\alpha_1 \alpha_2 \beta_1 \beta_2 + \alpha_2 \alpha_3 \beta_2 \beta_3 + \alpha_3 \alpha_1 \beta_3 \beta_1) \end{aligned} \quad (30)$$

Przy położeniu wektora magnetyzacji domeny wzdłuż kierunku [100] odształcenie magnetostrykcyjne λ_{100} (stała magnetostrykcji) wzdłuż tego kierunku otrzymuje się przez podstawienie $\alpha_1 = \beta_1 = 1$ oraz $\alpha_2 = \alpha_3 = \beta_2 = \beta_3 = 0$, czyli

$$\lambda_{100} = -\frac{2}{3} \frac{B_1}{c_{11} - c_{12}} \quad (31)$$

Analogicznie w przypadku położenia wektora magnetyzacji wzdłuż kierunku [111] odształcenie magnetostrykcyjne λ_{111} wzdłuż tego kierunku otrzymuje się przez podstawienie $\alpha_i = \beta_i = 1/\sqrt{3}$ ($i = 1, 2, 3$);

$$\lambda_{111} = -\frac{1}{3} \frac{B_2}{c_{44}} \quad (32)$$

Podstawiając stałe magnetostrykcji λ_{100} i λ_{111} do zależności (2.30) otrzymuje się wyrażenie określające magnetostrykcję kryształu regularnego:

$$\begin{aligned} \lambda_{si} = \frac{\Delta l}{l} = & \frac{3}{2} \lambda_{100} \left(\alpha_1^2 \beta_1^2 + \alpha_2^2 \beta_2^2 + \alpha_3^2 \beta_3^2 - \frac{1}{3} \right) \\ & + 3 \lambda_{111} (\alpha_1 \alpha_2 \beta_1 \beta_2 + \alpha_2 \alpha_3 \beta_2 \beta_3 + \alpha_3 \alpha_1 \beta_3 \beta_1) \end{aligned} \quad (33)$$

Podstawiając zależności (17), (18), (19) do (31) i (32), otrzymuje się dla sieci regularnej prostej

$$\lambda_{100} = -\frac{2}{3} \frac{N}{c_{11} - c_{12}} \left(\frac{\partial l}{\partial r} \right) r_0; \quad \lambda_{111} = -\frac{4}{3} \frac{N}{c_{44}} l \quad (34)$$

a dla sieci regularnej przestrzennie centrowanej

$$\lambda_{100} = -\frac{16}{9} \frac{N}{c_{11} - c_{12}} l; \quad \lambda_{111} = -\frac{16}{27} \frac{N}{c_{44}} \left[l + \left(\frac{\partial l}{\partial r} \right) r_0 \right] \quad (35)$$

natomiast dla sieci regularnej płasko centrowanej

$$\lambda_{100} = -\frac{1}{3} \frac{N}{c_{11} - c_{12}} \left[6l + \left(\frac{\partial l}{\partial r} \right) r_0 \right]; \quad \lambda_{111} = -\frac{2}{3} \frac{N}{c_{44}} \left[2l + \left(\frac{\partial l}{\partial r} \right) r_0 \right] \quad (36)$$

Zależność (33), o charakterze fenomenologicznym nazywana w literaturze równaniem dwóch stałych, wyprowadzona została między innymi przez Lee [8] i Kittela [97]. W literaturze [8, 10] zależność (33) podawana jest również w postaci zapisu wprowadzającego stałe h_1 i h_2 , przy czym $h_1 = 3/2 \lambda_{100}$ i $h_2 = 3/2 \lambda_{111}$.

Zależność (33) określa odkształcenie magnetostrykcyjne nasycenia λ_{si} kryształu regularnego, w kierunku wyznaczonym w odniesieniu do osi kryształu, kosinusami kierunkowymi $\beta_1, \beta_2, \beta_3$, jeżeli jest on magnesowany od stanu rozmagnesowania do nasycenia w kierunku określonym kosinusami kierunkowymi $\alpha_1, \alpha_2, \alpha_3$.

W stanie nasycenia cały kryształ staje się z definicji pojedynczą domeną, w której magnetyzacja M_s jest równoległa do kierunku przyłożonego pola. Natomiast stan rozmagnesowania nie jest jednoznacznie zdefiniowany, ponieważ wymagany jest jedynie taki warunek, ażeby suma wektorów magnetyzacji była równa zero. Warunek ten może spełniać dowolna liczba konfiguracji odpowiadająca wypadkowej magnetyzacji równej zero. Zależność (33) opiera się na arbitralnym założeniu, że w stanie rozmagnesowania domeny dla poszczególnych kierunków magnetyzacji M_s mają jednakowe objętości. Przykładowo w kryształach żelaza odnosi się to do wszystkich sześciu rodzajów domen – [100], [100], [010], [010], [001], [001]. Jeżeli założony stan „idealnego rozmagnesowania” nie jest zrealizowany, wzór (33) nie jest słuszny i zmierzone wartości magnetostrykcji będą różniły się od obliczonych. Różnice w stanach rozmagnesowania próbek mogą być również przyczyną różnych wartości magnetostrykcji, otrzymanych doświadczalnie przez różnych autorów dla tych samych materiałów. Zagadnienie zdefiniowania stanu rozmagnesowania próbki i jego wpływ na wyniki doświadczalne w pomiarach właściwości magnetycznych omówiono w pracy [100].

Symbol λ_{si} wprowadził do zależności (33) Cullity [9], oznaczając indeksem i stan idealnego rozmagnesowania. Umożliwia to rozróżnienie dwóch rodzajów magnetostrykcji:

- 1) λ_{si} – wartość zmierzona w warunkach idealnego rozmagnesowania. Wartość ta jest stałą materiałową określoną zależnością (33), w której stałe λ_{100} i λ_{111} są również wartościami λ_{si} dla kierunków krystalograficznych [100] i [111], ($\lambda_{si[100]}$, $\lambda_{si[111]}$),
- 2) λ_s – wartość zmierzona i prawdziwa tylko dla danej mierzonej próbki o „nieidealnym” i bliżej niezdefiniowanym stanie rozmagnesowania, w którym wszystkie możliwe rodzaje domen różnią się objętością. Oznacza to preferowa-

ny stan orientacji domen, pomimo wypadkowej magnetyzacji równej zeru. Stan taki może wystąpić zarówno w monokryształach, jak i w pojedynczych ziarnach polikryształów oraz w zorientowanej strukturze ziaren w materiale polikrystalicznym. Wartość λ_s silnie zależy od magnetycznej, termicznej i mechanicznej historii próbki.

W praktycznym przypadku poszukiwania odkształcenia magnetostrykcyjnego w kierunku zgodnym z polem magnesującym ($\alpha_i = \beta_i$, $i = 1, 2, 3$), zależność (33) upraszcza się do postaci:

$$\lambda_{si} = \lambda_{100} + 3(\lambda_{111} - \lambda_{100})(\alpha_1^2\alpha_2^2 + \alpha_2^2\alpha_3^2 + \alpha_3^2\alpha_1^2) \quad (37)$$

Zakładając, że magnetostrykcja jest izotropowa, tj. $\lambda_{100} = \lambda_{111} = \lambda_{si}$, zależność (33) otrzyma postać:

$$\lambda_\theta = \frac{3}{2}\lambda_{si} + \left[(\alpha_1\beta_1 + \alpha_2\beta_2 + \alpha_3\beta_3)^2 - \frac{1}{3} \right] = \frac{3}{2}\lambda_{si} \left(\cos^2\theta - \frac{1}{3} \right) \quad (38)$$

przy czym λ_θ jest odkształceniem magnetostrykcyjnym w kierunku tworzącym kąt θ z kierunkiem magnetyzacji i pola.

Wartości stałych magnetosprężystych B_1 i B_2 oraz stałych magnetostrykcji λ_{100} i λ_{111} ferrytów, żelazowego (magnetytu), kobaltowego, niklowego, manganowego, litowego oraz ferrytów: kobaltowo-manganowego, kobaltowo-niklowego, kobaltowo-cynkowego, manganowo-cynkowego podano w pracy [101]. Wartości stałych sprężystych c_{11} , c_{12} , c_{44} żelaza i niklu oraz ferrytu niklowego podano odpowiednio w pracach [10] oraz [102]. Dane doświadczalne wartości stałych magnetostrykcji λ_{100} i λ_{111} ferrytów Mn-Zn zawierają prace [10, 103], ferrytów Ni-Zn [10, 104], ferrytów Co-Zn [10], ferrytów Co-Mn [105].

2.2.3. MAGNETOSTRYKCJA POLIKRYSTAŁÓW

Magnetostrykcja nasycenia próbki polikrystalicznej określona jest stałą λ_p [9]. Zależy ona od właściwości magnetostrykcyjnych poszczególnych kryształów oraz ich rozkładu geometrycznego ze względu na obecność wyróżnionych kierunków. Jeżeli orientacja ziaren jest całkowicie losowa, to λ_p może być wyznaczona przez uśrednienie magnetostrykcji tych ziaren względem poszczególnych kierunków. W procesie magnesowania polikryształu poszczególne ziarna, ze względu na różne orientacje w stosunku do pola magnesującego, odkształcają się magnetostrykcyjnie o inne wartości niż ziarna sąsiednie. W procesie tym można wyróżnić dwa przypadki: 1 – naprężenia w poszczególnych ziarnach są stałe, a odkształcenia różne, 2 – odkształcenia są jednakowe, a zmieniają się naprężenia.

Zwykle za fizycznie bardziej uzasadniony rozpatrywany jest przypadek jednakowych odkształceń.

W przypadku kryształów regularnych przez uśrednienie zależności (37) względem wszystkich możliwych kierunków wektora M_s otrzymuje się średnią wartość magnetostrykcji λ_{si} (w kierunku pola), [9, 8]:

$$\bar{\lambda}_{si} = \frac{2\lambda_{100} + 3\lambda_{111}}{5} \quad (39)$$

Wzór (39) prawidłowy jest również dla ferrytów.

Callen i Goldberg [67], przyjmując założenie pośrednie między przypadkiem jednakowych naprężeń a jednakowych odkształceń, otrzymali następującą zależność

$$\bar{\lambda}_{si} = \lambda_{111} + \left(\frac{\frac{2}{5} - \ln c}{8} \right) (\lambda_{100} - \lambda_{111}) \quad (40)$$

przy czym $c = 2c_{44}/(c_{11} - c_{12})$, a c_{44} , c_{11} , c_{12} są stałymi sprężystymi monokryształu. Kryształ jest sprężyste izotropowy, jeżeli $c = 1$, a więc $c_{44} = (c_{11} - c_{12})/2$ i w tym przypadku zależności (40) i (39) stają się identyczne. Zależność (39) w wielu przypadkach polikrystalicznych metali, stopów oraz ferrytów daje dobrą zgodność wartości obliczonych z doświadczalnymi.

Odkształcenie magnetostrykcyjne polikryształu λ_θ wzdłuż kierunku tworzącego kąt θ z kierunkiem magnetyzacji może być określone przez wyrażenie (38), w którym λ_{si} zastąpione zostało przez λ_{pi} :

$$\lambda_\theta = \frac{3}{2} \lambda_{pi} \left(\cos^2 \theta - \frac{1}{3} \right) \quad (41)$$

Zależność (41) jest słuszna dla materiału izotropowego i stosowanie jej wymaga próbek bez uprzywilejowanych orientacji ziaren.

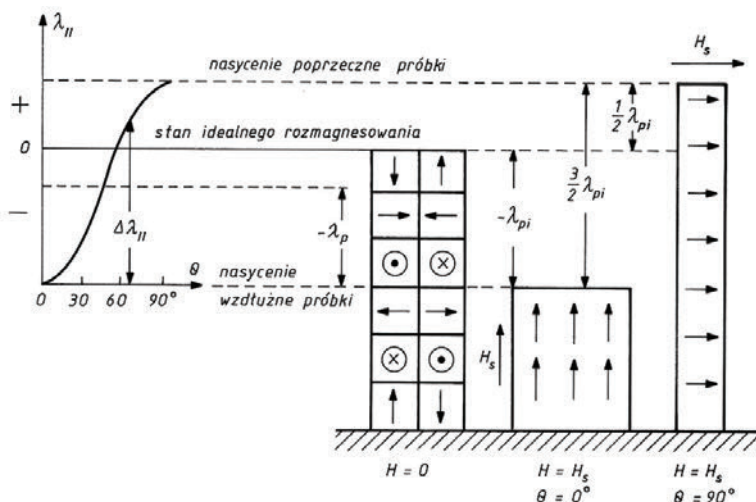
Zależność (41) ma istotne znaczenie praktyczne ze względu na możliwość pomiaru λ_{pi} . Odpowiednie zastosowanie tej zależności umożliwia pomiar λ_{pi} niezależnie od stanu rozmagnesowania próbki.

Rozważmy próbkę „idealnie” rozmagnesowaną, a następnie namagnesowaną do nasycenia wzdłuż długości polem H_s (rys. 5). W nasyceniu wszystkie wektory magnetyzacji są zgodne z kierunkiem pola H_s ; zanikają ściany domenowe i próbka staje się jednodomenowa. Wzdłużne odkształcenie magnetostrykcyjne próbki (w rozpatrywanym przypadku ujemne) zgodne z kierunkiem pola H_s według zależności (41) odpowiada $\lambda(\theta = 0) = \lambda_{pi}$.

Przyjmując to odkształcenie próbki za stan początkowy, otrzymuje się w wyniku obrotu pola H_s o kąt θ (w stosunku do podłużnej osi próbki) następującą zmianę $\Delta\lambda_{||}$ wzdłużnego odkształcenia próbki:

$$\Delta\lambda_{\parallel} = \lambda_{\theta}(\theta = \theta) - \lambda_{\theta}(\theta = 0) = +\frac{3}{2}\lambda_{pi}\left(\cos^2\theta - \frac{1}{3}\right) - \lambda_{pi} = -\frac{3}{2}\lambda_{pi}\sin^2\theta \quad (42)$$

Jeżeli wartość λ_{pi} jest ujemna (jak w przypadku ferrytów), to $\Delta\lambda_{\parallel} = 3/2\lambda_{pi}\sin^2\theta$. Zmiany $\Delta\lambda_{\parallel}$ w funkcji kąta obrotu θ przedstawiono na rysunku 5. Jeżeli kąt θ zmieni się o 90° , to odpowiadająca mu zmiana $\Delta\lambda_{\parallel} = 3/2\lambda_{pi}$.



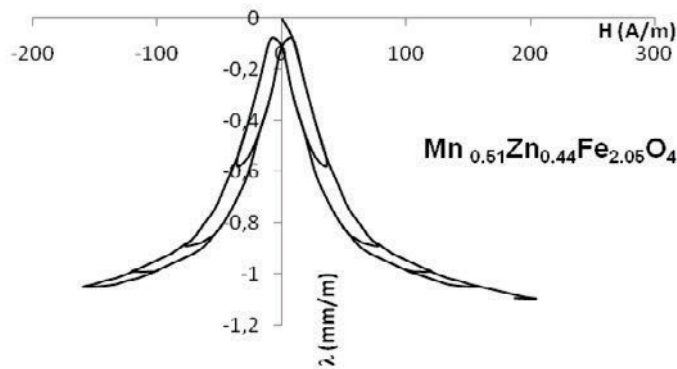
Rys. 5. Zmiany magnetostrykcji wzdłużnej λ_{\parallel} próbki w funkcji kąta θ obrotu pola H_s oraz schematyczne zmiany wymiarów wzdłużnych próbki (przypadek ujemnej wartości λ_{pi}) [98]

Fig. 5. Magnetostrictive strains λ_{\parallel} as a function of θ angle of the rotation of H_s and schematic changes of the elongation of the sample [98]

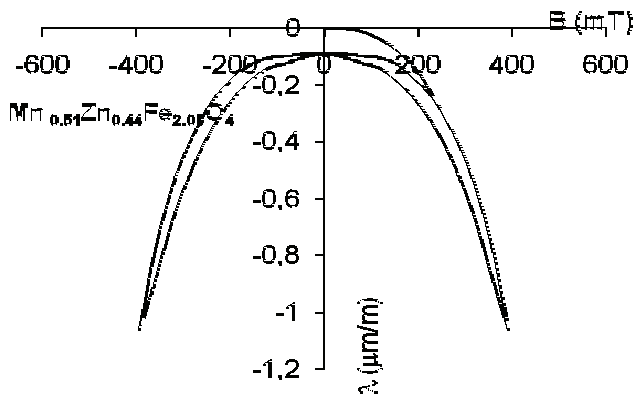
Jeżeli stan rozmagnesowania próbki nie jest „idealny”, to następuje przesunięcie zera na rysunku 5 w dół lub w górę i występująca wartość λ_{pi} staje się wartością λ_p . Ma to istotne znaczenie praktyczne podczas interpretacji wyników pomiaru magnetostrykcji, ponieważ tylko w stanie „idealnego rozmagnesowania” próbki wzdłużne odkształcenie λ_{\parallel} w wyniku namagnesowania próbki polem H_s jest równe wartości λ_{pi} . W każdym innym stanie rozmagnesowania doświadczalne wartości λ_{\parallel} , zmierzone dla tego samego materiału (przy nasyceniu próbki) jako wartości λ_p , mogą różnić się znacznie między sobą. Może to być przyczyną różnych wartości λ_p otrzymanych doświadczalnie dla tych samych materiałów przez różnych autorów. W skrajnym przypadku, jeżeli w stanie rozmagnesowania już przed pomiarem λ_{\parallel} powstanie, w kierunku równoległym do pola, jednoosiowa struktura domenowa (np. w wyniku działania naprężeń od sił zewnętrznych), to otrzymane wartości λ_{\parallel} przy polu H_s będą zbliżone do zera, co nie oznacza, że wartość magnetostrykcji nasycenia tego materiału jest bliska zera.

Wylimitowanie wspomnianych błędów jest możliwe przez pomiar zmiany $\Delta\lambda_{\parallel}$ odkształcenia wzdłużnego powodowanej rotacją wektora H_s o znany kąt θ . Ze względów technicznych przyjmuje się zwykle $\theta = 90^\circ$. Dane doświadczalne magnetostrykcyjnych właściwości polikrystalicznych ferrytów Ni-Zn zawierających kobalt przedstawiono w pracach [109, 110], natomiast zależności temperaturowe tych właściwości w pracy [111].

Przykładowe wyniki badań własnych autorów, magnetostrykcyjnych właściwości wybranych ferrytów podano na rysunkach 6–13. Właściwości magnetostrykcyjne $\lambda(H)$ ferrytu $\text{Mn}_{0,51}\text{Zn}_{0,44}\text{Fe}_{2,05}\text{O}_4$, przedstawiające krzywą pierwotną i pętlę histerezy magnetostrykcyjnej, podano na rysunku 6. Na rysunku 7 podano natomiast krzywą pierwotną i pętlę histerezy magnetostrykcyjnej w zależności od indukcji magnetycznej B , (charakterystyki $\lambda(B)$), dla tego samego ferrytu [106].

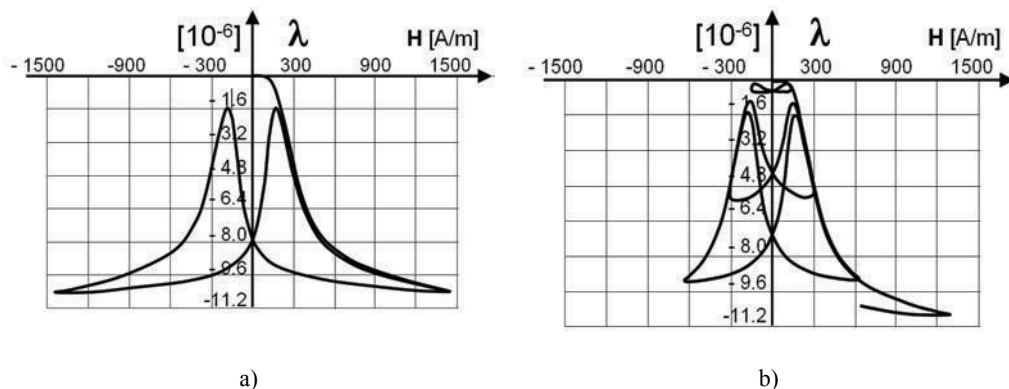


Rys. 6. Krzywa pierwotna i pętle histerezy magnetostrykcyjnej $\lambda(H)$ ferrytu $\text{Mn}_{0,51}\text{Zn}_{0,44}\text{Fe}_{2,05}\text{O}_4$ [106]
 Fig. 6. The initial curve and magnetostriction hysteresis loops $\lambda(H)$ of the $\text{Mn}_{0,51}\text{Zn}_{0,44}\text{Fe}_{2,05}\text{O}_4$, ferrite [106]



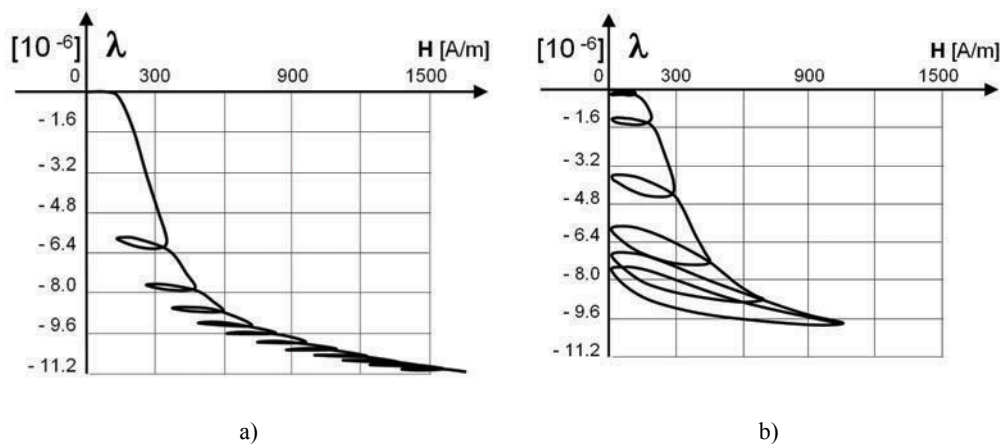
Rys. 7. Krzywa pierwotna i pętle histerezy magnetostrykcyjnej $\lambda(B)$ ferrytu $\text{Mn}_{0,51}\text{Zn}_{0,44}\text{Fe}_{2,05}\text{O}_4$ [106]
 Fig. 7. The initial curve and magnetostriction hysteresis loops $\lambda(B)$ of the $\text{Mn}_{0,51}\text{Zn}_{0,44}\text{Fe}_{2,05}\text{O}_4$, ferrite [106]

Charakterystyki magnetostrykcyjne ferrytu niklowego $\text{Co}_{0,004}\text{Cu}_{0,12}\text{Ni}_{0,866}\text{Fe}_{2,01}\text{O}_{4,02}$ przedstawiono na rysunkach 8 i 9. Na rysunku 8 podano krzywą pierwotną i pętlę histerezy magnetostrykcji (a) oraz cząstkowe pętle histerezy magnetostrykcji (b), natomiast na rysunku 9 przedstawiono rewersyjne krzywe magnetostrykcji dla tego samego ferrytu.



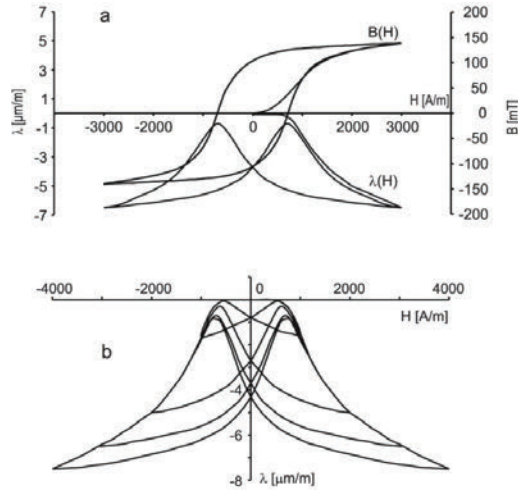
Rys. 8. Krzywa pierwotna i pętla histerezy magnetostrykcji (a) oraz cząstkowe pętle histerezy magnetostrykcji (b) ferrytu $\text{Co}_{0,004}\text{Cu}_{0,12}\text{Ni}_{0,866}\text{Fe}_{2,01}\text{O}_{4,02}$, [107]

Fig. 8. Initial curve, major and minor magnetostriction hysteresis loops (butterfly loops) as the functions of static magnetic field of the $\text{Co}_{0,004}\text{Cu}_{0,12}\text{Ni}_{0,866}\text{Fe}_{2,01}\text{O}_{4,02}$ ferrite [107]



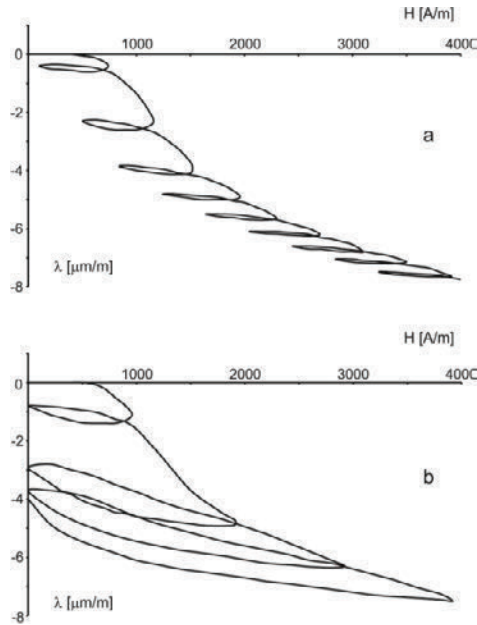
Rys. 9. Rewersyjne pętle histerezy magnetostrykcji dla amplitudy pola magnesującego 240 A/m (a) i krzywe magnetostrykcji po każdym powrocie pola magnesującego do zera (b), ferrytu $\text{Co}_{0,004}\text{Cu}_{0,12}\text{Ni}_{0,866}\text{Fe}_{2,01}\text{O}_{4,02}$ [107]

Fig. 9. Reversible magnetostriction loops with the amplitude of the magnetic field equal to 240 A/m (a) and magnetostriction curve after each returning to zero (b) of the $\text{Co}_{0,004}\text{Cu}_{0,12}\text{Ni}_{0,866}\text{Fe}_{2,01}\text{O}_{4,02}$ ferrite [107]



Rys. 10. Krzywa pierwotna i pętla histerezy $B(H)$ oraz krzywa pierwotna i pętla histerezy $\lambda(H)$ (a),
 cząstkowe pętli histerezymagnetostrykcji $\lambda(H)$ ferrytu nikowego (b) [108]

Fig. 10. Initial curve, induction hysteresis loop and magnetostriction hysteresis loop (a),
 major and minor magnetostriction butterfly of Ni-ferrite (b) [108]

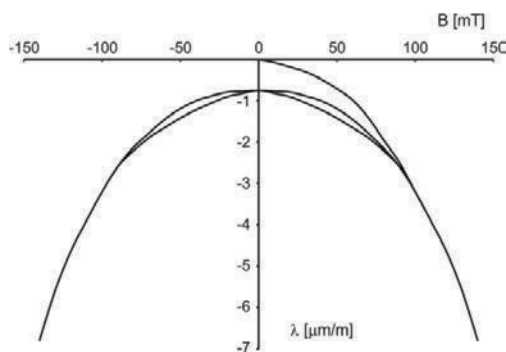


Rys. 11. Rewersyjne pętli histerezy magnetostrykcji dla amplitudy pola magnesującego 400 A/m (a)
 i krzywe magnetostrykcji po każdym powrocie pola magnesującego do zera ferrytu nikowego (b) [108]

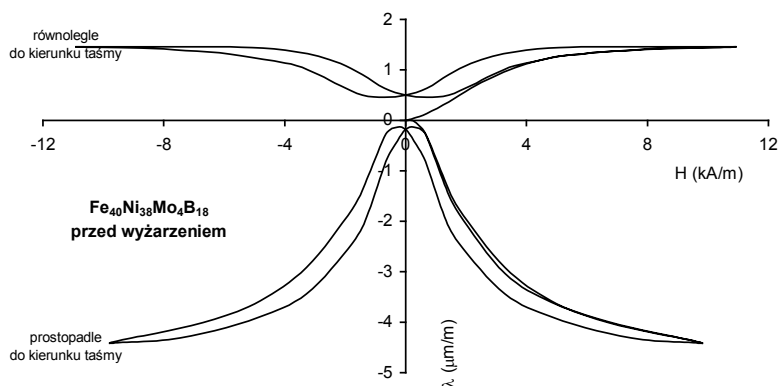
Fig. 11. Magnetostriction curves with the increasing of bias field at the constant 400 A/m amplitude
 of the magnetic field (a) and magnetostriction curves after each returning to zero of the magnetic field
 of Ni-ferrite [108]

Właściwości magnetostrykcyjne ferrytu niklowego NiFe_2O_4 podano na rysunkach 10–2.12 [108]. Na rysunku 10 podano krzywe pierwotne i pętle histerezy magnetostrykcyjnej $\lambda(H)$ w zależności od pola magnesującego H , na rys. 11 pokazano natomiast rewersyjne pętle magnetostrykcyjnej dla tego samego ferrytu. Krzywe pierwotne i pętle histerezy magnetostrykcyjnej dla ferrytu niklowego, w zależności zmian indukcji magnetycznej B , podano na rysunku 12.

Krzywą pierwotną i pętlę histerezy magnetostrykcyjnej $\lambda(B)$ w zależności od zmian indukcji magnetycznej B dla ferrytu niklowego podano na rysunku 12, natomiast na rysunku 13 podano krzywą pierwotną i pętlę histerezy magnetostrykcyjnej mierzonej w kierunku prostopadłym i równoległym do kierunku taśmy z magnetyka amorficznego o składzie $\text{Fe}_{40}\text{Ni}_{38}\text{Mo}_4\text{B}_{18}$ w stanie wyjściowym (bez przeprowadzania relaksacji termicznej).



Rys. 12. Krzywa pierwotna i pętla histerezy $\lambda(B)$ ferrytu niklowego[108]
Fig. 12. Magnetostriction hysteresis $\lambda(B)$ of Ni-ferrite



Rys. 13. Krzywe pierwotne i pętle histerezy magnetostrykcyjnej mierzonej w kierunku prostopadłym i równoległym do kierunku taśmy z magnetyka amorficznego o składzie $\text{Fe}_{40}\text{Ni}_{38}\text{Mo}_4\text{B}_{18}$ w stanie wyjściowym [112]
Fig. 13. Initial curve and hysteresis loop of amorphous alloy in as-quenched state measured in both parallel and perpendicular directions [112]

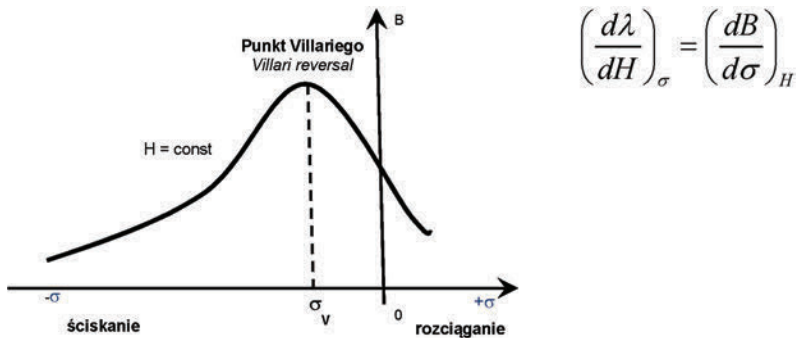
3. WŁAŚCIWOŚCI MAGNETOSPREŻYSTE MAGNETYKÓW. ZJAWISKO VILLARIEGO

3.1. WPROWADZENIE

Zmianom stanu magnetycznego materiału towarzyszy zazwyczaj deformacja jego sieci krystalicznej. Następstwem tej deformacji jest omawiany wcześniej efekt magnetostrykcyjny oraz naprężenia własne w materiale. Termodynamicznie odwrotnym zjawiskiem jest efekt magnetosprężysty (nazywany również odwrotnym efektem magnetostrykcyjnym), czyli zmiana stanu magnetycznego magnetyka pod wpływem naprężeń wywołanych siłami zewnętrznymi. Efekt magnetosprężysty został odkryty przez Villariego w 1863 r.

Efekt ten może być obserwowany jako zmiana kształtu pętli histerezy magnetycznej lub jako zmiany indukcji magnetycznej B przy danym natężeniu pola magnesującego H , po przyłożeniu siły zewnętrznej do próbki z materiału ferro- lub ferrimagnetycznego. Zmiany parametrów magnetycznych materiału należy jednak odnosić do wartości naprężeń (σ) w materiale, gdyż to one, nie wartość przyłożonej siły, decydują o wielkości efektu magnetosprężystego.

Stanowi, w którym osiąga się minimum energii swobodnej materiału w wyniku działania naprężeń σ , odpowiada maksimum na krzywej $B(\pm\sigma)_H$. Maksimum to, w którym $(\partial B/\partial\sigma)_H = 0$, nosi nazwę punktu Villariego (ang. *Villari reversal*), a samo zjawisko występowania ekstremum na krzywej $B(\pm\sigma)_H$ – efektem Villariego.



Rys. 14. Schematyczna zależność indukcji magnetycznej B od naprężeń $\pm\sigma$ (przy $H = \text{const}$) w zakresie występowania zjawiska Villariego

Fig. 14. Stress $\pm\sigma$ dependence of flux density B (for constant magnetizing field H) within Villari effect range

Efekty magnetostrykcyjny i magnetosprężysty powiązane są zależnością $(\partial\lambda/\partial H)_\sigma = (\partial B/\partial\sigma)_H = d$, przy czym d jest czułością piezomagnetyczną magnetyka [25]. Naj-

mniejszą czułość piezomagnetyczną materiał wykazuje w punkcie Villariego, tzn. wówczas, gdy przyrost naprężeń σ nie powoduje zmian indukcji B materiału.

3.2. WPŁYW NAPRĘŻEŃ NA INDUKCJĘ MAGNETYCZNĄ

Całkowitą energię swobodną E próbki znajdującej się w polu magnetycznym można zapisać jako sumę poszczególnych energii swobodnych

$$E = E_H + E_D + E_K + E_\sigma + E_W + E_0 \quad (43)$$

przy czym: E_H jest energią związaną z magnesowaniem próbki w polu H , E_D energią od magnesowania związaną z własnym polem próbki, E_K energią anizotropii magnetycznej, E_σ energią magnetosprężystą, E_W energią wymiany, E_0 są składnikami energii związanymi z niejednorodnością próbki.

Stabilny stan równowagi dla namagnesowanej próbki wyznacza się ze zminimalizowania wyrażenia (43) ze względu na energię swobodną. Minimalizacja ta prowadzi do nieliniowych równań różniczkowych cząstkowych [113]. Po wprowadzeniu pewnych założeń upraszczających z zależności (43) można wyprowadzić opis zachowania się materiału w obecności pola magnesującego i naprężeń σ .

Jeżeli rozważa się próbkę jednorodną, bez przyłożonego zewnętrznego pola magnesującego, to $E_H = 0$ i $E_0 = 0$. Energię wymiany E_W można pominąć ze względu na jej izotropowość oraz krótki zasięg oddziaływania. Pominąć również można energię od magnesowania E_D , ze względu na konfigurację domen. Praktyczne badania można więc ograniczyć do rozpatrzenia wpływu dwóch pozostałych składników energii: E_K – energii anizotropii magnetycznej oraz E_σ – energii magnetosprężystej.

W przypadku braku naprężeń kierunek magnetyzacji M_s wymuszony jest anizotropią magnetyczną E_K . Dla kryształów sieci regularnej można ją wyrazić zależnością [9]

$$E_K = K_1(\alpha_1^2\alpha_2^2 + \alpha_2^2\alpha_3^2 + \alpha_3^2\alpha_1^2) + K_2\alpha_1^2\alpha_2^2\alpha_3^2 + K_3(\alpha_1^4\alpha_2^4 + \dots) \quad (44)$$

przy czym: K_1, K_2, K_3 są stałymi anizotropii magnetycznej, $\alpha_1, \alpha_2, \alpha_3$ odpowiednimi kosinusami kątów określających kierunek magnetyzacji M_s w stosunku do osi krytalograficznych.

Utożsamianie jednak E_K tylko ze stałą K_1 , jak wykazał Kulikowski [114], nie zawsze jest słuszne i to nie tylko formalnie przez pominięcie dalszych wyrazów w szeregu.

Jeżeli w materiale istnieją naprężenia σ , wywołane działaniem sił zewnętrznych, przy czym $\gamma_1, \gamma_2, \gamma_3$ są odpowiednimi kosinusami kątów określających kierunek tych naprężeń w stosunku do osi krytalograficznych, to oznaczając składowe tensora naprężeń przez $\sigma_{ij} = \sigma\gamma_i\gamma_j$, otrzymuje się odpowiednie składowe odkształceń e_{ij} [10]:

$$e_{xx} = \sigma[s_{11}\gamma_1^2 + s_{12}(\gamma_2^2 + \gamma_3^2)\dots] \quad (45)$$

$$e_{xy} = \sigma s_{44}\gamma_1\gamma_2$$

przy czym s_{11} , s_{12} , s_{44} są współczynnikami sprężystości.

Podstawiając zależności (45) do (16) otrzymuje się:

$$E_\sigma = B_1\sigma(s_{11} - s_{12})\left(\alpha_1^2\gamma_1^2 + \alpha_2^2\gamma_2^2 + \alpha_3^2\gamma_3^2 - \frac{1}{3}\right) \quad (46)$$

$$+ B_2\sigma s_{44}(\alpha_1\alpha_2\gamma_1\gamma_2 + \alpha_2\alpha_3\gamma_2\gamma_3 + \alpha_3\alpha_1\gamma_3\gamma_1)$$

Wyznaczając stałe B_1 i B_2 w zależności od λ_{100} i λ_{111} odpowiednio ze wzorów (31) i (32) oraz uwzględniając relacje między c_{11} , c_{12} i c_{44} oraz s_{11} , s_{12} i s_{44}

$$c_{11} = \frac{s_{11} + s_{12}}{(s_{11} - s_{12})(s_{11} + 2s_{12})}; \quad c_{12} = \frac{-s_{12}}{(s_{11} - s_{12})(s_{11} + 2s_{12})}; \quad c_{44} = \frac{1}{s_{44}} \quad (47)$$

otrzymuje się

$$E_\sigma = -\frac{3}{2}\lambda_{100}\sigma(\alpha_1^2\gamma_1^2 + \alpha_2^2\gamma_2^2 + \alpha_3^2\gamma_3^2) \quad (48)$$

$$- 3\lambda_{111}\sigma(\alpha_1\alpha_2\gamma_1\gamma_2 + \alpha_2\alpha_3\gamma_2\gamma_3 + \alpha_3\alpha_1\gamma_3\gamma_1)$$

Po wprowadzeniu do magnetyka naprężeń σ kierunek magnetyzacji M_s wymuszony jest zarówno przez σ , jak i anizotropię magnetokrystaliczną, a energia swobodna próbki dla kryształów regularnych określona jest zależnością [9]:

$$E = K_1(\alpha_1^2\alpha_2^2 + \alpha_2^2\alpha_3^2 + \alpha_3^2\alpha_1^2) + K_2\alpha_1^2\alpha_2^2\alpha_3^2 + K_3(\alpha_1^4\alpha_2^4 + \dots) \quad (49)$$

$$- \frac{3}{2}\lambda_{100}\sigma(\alpha_1^2\gamma_1^2 + \alpha_2^2\gamma_2^2 + \alpha_3^2\gamma_3^2)$$

$$- 3\lambda_{111}\sigma(\alpha_1\alpha_2\gamma_1\gamma_2 + \alpha_2\alpha_3\gamma_2\gamma_3 + \alpha_3\alpha_1\gamma_3\gamma_1)$$

przy czym: α_1 , α_2 , α_3 są kosinusami kątów określających kierunek magnetyzacji M_s w stosunku do osi krystalograficznych, γ_1 , γ_2 , γ_3 – kosinusami kątów określających kierunek naprężeń σ w stosunku do osi krystalograficznych, σ – naprężeniami wywołanymi działaniem sił zewnętrznych λ_{100} , λ_{111} – stałymi magnetostrykcji, odpowiednio w kierunku [100] i [111].

Pierwsze dwa człony w wyrażeniu (49) reprezentują energię anizotropii magneto-krystalicznej, natomiast dwa następne są związane z odkształceniem magnetostrycyjnym i naprężeniami σ , stanowiąc energię magnetosprężystą E_σ [9]. Położenie równowagi wektora magnetyzacji M_s jest określone przez minimum energii swobodnej E i stanowi funkcję stałych anizotropii K_1 , K_2 , K_3 ..., stałych magnetostrykcji λ_{100} i λ_{111} oraz naprężeń σ dla każdego danego kierunku naprężeń, określonego przez γ_1 , γ_2 , γ_3 .

Można zauważyć, że położenie wektora M_s będzie determinowane bardziej przez anizotropię magnetokrystaliczną, jeżeli K_1, K_2, K_3, \dots są większe niż $\lambda_{100} \sigma$ i $\lambda_{111} \sigma$, natomiast gdy te relacje są odwrotne, wówczas naprężenia σ wymuszają kierunek wektora M_s .

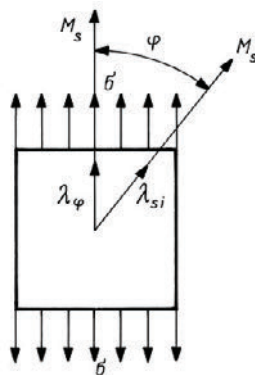
Jeżeli magnetostrykcja jest izotropowa, a więc $\lambda_{100} = \lambda_{111} = \lambda_{si}$, dwa ostatnie człony sprowadzają się do prostego wyrażenia na energię magnetosprężystą

$$E_\sigma = -\frac{3}{2} \lambda_{si} \sigma \cos^2 \varphi \quad (50)$$

w którym kąt φ jest kątem pomiędzy M_s a σ .

Rozważenie wyrażenia (50) oddaje jego sens fizyczny. Załóżmy, że naprężenia rozciągające σ są przyłożone do próbki (rys. 15) oraz że początkowo wektor magnetyzacji M_s jest równoległy do kierunku działania naprężeń.

Niech wektor M_s zostanie obrócony o kąt φ . Wówczas materiał skróci się wzdłuż kierunku działania naprężeń, ponieważ λ_φ jest mniejsze niż λ_{si} , jak pokazano na rysunku 15 (jeżeli λ_{si} jest dodatnie). To skrócenie w obecności naprężeń rozciągających oznacza, że praca została zmagazynowana w materiale. Praca ta jako energia magnetosprężysta materiału dla nieskończenie małego obrotu jest dana wyrażeniem $dE_\sigma = -\sigma d\lambda$.



Rys. 15. Obrót wektora magnetyzacji M_s w obecności naprężeń (schematycznie)
Fig. 15. Unit cube of material with isotropic magnetostriction

Tak więc

$$\int_0^{E_\sigma} dE_\sigma = - \int_{\lambda_{si}}^{\lambda_\varphi} d\lambda \quad (51)$$

$$E_\sigma = -\sigma \left[\frac{3}{2} \lambda_{si} \left(\cos^2 \varphi - \frac{1}{3} \right) - \lambda_{si} \right] = \frac{3}{2} \lambda_{si} \sigma \sin^2 \varphi \quad (52)$$

Ta postać wyrażenia na E_σ wskazuje, że dla dodatniego iloczynu $\lambda_{si}\sigma$ energia magnetosprężysta E_σ jest równa zero, jeżeli M_s i σ są równoległe, i rośnie do maksimum wynoszącego $3/2\lambda_{si}\sigma$, gdy M_s i σ tworzą kąt prosty

Dla ujemnego iloczynu $\sigma\lambda_{si}$ minimum energii E_σ występuje dla prostopadłości M_s i σ .

Zakładając, że anizotropia magnetokrystaliczna jest słaba, kierunek wektora M_s , przy nieobecności pola magnesującego H będzie głównie wymuszany przez naprężenia. W przypadku próbek polikrystalicznych (bez preferowanych orientacji) może być stosowana zależność (50). Z zależności tej wynika, że sposób, w jaki materiał reaguje na naprężenia, zależy jedynie od znaku iloczynu $\lambda_{si}\sigma$, wobec czego:

- Jeżeli $\lambda_{si}\sigma > 0$ (czyli dla materiału o dodatniej magnetostrykcji λ_{si} poddanego działaniu naprężeń rozciągających $+\sigma$ lub dla materiału o ujemnej magnetostrykcji $-\lambda_{si}$ poddanego naprężeniom ściskającym $-\sigma$), energia magnetosprężysta E_σ osiąga minimum, gdy kierunek działania naprężeń σ i kierunek wektora magnetyzacji M_s są równoległe lub antyrównoległe ($\varphi = 0^\circ$ lub 180°)
- Jeżeli $\lambda_{si}\sigma < 0$ (czyli dla materiału o dodatniej magnetostrykcji λ_{si} poddanego działaniu naprężeń ściskających $-\sigma$ lub dla materiału o ujemnej magnetostrykcji $-\lambda_{si}$ poddanego naprężeniom rozciągającym $+\sigma$), energia magnetosprężysta osiąga minimum, gdy kierunek wektora magnetyzacji M_s i kierunek działania naprężeń σ tworzy kąt prosty ($\varphi = 90^\circ$ lub 270°).

Wprowadzone do magnetyka, będącego w stanie rozmagnesowania, naprężenia mogą powodować ruch ścian domenowych. Ruch ten jednak musi być taki, ażeby zapewniał zerową magnetyzację próbki jako całości. Warunek ten nie jest trudny do spełnienia, ponieważ istnieje nieskończona liczba struktur, dla których wypadkowa magnetyzacja M jest równa zero. Naprężenia te mogą jednak kreować osie łatwego namagnesowania.

Tak więc w przypadku działania naprężeń anizotropia naprężeń musi być rozważana równocześnie z pozostałymi rodzajami anizotropii, jak np. anizotropia magnetokrystaliczna lub kształtu (formy). Można zatem zapisać, że energią anizotropii naprężeń K_μ^σ jest energia magnetosprężysta

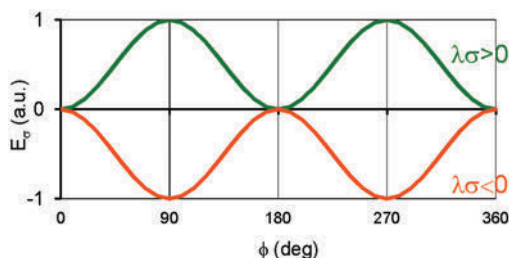
$$K_\mu^\sigma = E_\sigma = K_\sigma \sin^2 \varphi \quad (53)$$

przy czym stała anizotropii naprężeń K_σ , na podstawie zależności (52), dana jest przez $3/2\lambda_{si}\sigma$.

Na rysunku 16 przedstawiono unormowane zmiany energii magnetosprężystej E_σ materiału w funkcji kąta między kierunkiem wektora naprężeń a kierunkiem wektora magnetyzacji.

Przyjmując, że anizotropia magnetokrystaliczna jest stała dla danego materiału, natomiast anizotropia naprężeń rośnie liniowo ze wzrostem naprężeń, można wyzna-

czyć wartość naprężeń, w których oba rodzaje energii (E_K i E_σ) będą sobie równe. Naprężenia te mogą być z dużym przybliżeniem (wynikającym z faktu pominięcia stałych K_2 , K_3) wyznaczone przez podzielenie wyrażenia na wartość energii magneto-sprężystej $E_\sigma = (3/2)\lambda_{si}\sigma$ (dla $\cos^2\varphi = 1$) przez stałą K_1 energii anizotropii magnetokry-stalicznej.



Rys. 16. Unormowane zmiany energii magneto-sprężystej E_σ materiału w funkcji kąta między kierunkiem wektora naprężeń, a kierunkiem wektora magnetyzacji
Fig. 16. Normalized changes of magnetoelastic energy E_σ as a function of an angle between vectors of stresses and magnetization

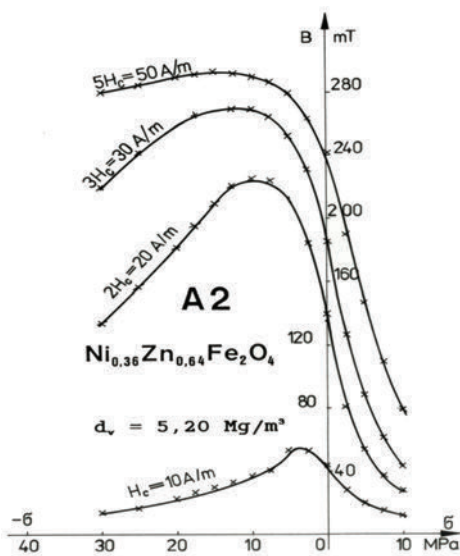
Przykładowo dla ferrytu gęstego o składzie chemicznym $Ni_{0,36}Zn_{0,64}Fe_2O_4$ ($K_1 = -2,7 \cdot 10^3 \text{ J/m}^3$, $\lambda_{si} = -2,5 \cdot 10^{-6}$) otrzyma się $\sigma = 2K_1/3\lambda_{si} = 720 \text{ MPa}$. Tak więc naprężenia, przy których $E_\sigma \cong E_K$, są znacznie większe od maksymalnej (ze względu na wytrzymałość ferrytu na ściskanie) wartości $\sigma = 100 \text{ MPa}$ w badanym zakresie zmian σ [98] i na tyle duże, że ich przyłożenie mogłoby doprowadzić do zniszczenia próbek. Oznacza to, że energia anizotropii magnetokryształicznej jest znacznie większa od anizotropii naprężeń i że naprężenia (w badanym zakresie zmian od 0 do 100 MPa) jedynie uprzywilejowują te kierunki łatwego magnesowania, które tworzą małe (w przypadku $\lambda_{si}\sigma > 0$) lub zbliżone do 90° (w przypadku $\lambda_{si}\sigma < 0$) kąty z kierunkiem działania naprężeń.

Na rysunku 17 przedstawiono doświadczalne charakterystyki $B(\pm\sigma)_H$ dla ferrytu gęstego o składzie chemicznym $Ni_{0,31}Zn_{0,69}Fe_2O_4$; magnetostrykcja λ_p tego ferrytu ma wartość ujemną.

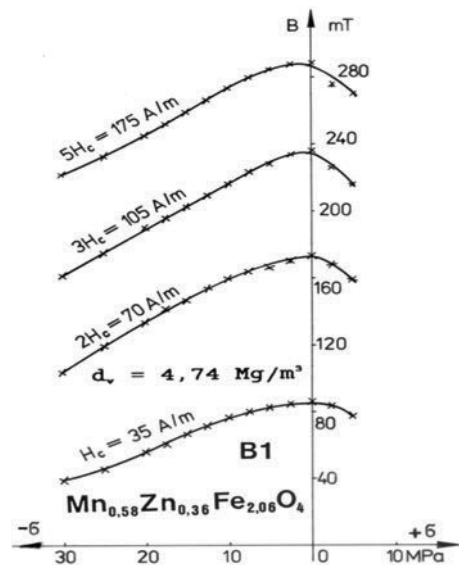
W początkowym zakresie zmian naprężeń ściskających $-\sigma$ indukcja magnetyczna B (dla ustalonej wartości pola H) pod wpływem działania naprężeń wzrasta i przy pewnej wartości naprężeń $\sigma = \sigma_v$ osiąga maksimum, czyli punkt Villariego. Po dalszym zwiększaniu wartości naprężeń (za punktem Villariego) indukcja B maleje.

Przed punktem Villariego ferryt reaguje zgodnie z mechanizmem opisanym dla przypadku $\lambda_{si}\sigma > 0$, jako że $\lambda_{si} < 0$ i $\sigma < 0$. Natomiast za punktem Villariego ten sam ferryt zaczyna reagować tak, jak w przypadku ujemnej wartości iloczynu $\lambda_{si}\sigma$ (indukcja B maleje ze wzrostem σ), pomimo że charakter naprężeń nie zmienił się (materiał jest nadal ściskany). Zachodzi proces odwrotny do tego, jaki był przy wzroście naprę-

żeń w zakresie od 0 do wartości σ_v . Ponieważ charakter naprężeń nie zmienił się, przyczyny zmiany procesu za punktem Villariego należy dopatrywać się w zmianie znaku magnetostrykcji. Oznacza to, że magnetostrykcja $\lambda_{||}$ w punkcie Villariego jest równa zero. Pewnym potwierdzeniem tego może być przypadek ferrytu $Mn_{0,58}Zn_{0,36}Fe_{2,06}O_4$ o zbliżonej do zera magnetostrykcji; efekt Villariego pozornie tu nie występuje, w rzeczywistości ma on miejsce dla $\sigma_v \approx 0$; dlatego zależność $B(\pm\sigma)_H$ dla $\sigma > 0$ jest monotoniczna. Przykład charakterystyki $B(\pm\sigma)_H$ dla tego ferrytu w naprężeniach zarówno ściskających, jak i rozciągających ilustruje rysunek 18. Natomiast na rysunku 19 podano analogiczne charakterystyki dla stopu amorficznego o składzie $Fe_{78}Si_{13}B_9$ w stanie wyjściowym i po wyżarzaniu.

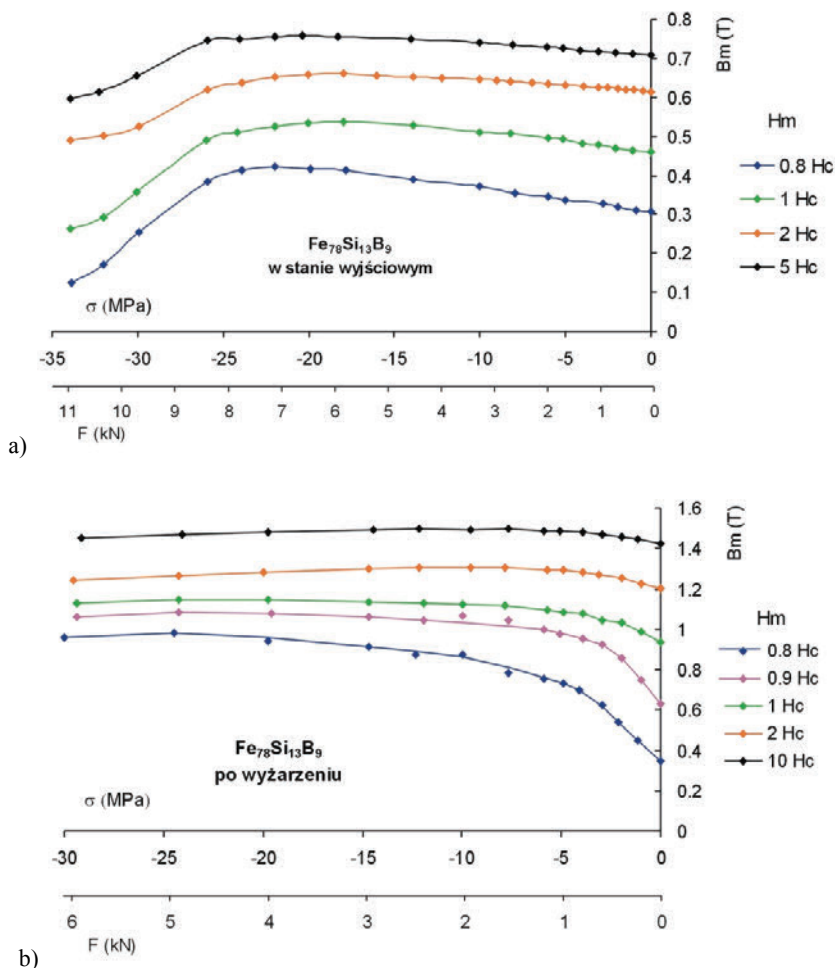


Rys. 17. Charakterystyki $B(\pm\sigma)_H$ ferrytu $Ni_{0,36}Zn_{0,64}Fe_2O_4$ [98]
Fig. 17. Magnetoelastic $B(\pm\sigma)_H$ characteristics of $Ni_{0,36}Zn_{0,64}Fe_2O_4$ ferrite [98]



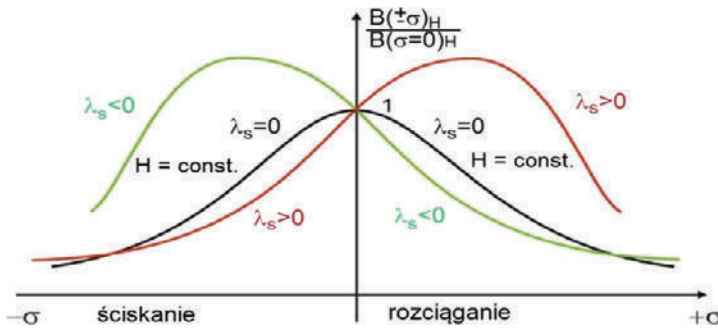
Rys. 18. Charakterystyki $B(\pm\sigma)_H$ ferrytu $Mn_{0,58}Zn_{0,36}Fe_{2,06}O_4$ [98]
Fig. 18. Magnetoelastic $B(\pm\sigma)_H$ characteristics of $Mn_{0,58}Zn_{0,36}Fe_{2,06}O_4$ ferrite [98]

Dogodnym w ujęciu technicznym sposobem oceny ilościowej efektu magnetoelastycznego są badania wpływu naprężeń (σ) na wartość indukcji magnetycznej B w materiale, dla ustalonej wartości zewnętrznego pola magnesującego H_m . W ten sposób można wyznaczyć czułość magnetoelastyczną materiału definiowaną jako względną zmianę indukcji magnetycznej (lub przenikalności magnetycznej materiału) pod wpływem jednostkowych naprężeń.



Rys. 19. Charakterystyki magnetoelastyczne $B(\sigma)_{H_m}$ rdzenia pierścieniowego ze stopu amorficznego o składzie $\text{Fe}_{78}\text{Si}_{13}\text{B}_9$: a) w stanie wyjściowym, b) po wyżarzaniu
 Fig. 19. Magnetoelastic $B(\sigma)_{H_m}$ characteristics of ring-shaped core made of $\text{Fe}_{78}\text{Si}_{13}\text{B}_9$ amorphous alloy: a) in as-quenched state, b) after annealing

Uogólnione ujęcie magnetoelastycznego zjawiska Villariego dla magnetyków magnetycznie miękkich podano w pracy [98]. Ujęcie to pod kątem fizycznym daje wspólny obraz zjawisk, zależności i kryteriów oceny właściwości magnetoelastycznych $B(\pm\sigma, H)$ i magnetostrykcyjnych $\lambda(\pm\sigma, H)$ magnetyka magnetycznie miękkiego. Zgodnie z nim zachowanie się magnetyka magnetycznie miękkiego w obecności naprężeń w całym zakresie ich zmian zarówno ściskających $-\sigma$, jak i rozciągających $+\sigma$ można opisać jedną z trzech (zależnie od znaku magnetostrykcji λ_s , tego magnetyka) charakterystycznych krzywych, przedstawionych na rysunku 20.



Rys. 20. Zależność unormowanej indukcji magnetycznej od naprężeń w materiałach magnetycznych o różnej wartości magnetostrykcji nasycenia [98]
 Fig. 20. Stress Dependence of normalized flux density in magnetic materials with different value of saturation magnetostriction [98]

Stromość tych krzywych, poza ekstremum, zależy od stosunku energii magneto-sprężystej do sumy pozostałych składników energii swobodnej. Z tego względu ekstrema związane z punktem Villariego nie dla wszystkich magnetyków magnetyczne miękkich muszą być wyraźnie zaznaczone i dla niektórych z nich (szczególnie w przypadku $H \gg H_c$) mogą być płaskie lub przejawiać się tylko jako stan nasycenia.

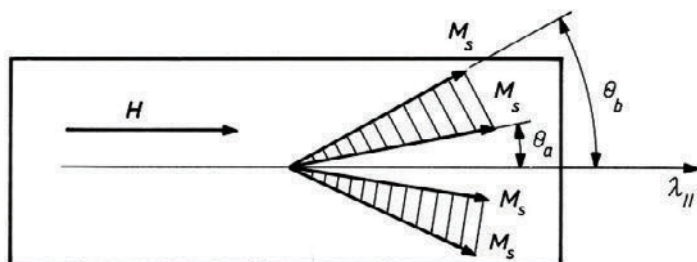
Źródłem efektu magnetosprężystego jest zmiana łącznej energii swobodnej w materiale pod wpływem naprężeń od sił zewnętrznych. Na łączną energię swobodną materiału magnetycznego składają się takie czynniki, jak:

- Energia anizotropii magnetokrystalicznej materiału związana ze strukturą sieci krystalicznej materiału.
- Energia od magnesowania spowodowana rozproszeniem pola magnetycznego na nieciągłościach materiału próbki.
- Energia wymiany związana ze strukturą atomową materiału.
- Energia magnetosprężysta związana ze zjawiskiem magnetosprężystym.

Na podstawie analizy łącznej energii swobodnej [115] można stwierdzić, że dla małych wartości naprężeń (σ) kierunek zmian indukcji zależy od znaku iloczynu $\lambda_s \sigma$. Gdy wartość tego iloczynu jest dodatnia, wtedy pod wpływem naprężeń wartość indukcji magnetycznej rośnie, natomiast gdy jest ujemna – maleje. Ponadto wielkość efektu magnetosprężystego jest na ogół tym większa, im większa jest wartość magnetostrykcji materiału (i związana z nią energia magnetosprężysta). Zwiększenie wartości energii anizotropii magnetokrystalicznej lub energii od magnesowania powoduje zmniejszenie wielkości efektu magnetosprężystego, ponieważ udział energii magnetosprężystej w łącznym bilansie energii swobodnej materiału maleje.

3.3. WPŁYW NAPRĘŻEŃ NA MAGNETOSTRYKCJĘ

Naprężenia wywołane działaniem sił zewnętrznych powodują nie tylko zmiany indukcji magnetycznej, ale mogą również, przez mechanostrykcję [116], powodować znaczne zmiany magnetostrykcji. Zjawisko to można wytłumaczyć istnieniem w stanie rozmagnesowania preferowanej orientacji domen jako następstwa działania naprężeń. W przypadku dodatniej wartości iloczynu ($\lambda_{si}\sigma$) naprężenia wywołane siłami zewnętrznymi będą preferowały strukturę domenową, w której kierunek działania naprężeń σ i kierunek wektora magnetyzacji M_s są równoległe lub antyrównoległe ($\phi = 0^\circ$ lub 180°). Przy takiej strukturze domenowej (charakterystycznej dla kryształów o anizotropii jednoosiowej) już niewielka wartość pola magnesującego, działającego zgodnie z kierunkiem naprężeń, może spowodować w wyniku przesunięć ścian 180° nasycenie próbki; nie powodując jednak odkształceń magnetostrykcyjnych λ . Oznacza to, że im większe jest równoległe i antyrównoległe uporządkowanie struktury domenowej w próbce wzdłuż kierunku działania naprężeń, tym mniejsze jest odkształcenie magnetostrykcyjne wzdłuż tego kierunku.



Rys. 21. Rozkład wektorów magnetyzacji w próbce (schematycznie) [98]

Fig. 21. Distribution of the M_s vectors in the sample [98]

Jeżeli stan początkowy jest stanem „idealnego rozmagnesowania”, to magnetostrykcję nasycenia λ_θ , mierzoną w kierunku tworzącym kąt θ z magnetyzacją M_s , opisuje znana już zależność

$$\lambda_\theta = \frac{3}{2} \lambda_{si} \left(\cos^2 \theta - \frac{1}{3} \right) \quad (54)$$

Jeżeli założy się, że stan początkowy jest stanem „idealnego rozmagnesowania”, natomiast stan końcowy nie odpowiada nasyceniu, a kierunki wektorów magnetyzacji domen zawarte są pomiędzy kątami θ_a i θ_b do osi próbki (rys. 21), to magnetostrykcja $\lambda_{||}$ wzdłuż osi próbki może być wtedy opisana zależnością

$$\lambda_{||} = \frac{3}{2} \lambda_{si} \left(\overline{\cos^2 \theta} - \frac{1}{3} \right) \quad (55)$$

przy czym $\overline{\cos^2 \theta}$ jest wartością średnią wszystkich orientacji wektora M_s w stanie końcowym (głównie orientacji pomiędzy kątami θ_a i θ_b).

Jeżeli stan magnetyczny próbki zmienia się od stanu 1 do stanu 2, z których każdy zdefiniowany jest przez średnią wartość odpowiednio $\cos^2 \theta_1$ i $\cos^2 \theta_2$, to magnetostrykcja $\lambda_{||}$ jako następstwo zmiany tych stanów jest równa

$$\lambda_{||} = \frac{3}{2} \lambda_{si} \left[\left(\overline{\cos^2 \theta_2} - \frac{1}{3} \right) - \left(\overline{\cos^2 \theta_1} - \frac{1}{3} \right) \right] = \frac{3}{2} \lambda_{si} (\overline{\cos^2 \theta_2} - \overline{\cos^2 \theta_1}) \quad (56)$$

Zakładając, że stan końcowy 2 odpowiada nasyceniu, to dla wszystkich domen $\theta_2 = 0$ i $\cos^2 \theta_2 = 1$, otrzymuje się

$$\lambda_{(||s)} = \frac{3}{2} \lambda_{si} (1 - \overline{\cos^2 \theta_1}) \quad (57)$$

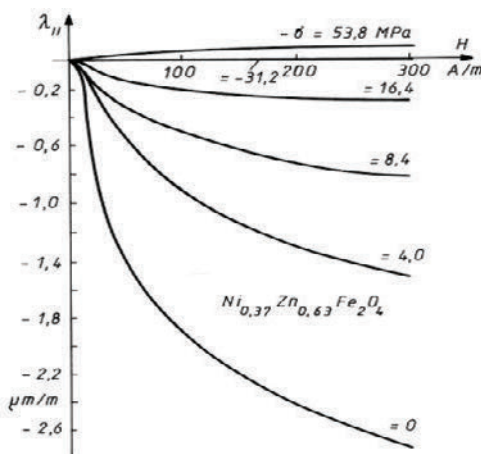
Zależność (57), podana przez Cullity'ego [9], umożliwia wyznaczenie magnetostrykcji w nasyceniu dla każdego stanu początkowego o znanym rozkładzie wektorów magnetyzacji.

Przykładowo, jeżeli w stanie „idealnego rozmagnesowania” polikryształu domeny zororientowane są przypadkowo, to średnia wartość $\cos^2 \theta_1 = 1/3$ i wówczas z zależności (57) $\lambda_{||s} = \lambda_{si}$. W rozmagnesowanej próbce magnetyka, poddanej działaniu naprężeń ściskających o wartości odpowiadającej punktowi Villario ($\sigma = \sigma_v$), wektory magnetyzacji są równoległe do osi działania naprężeń, a więc $\cos^2 \theta_1 = 1$. Zmierzona w tych warunkach magnetostrykcja $\lambda_{||s}$, zgodnie z (57), powinna być równa zero. Natomiast w tym samym magnetyku poddanym w stanie rozmagnesowania działaniu naprężeń rozciągających o odpowiednio dużej wartości – wektory magnetyzacji tworzą kąt proste do osi naprężeń, czyli $\cos^2 \theta_1 = 0$, a więc $\lambda_{||s} = 3/2 \lambda_{si}$.

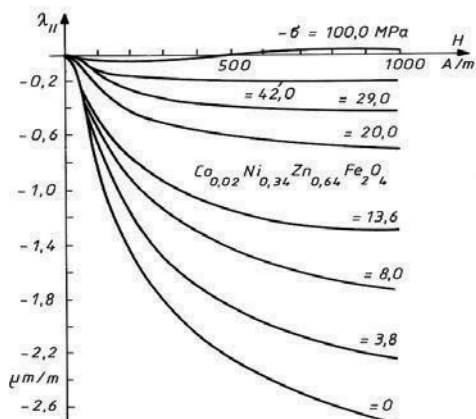
Wyniki badań doświadczalnych materiałów metalowych [55, 69, 117–122] potwierdzają opisywany charakter zmian magnetostrykcji w obecności naprężeń wywołanych siłami zewnętrznymi.

Tremolet [118], rozważając w przypadku żelaza wpływ naprężeń na strukturę domonową, rozróżnił dwa procesy: 1) przesunięcia ścian 90° , 2) obroty wektorów magnetyzacji. Mechanizm pierwszy występuje w zakresie naprężeń porównywalnych z naprężeniami własnymi materiału (w przypadku materiałów magnetycznie miękkich dla małych wartości naprężeń). Mechanizm drugi występuje w zakresie naprężeń, w których energia magnetosprężysta E_σ jest porównywalna z energią anizotropii magnetokrystalicznej E_K ; w przypadku żelaza występuje to dla $\sigma \cong 250$ MPa. Dla porównania na przykład w przypadku ferrytów, według dokonanych wcześniej obliczeń, wartość ta wynosi około 750 MPa. W odniesieniu do ferrytów należałoby więc w badanym zakresie naprężeń do 100 MPa uwzględniać mechanizm pierwszy. Tremolet, rozważając mechanizm pierwszy w żelazie, wykazał teoretycznie, a następnie potwierdził doświadczalnie, zmianę znaku magnetostrykcji $\lambda_{||}$ w zależności od naprężeń ściskających.

Na rysunkach 22 oraz 23 podano wyniki doświadczalne badań własnych [98] wpływu naprężeń na magnetostrykcję wybranych ferrytów.



Rys. 22. Charakterystyki $\lambda_{||}(H)_{-\sigma}$ ferrytu $\text{Ni}_{0,37}\text{Zn}_{0,63}\text{Fe}_2\text{O}_4$ [98]
 Fig. 22. Characteristics $\lambda_{||}(H)_{-\sigma}$ of the $\text{Ni}_{0,37}\text{Zn}_{0,63}\text{Fe}_2\text{O}_4$ ferrite [98]



Rys. 23. Charakterystyki $\lambda_{||}(H)_{-\sigma}$ ferrytu $\text{Co}_{0,02}\text{Ni}_{0,34}\text{Zn}_{0,64}\text{Fe}_2\text{O}_4$ [98]
 Fig. 23. Characteristics $\lambda_{||}(H)_{-\sigma}$ of the $\text{Co}_{0,02}\text{Ni}_{0,34}\text{Zn}_{0,64}\text{Fe}_2\text{O}_4$ ferrite [98]

Z zakresu wpływu naprężeń mechanicznych na magnetostrykcję magnetyków amorficznych ukazały się prace teoretyczne i doświadczalne Szymczaka i in. [123, 124], Lachowicza i in. [125, 126] i inne [127–132] potwierdzające, że przy pewnych wartościach naprężeń magnetostrykcja zmienia znak. Ukazały się również prace doświadczalne dotyczące wpływu naprężeń mechanicznych na magnetostrykcję stopów metali ziem rzadkich [133–135].

Przedstawione wyniki potwierdzają, że zjawisko magnetostrykcji było intensywnie badane ze względu na duże znaczenie zarówno w ujęciu analiz teoretycznych, jak i zastosowań praktycznych. Oba te aspekty wymagają jednak modelu fizycznego. Z aktualnie znanych modeli jedynie model Jilesa–Athertona–Sabluka umożliwia powiązanie zjawisk fizycznych z pętlami histerezy magnetycznej i magnetosprężystej.

4. MODEL JILESA–ATHERTONA–SABLIKA

Przedstawiony w poprzednich rozdziałach jakościowy model, opisujący procesy magnetomechaniczne, umożliwia zrozumienie istoty tych zjawisk. W powiązaniu z informacjami o magnetostrykcyjnych charakterystykach magnetyków umożliwia także przewidzenie tendencji wpływu naprężeń mechanicznych na charakterystyki magnesowania materiałów magnetycznie miękkich. Jednak technicznie do wykorzystania w procesie modelowania na potrzeby optymalizacji komponentów mechatronicznych dużo korzystniejsze są modele ilościowe. W literaturze opisane są cztery modele, które umożliwiają powiązanie wybranych właściwości fizycznych magnetyka z jego charakterystykami użytkowymi [136]:

- model Stonera–Wohlfartha [137],
- model Globusa [138],
- model Preisacha [139],
- model Jilesa–Athertona–Sabluka [140].

Model Stonera–Wohlfartha jest z powodzeniem stosowany do modelowania charakterystyk materiałów magnetycznie twardych. Ze względu na swoje założenia fizyczne w zasadzie nie powinien być stosowany do modelowania materiałów magnetycznie miękkich.

Model Globusa został opracowany specjalnie na potrzeby modelowania rdzeni pierścieniowych z izotropowych materiałów magnetycznie miękkich, takich jak ferryty. Jednak nie można go stosować w ilościowym opisie charakterystyk magnetostrykcyjnych czy magnetosprężystych.

Aktualnie intensywnie rozwijany model Preisacha jest szczególnie użyteczny w przypadku cienkich warstw magnetycznych, szczególnie tych wykorzystywanych w urządzeniach do cyfrowego zapisu informacji. Jednak ze względu na swój charakter ma on ograniczone zastosowanie w przypadku potrzeby modelowania zjawisk magnetostrykcyjnych czy magnetosprężystych w litych materiałach magnetycznie miękkich.

Model Jilesa–Athertona–Sabluka jest najbardziej użytecznym modelem procesu magnesowania ze względu na jego wyraźne powiązanie z opisem procesów fizycznych, w szczególności minimalizacją energii swobodnej materiału magnetycznie miękkiego i analizę potencjałów termodynamicznych. Ponadto model ten może być stosowany zarówno w odniesieniu do magnetyków izotropowych, jak i anizotropowych, co wyróżnia go na tle innych modeli.

4.1. MODELOWANIE CHARAKTERYSTYK MAGNETYCZNYCH

Założenia modelu Jilesa–Athertona–Sabluka są oparte na analizie potencjałów termodynamicznych materiału magnetycznie miękkiego [141]. W ujęciu fizycznym potencjały te opisują przemiany termodynamiczne i dane są zależnościami:

$$A = G + \mu_0 HM \quad (58)$$

$$G = U - TS + \frac{3}{2} \sigma \lambda \quad (59)$$

$$U = \frac{1}{2} \alpha \mu_0 M^2 \quad (60)$$

w których przez A oznaczono energię swobodną Helmholtza, czyli część energii wewnętrznej materiału, która może być w danym procesie termodynamicznym uwolniona na zewnątrz układu. Energia swobodna Gibbsa (zwana także entalpią swobodną) G jest maksymalną pracą nieobjętościową, którą może wykonać materiał podczas przemiany izotermicznej. Natomiast energia wewnętrzna materiału U to suma energii oddziaływań międzycząsteczkowych i wewnątrzcząsteczkowych układu, natomiast S to entropia materiału, opisującego część energii swobodnej, która nie może być oddana na zewnątrz jako praca. Ponadto w równaniach potencjałów termodynamicznych przez M oznaczona jest magnetyzacja materiału, H to zewnętrzne pole magnesujące, $\mu_0 = 4\pi 10^{-7}$ H/m to przenikalność magnetyczna próżni (stała magnetyczna), T – temperatura materiału (w kelwinach), σ – naprężenia w materiale, λ zaś – wartość odkształcenia magnetostrykcyjnego. Współczynnik α opisuje wartość sprzężenia międziodomenowego zgodnie z modelem Blocha [140].

Wartość efektywnego pola magnesującego H_e oddziałującego na poszczególne domeny w materiale ferromagnetycznym, które ma kluczowe znaczenie z perspektywy budowy modelu Jilesa–Athertona–Sabluka, można wyznaczyć bezpośrednio z równań potencjałów termodynamicznych [142]

$$H_e = \frac{1}{\mu_0} \left(\frac{\partial A}{\partial M} \right)_T = H + \alpha M + \frac{3}{2} \sigma \left(\frac{\partial \lambda}{\partial M} \right)_T \quad (61)$$

natomiast wartość magnetyzacji M w materiale dana jest zależnością [143]:

$$M = M_s \cos \theta \quad (62)$$

przy czym kąt θ jest kątem między kierunkiem magnetyzacji M w materiale a kierunkiem magnetyzacji spontanicznej M_s domeny (magnetyzacji nasycenia); wartość magnetostrykcji λ w materiale magnetycznym opisana jest zależnością [144]

$$\lambda = \frac{3}{2} \lambda_s \cos^2 \theta \quad (63)$$

w której λ_s jest współczynnikiem magnetostrykcji nasycenia materiału magnetycznego. Po przekształceniu równań (62) i (63) zależność między magnetostrycją λ a magnetyzacją M opisuje funkcja $\lambda(M)$:

$$\lambda(M) = \frac{3}{2} \lambda_s \frac{M^2}{M_s^2} \quad (64)$$

W rezultacie wartość efektywnego pola magnesującego H_e można wyznaczyć z zależności (61) i (64), otrzymując

$$H_e = H + \alpha M + \frac{9\lambda_s\sigma}{2M_s^2} M = H + M \left(\alpha + \frac{9\lambda_s\sigma}{2M_s^2} \right) \quad (65)$$

Na podstawie zależności (65) można więc ogólnie wprowadzić współczynnik $\tilde{\alpha}$, zdefiniowany jako [145]

$$\tilde{\alpha} = \alpha + \frac{9\lambda_s\sigma}{2M_s^2} \quad (66)$$

Po wprowadzeniu współczynnika $\tilde{\alpha}$ do zależności (65) pole efektywne H_e jest opisane zależnością

$$H_e = H + \tilde{\alpha} M \quad (67)$$

Należy podkreślić, że jeśli materiał nie jest poddany działaniu naprężeń od sił zewnętrznych ($\sigma = 0$), to zależność (66) upraszcza się do postaci $\tilde{\alpha} = \alpha$.

W przypadku materiałów paramagnetycznych proces magnesowania (który ma charakter bezhisterezowy) przebiega zgodnie z rozkładem Boltzmana i jest opisany zależnością [146]:

$$M_{\text{para}} = M_s \frac{\int_0^\pi e^{\frac{-E_m(\theta)}{k_B T}} \sin \theta \cos \theta d\theta}{\int_0^\pi e^{\frac{-E_m(\theta)}{k_B T}} \sin \theta d\theta} \quad (68)$$

w której k_B jest stałą Boltzmana, T – temperaturą, kąt θ zaś – kątem pomiędzy kierunkiem pola magnesującego a kierunkiem momentu magnetycznego w atomie. Energia $E_m(\theta)$ jest energią atomowego momentu magnetycznego ustawionego pod kątem θ do pola magnesującego H , która jest dana zależnością [146]

$$E_m(\theta) = -\mu_0 m_{at} H \cos \theta \quad (69)$$

W przypadku izotropowego ferromagnetyka o strukturze domenowej magnetyzacja bezhisterezowa wyznaczana jest analogicznie do sposobu zależności (68) [140]:

$$M_{\text{iso}} = M_s \frac{\int_0^\pi e^{\frac{-E_d(\theta)}{k_B T}} \sin \theta \cos \theta d\theta}{\int_0^\pi e^{\frac{-E_d(\theta)}{k_B T}} \sin \theta d\theta} \quad (70)$$

przy czym $E_d(\theta)$ jest energią domeny, której wektor magnetyzacji m_d usytuowany jest pod kątem θ do kierunku efektywnego pola magnesującego H_e , opisanego zależnością (65). Wartość energii $E_d(\theta)$ może być zatem wyznaczona z zależności [147]

$$E_d(\theta) = -\mu_0 m_d H_e \cos \theta \quad (71)$$

analogicznej do zależności (69) [148]. Średnia magnetyzacja m_d pojedynczej domeny może być opisana zależnością

$$m_d = \frac{M_s}{N} \quad (72)$$

Przy czym N jest średnią liczbą domen w jednostce objętości materiału ferromagnetycznego. Dlatego też można uznać, że parametr m_d pośrednio określa gęstość ścian domenowych w magnetyku. Przyjmując, że [140]

$$a = \frac{N k_B T}{\mu_0 M_s} \quad (73)$$

można wyznaczyć funkcje pierwotne funkcji całkowanych w zależności (70), co prowadzi do zależności określającej wartość magnetyzacji bezhisterezowej M_{iso} w magnetyku izotropowym [148, 140]

$$M_{\text{iso}} = M_s \left[\coth h \left(\frac{H_e}{a} \right) - \left(\frac{a}{H_e} \right) \right] \quad (74)$$

Należy podkreślić, że magnetyzacja bezhisterezowa może być wyznaczana z zależności (74) wyłącznie w przypadku magnetyków izotropowych, czyli materiałów, w których właściwości magnetyczne nie zależą od kierunku ich magnesowania.

W przypadku, gdy materiał magnetyczny wykazuje właściwości anizotropowe, jego magnetyzacja bezhisterezowa M_{iso} wyznaczana jest w sposób analogiczny do zależności (70). W publikacji magnetyzacja bezhisterezowa M_{aniso} opisana jest zależnością [149]

$$M_{\text{aniso}} = M_s \left[\frac{\int_0^\pi e^{E(1)+E(2)} \sin \theta \cos \theta d\theta}{\int_0^\pi e^{E(1)+E(2)} \sin \theta d\theta} \right] \quad (75)$$

energii $E(1)$ i $E(2)$ zaś opisują zależności [150]

$$E(1) = \frac{H_e}{a} \cos \theta - \frac{K_{\text{an}}}{M_s \mu_0 a} \sin^2(\psi - \theta) \quad (76)$$

$$E(2) = \frac{H_e}{a} \cos \theta - \frac{K_{\text{an}}}{M_s \mu_0 a} \sin^2(\psi + \theta) \quad (77)$$

przy czym K_{an} jest anizotropią magnetyka, kąt ψ jest kątem pomiędzy osią łatwego magnesowania materiału magnetycznego a kierunkiem pola magnesującego H_e w magnetyku. Niestety w odniesieniu do funkcji całkowanych w zależności (75) nie są znane funkcje pierwotne, dlatego równanie to może być rozwiązywane jedynie metodami numerycznymi. W praktyce, ponieważ całkowana funkcja ma charakter okresowy, do jej rozwiązania stosuje się całkowanie numeryczne z zastosowaniem wykorzystania aproksymacji Gussa–Konroda [151].

Dokładna analiza równania (75) podanego w [149] prowadzi do wniosku, że do równania tego zakradł się błąd obliczeniowy. Prawidłowa jego postać to

$$M_{\text{aniso}} = M_s \left[\frac{\int_0^\pi e^{\frac{E(1)+E(2)}{2}} \sin \theta \cos \theta d\theta}{\int_0^\pi e^{\frac{E(1)+E(2)}{2}} \sin \theta d\theta} \right] \quad (78)$$

Tylko w takiej postaci równanie (78), dla gęstości średniej energii anizotropii K_{an} równej zeru, redukuje się do równania Langevina danego zależnością (74).

Ponieważ większość stosowanych w praktyce materiałów, w tym magnetyki o strukturze amorficznej, są materiałami zawierającymi zarówno fazę anizotropową, jak i izotropową; ich magnetyzacja bez uwzględnienia histerezy magnetycznej M_{an} [149] jest opisywana przez sumę ważoną magnetyzacji izotropowej M_{iso} oraz magnetyzacji anizotropowej M_{aniso} i wyznaczana zgodnie z zależnością [148] jako

$$M_{\text{an}} = tM_{\text{aniso}} + (1-t)M_{\text{iso}} \quad (79)$$

przy czym $t \in \langle 0,1 \rangle$ opisuje zawartość fazy anizotropowej w materiale.

W modelu Jilesa–Athertona–Sabluka magnetyzacja nieodwracalna M_{irr} opisuje składową magnetyzacji materiału związaną z ruchem ścian domenowych w magnetyku. Po

zniesieniu pola magnesującego H ściany domenowe nie wracają samoistnie do poprzedniego położenia ze względu na obecność punktów zaczepienia ścian domenowych, będących na przykład dyslokacjami czy granicami ziaren w materiałach polikrystalicznych. Magnetyzacja nieodwracalna M_{irr} dana jest równaniem różniczkowym [140]

$$\frac{dM_{\text{irr}}}{dH_e} = \delta_M \frac{M_{\text{an}} - M_{\text{irr}}}{k\delta} \quad (80)$$

przy czym parametr $\delta = +1$ dla $dH/dt \geq 0$, oraz $\delta = -1$ dla $dH/dt < 0$ [140]. Parametr δM przyjmuje wartość 0, gdy $M_{\text{an}} - M > 0$ i $\delta = -1$, jak również jest równy 0 w przypadku, gdy $M_{\text{an}} - M < 0$ i $\delta = +1$. W pozostałych przypadkach $\delta M = 1$. Zastosowanie parametru δM , który zeruje zmiany magnetyzacji nieodwracalnej, stwarza możliwość uniknięcia nieuzasadnionej w ujęciu fizycznym, gdy w quasi-statycznej pętli histerezy magnetycznej przenikalność rewersyjna jest mniejsza niż 0 [152–154].

Współczynnik k w zależności (81) opisuje średnią energię konieczną do przesunięcia ściany domenowej przez punkt zaczepienia i opisany jest zależnością [147]:

$$k = \frac{\langle p \rangle \langle \varepsilon_\pi \rangle}{2m_d \mu_0 (1 - c)} \quad (81)$$

przy czym $\langle p \rangle$ jest średnią gęstością punktów zaczepienia, $\langle \varepsilon_\pi \rangle$ jest średnią energią stuosiemdziesięciostopniowych ścian domenowych, c – współczynnikiem odwracalności procesu magnesowania [147], który dany jest zależnością [140]

$$c = \left(\frac{m\pi y^4}{4E} \right) c' \quad (82)$$

w której E jest średnią energią powierzchniową ścian domenowych, $2y$ – średnią odległością między punktami zaczepienia ścian domenowych c' zaś jest stałą. Liczbowo parametr c jest równy ilorazowi przenikalności początkowej materiału μ_i oraz przenikalności początkowej materiału z pominięciem pętli histerezy (przenikalności bezhisterezowej) μ_{ah} [140].

Magnetyzacja odwracalna M_{rev} związana jest przede wszystkim ze sprężystym odkształcaniem się ścian domenowych w trakcie procesu magnesowania i w modelu Jilesa–Athertona–Sabluka dana jest zależnością

$$M_{\text{rev}} = c(M_{\text{an}} - M_{\text{irr}}) \quad (83)$$

W rezultacie można wyznaczyć różniczkową zależność magnetyzacji M próbkii w funkcji natężenia pola magnesującego H

$$\frac{dM}{dH} = \frac{\delta_M}{(1+c)} \frac{(M_{\text{an}} - M)}{(\delta k - \hat{a}(M_{\text{an}} - M))} + \frac{c}{(1+c)} \frac{dM_{\text{an}}}{dH} \quad (84)$$

Niestety równanie to można rozwiązać jedynie metodami numerycznymi. W praktyce do rozwiązania tego równania stosuje się metodę Rungego–Kutty czwartego rzędu [155], przyjmując co najmniej 300 kroków obliczeniowych na pętłę histerezy magnetycznej. Alternatywnie można zastosować metodę Rungego–Kutty czwartego i piątego rzędu z adaptacyjnym doбором kroku obliczeń [156]. W modelu Jilesa–Athertona–Sablíka całkowita indukcja B w materiale magnetycznym związana jest z magnetyzacją odwracalną M_{rev} oraz magnetyzacją nieodwracalną M_{irr} zgodnie z zależnością [140]

$$B = \mu_0(M_{\text{rev}} + M_{\text{irr}}) \quad (85)$$

W ten sposób model Jilesa–Athertona–Sablíka opisuje proces magnesowania materiału magnetycznie miękkiego, łącząc parametry fizyczne magnetyka z jego właściwościami funkcjonalnymi.

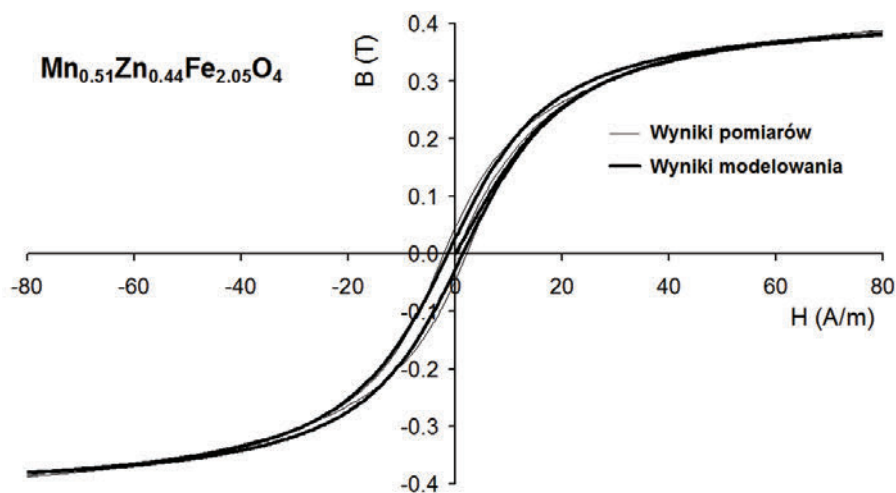
Podsumowując, zgodnie z oryginalnym modelem Jilesa–Athertona–Sablíka, właściwości materiału magnetycznie miękkiego w zakresie magnesowania przebiegiem wolnozmiennym (magnesowania quasistatycznego) są opisane przez następujące parametry:

- parametr a (wyrażany w A/m) opisujący, zgodnie z zależnością (73), gęstość domen w materiale magnetycznie miękkim,
- bezwymiarowy parametr $\tilde{\alpha}$, zgodnie z zależnością (66), opisujący sprzężenie międzydomenowe,
- parametr K_{an} (wyrażany w J/m^3), równy co do wartości energii anizotropii materiału magnetycznie miękkiego,
- kąt ψ , będący kątem pomiędzy kierunkiem pola magnesującego H a kierunkiem osi łatwego magnesowania w magnetyku wynikającej z jego anizotropii,
- bezwymiarowy parametr t opisujący udział fazy anizotropowej w magnetyku,
- parametr k (wyrażany w A/m) opisujący, zgodnie z zależnością (81), średnią energię potrzebną do przesunięcia ściany domenowej przez punkt zaczepienia,
- bezwymiarowy parametr c opisujący odwracalność procesu magnesowania (zgodnie z zależnością (82)),
- bezwymiarowy parametr λ_s równy jest współczynnikowi magnetostrykcji nasycenia.

W literaturze nie podano także do tej pory spójnego sposobu wyznaczania parametrów modelu Jilesa–Athertona–Sablíka. W pracy [157] przedstawiono jedynie wytyczne do wyznaczenia parametrów modelu metodą rekurencyjną na podstawie właściwości makroskopowych materiału, jednak wytyczne te dotyczą wyłącznie magnetyków izotropowych. Ponadto dokładność wyznaczenia parametrów modelu według podanych wytycznych umożliwi raczej zastosowanie ich jako punktu wyjściowego do wyznaczenia parametrów modelu w trakcie procesu optymalizacji niż do ich praktycznego zastosowania [158]. Podane w innych pracach metody wyznaczania

parametrów modelu także dotyczą wyłącznie magnetyków izotropowych [152, 159, 160, 161] i opierają się na zastosowaniu procesu optymalizacji wieloparametrycznej. Nie jest znana jednoznacznie zdefiniowana metodyka wyznaczania parametrów modelu w odniesieniu zarówno do materiałów izotropowych, jak i anizotropowych, co stanowi poważną barierę w praktycznym zastosowaniu modelu Jilesa–Athertona–Sabluka do opisu procesów magnesowania materiałów magnetycznie miękkich stosowanych w praktyce, szczególnie do wyznaczania ich charakterystyk jako rdzeni elementów indukcyjnych.

Na rysunku 24 przedstawiono wyniki modelowania pętli histerezy magnetycznej $B(H)$ ferrytu manganowo-cynkowego $Mn_{0,51}Zn_{0,44}Fe_{2,05}O_4$. Parametry modelu wyznaczono za pomocą optymalizacji z zastosowaniem strategii ewolucyjnej ($\mu + \lambda$) oraz optymalizacji gradientowej [162]. Można zaobserwować dużą zgodność wyników modelowania i wyników pomiarowych.



Rys. 24. Pętla histerezy magnetycznej $B(H)$ ferrytu manganowo-cynkowego $Mn_{0,51}Zn_{0,44}Fe_{2,05}O_4$:
wyniki pomiaru oraz obliczeń [176]

Fig. 24. Magnetic hysteresis loop $B(H)$ of $Mn_{0,51}Zn_{0,44}Fe_{2,05}O_4$ ferrite:
results of measurements and modelling [176]

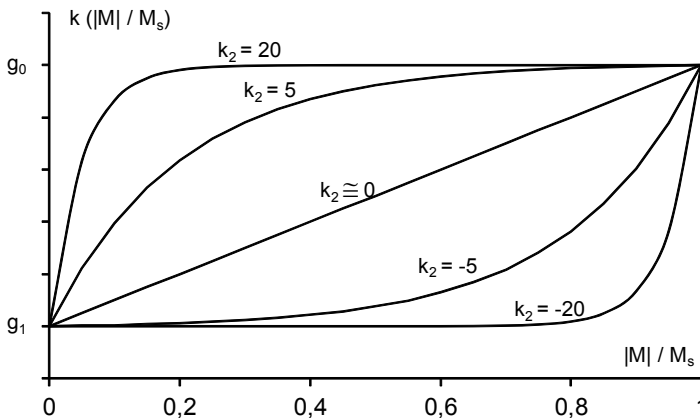
W modelu Jilesa–Athertona–Sabluka przyjęte jest założenie, że parametr k opisujący średnią energię konieczną do pokonania ściany domenowej zaczepu jest stały w całym zakresie pól magnesujących [147], co nie jest prawdziwe w ujęciu fizycznym [163, 164]. W trakcie procesu magnesowania średnia energia potrzebna do pokonania zaczepu ściany domenowej zmienia się w sposób monotoniczny [165, 166], co powinno być uwzględnione w opisie procesu magnesowania będącym podstawą modelu Jilesa–Athertona–Sabluka. Jednak znane z literatury rozwinięcie modelu Jilesa–Athertona–Sabluka w odniesieniu do materiałów izotropowych, w którym wybrane

parametry modelu ulegają zmianie w funkcji pola magnesującego [160, 161], nie mają fizycznego uzasadnienia. Opis zmian parametru k powinien być powiązany ze stanem fizycznym materiału, gdyż opisuje on wartość magnetyzacji materiału M , a nie wartość pola magnesującego H .

Zaproponowany przez autora nowy opis przebiegu zmian parametru k w funkcji magnetyzacji materiału M został podany w postaci zależności (86). Oddaje ona charakter zmian parametru k wynikających ze zmiany średniej energii koniecznej do pokonania zaczepu ściany domenowej [167].

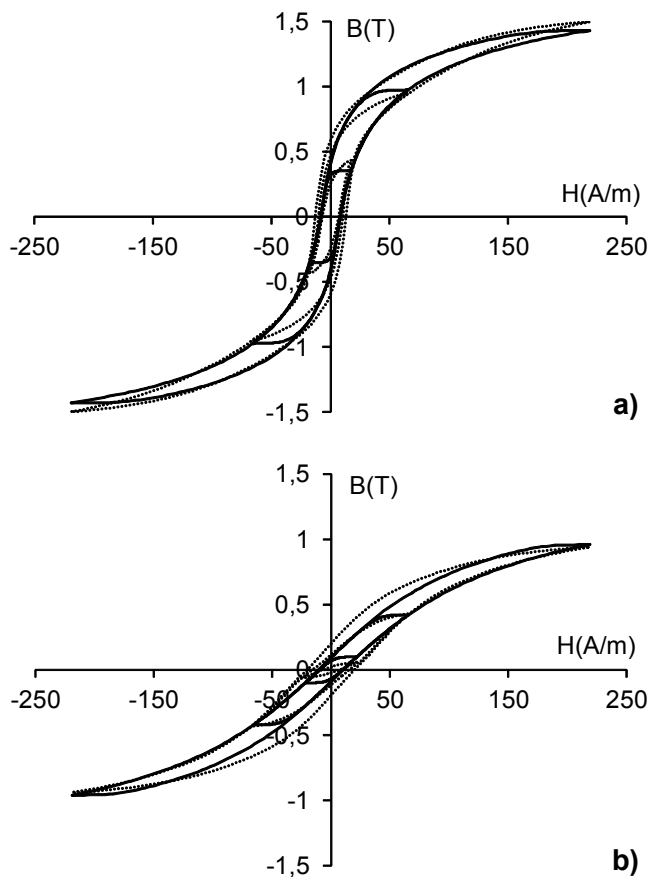
$$k(|M|/M_s) = k_0 + \frac{e^{k_2 (1-|M|/M_s)} - 1}{e^{k_2} - 1} (k_1 - k_0) \quad (86)$$

W zależności tej zmiany parametru $k(|M|/M_s)$ opisane są jako funkcja wektora trzech nowych parametrów: k_0 – określa wartość k w stanie remanencji, k_1 – określa wartość k w stanie nasycenia magnetycznego, natomiast k_2 – jest współczynnikiem kształtu przebiegu funkcji $k(|M|/M_s)$. Dla dodatnich wartości k_2 przebieg funkcji $k(|M|/M_s)$ jest wypukły do góry, dla ujemnych zaś – wypukły do dołu, jak podano na rysunku 25.



Rys. 25. Zależność funkcji $k(|M|/M_s)$ od parametrów k_0 , k_1 i k_2 [167]
 Fig. 25. Dependence of $k(|M|/M_s)$ on parameters k_0 , k_1 and k_2 [167]

Na rysunku 26 przedstawiono wyniki weryfikacji modelu rozszerzonego o zależność (86) w odniesieniu do silnie anizotropowej stali elektrycznej M130-27s. Weryfikację przeprowadzono w odniesieniu do wyników pomiarów eksperymentalnych zarówno w kierunku osi łatwego magnesowania, jak i w kierunku prostopadłym do tej osi. Także i w tym przypadku można zaobserwować dużą zgodność wyników modelowania i pomiarów.



Rys. 26. Pętla histerezy magnetycznej $B(H)$ anizotropowej stali elektrycznej M130-27s magnesowanej w kierunku: a) równoległym i b) prostopadłym do kierunku osi łatwego magnesowania.

Wyniki pomiarów – linia ciągła, wyniki modelowania – linia przerywana

Fig. 26. Magnetic hysteresis loop $B(H)$ of anisotropic electrical steel M130-27s magnetized in direction: a) parallel, b) perpendicular to the easy axis. Solid line – results of measurements, dotted line – results of modelling

4.2. MODELOWANIE CHARAKTERYSTYK MAGNETOSPREŻYSTYCH

W modelu Jilesa–Athertona–Sablíka przyjmuje się, że naprężenia mechaniczne σ wywoływane działaniem sił zewnętrznych wpływają jedynie na parametr $\tilde{\alpha}$, zgodnie z zależnością (66). Jest to jednak założenie znacznie uproszczone, ponieważ badania eksperymentalne wykazały, że naprężenia σ zmieniają także wartość magnetostrykcji nasycenia λ_s [98]. Ponadto pod wpływem naprężeń mechanicznych wartość magnetostrykcji nasycenia λ_s może nawet zmienić znak, co dodatkowo utrudnia opis ilościowych zmian współczynnika $\tilde{\alpha}$ w funkcji naprężeń. Dlatego wartość tego współczyn-

nika powinna być wyznaczana za pomocą analizy wyników pomiarów wpływu naprężeń na charakterystyki magnesowania materiałów magnetycznie miękkich.

Składowa gęstość energii anizotropii K_{an} , która jest generowana przez naprężenia σ , może być wyznaczona przez analizę energii magnetosprężystej E_σ . Po założeniu, że początkowo wektor magnetyzacji M jest równoległy do kierunku działania naprężeń σ , dla nieskończonego małego kąta ϕ obrotu wektora magnetyzacji, wykonana praca (jako energia magnetosprężysta E_σ) jest dana zależnością (87):

$$dE_\sigma = -\sigma d\lambda \quad (87)$$

w której przez $d\lambda$ oznaczono nieskończenie małą zmianę wartości magnetostrykcji w materiale. Wartość energii magnetosprężystej E_σ może być wyznaczona z zależności [168]

$$\int_0^{E_\sigma} dE_\sigma = -\sigma \int_{\lambda_s}^{\lambda_\phi} d\lambda \quad (88)$$

Zgodnie z analizą przedstawioną w [98] wartość magnetostrykcji λ_ϕ dana jest zależnością

$$\lambda_\phi = \frac{3}{2} \lambda_s \left(\cos^2 \phi - \frac{1}{3} \right) \quad (89)$$

przy czym ϕ jest kątem między osią łatwego magnesowania a kierunkiem magnetyzacji domeny. Wartość energii magnetosprężystej E_σ można wyznaczyć z zależności [98, 170]:

$$E_\sigma = -\sigma \left[\frac{3}{2} \lambda_s \left(\cos^2 \phi - \frac{1}{3} \right) - \lambda_s \right] = \frac{3}{2} \lambda_s \sigma \sin^2 \phi = K_\sigma \sin^2 \phi \quad (90)$$

co umożliwia wyznaczenie gęstości energii anizotropii K_σ generowanej naprężeniami mechanicznym σ jako

$$K_\sigma = \frac{3}{2} \lambda_s \sigma \quad (91)$$

Ponieważ wartość magnetostrykcji nasycenia λ_s zależy od naprężeń σ , zmiany K_σ (jak również całkowitą, średnią wartość energii anizotropii K_{an} , w funkcji naprężeń σ od sił zewnętrznych) należy wyznaczyć eksperymentalnie. Ponadto na skutek pojawienia się w bilansie energii swobodnej energii anizotropii magnetosprężystej K_σ , magnetyki izotropowe stają się anizotropowe. W takim przypadku ich magnetyzacja bezhisterezowa powinna być wyznaczana zgodnie z zależnością (78), a nie (74). Natomiast w przypadku magnetyków anizotropowych kierunek osi łatwego magnesowania Ψ może ulec zmianie pod wpływem naprężeń σ od sił zewnętrznych.

Działanie naprężeń generowanych przez siły zewnętrzne powoduje także zmiany udziału fazy anizotropowej t w magnetyku. Jednak liczbowo zmiany te są związane ze strukturą materiału i dlatego zmiany wartości parametru t w funkcji naprężeń powinny być wyznaczane na podstawie eksperymentalnych charakterystyk magnesowania materiału poddanego działaniu naprężeń.

Badania struktury domenowej wykorzystujące efekt Kerna wykazały, że naprężenia od sił zewnętrznych wpływają zarówno na kształt, jak i na liczbę domen w jednostce objętości magnetyka [171, 172]. Jednak do tej pory nie przedstawiono w literaturze liczbowego opisu zależności gęstości domen od naprężeń. Z tego względu wartość zmian współczynnika a w funkcji naprężeń od sił zewnętrznych musi być wyznaczana przez analizę wyników eksperymentalnych pomiarów pętli histerezy magnetycznej.

Gęstość powierzchniowa γ energii ściany domenowej w materiale magnetycznie miękkim jest dana jako suma niezależnej od naprężeń mechanicznych energii wymiany E_{ex} oraz energii anizotropii E_a w magnetyku [173]:

$$E_a = K_{\text{an}} \delta \quad (92)$$

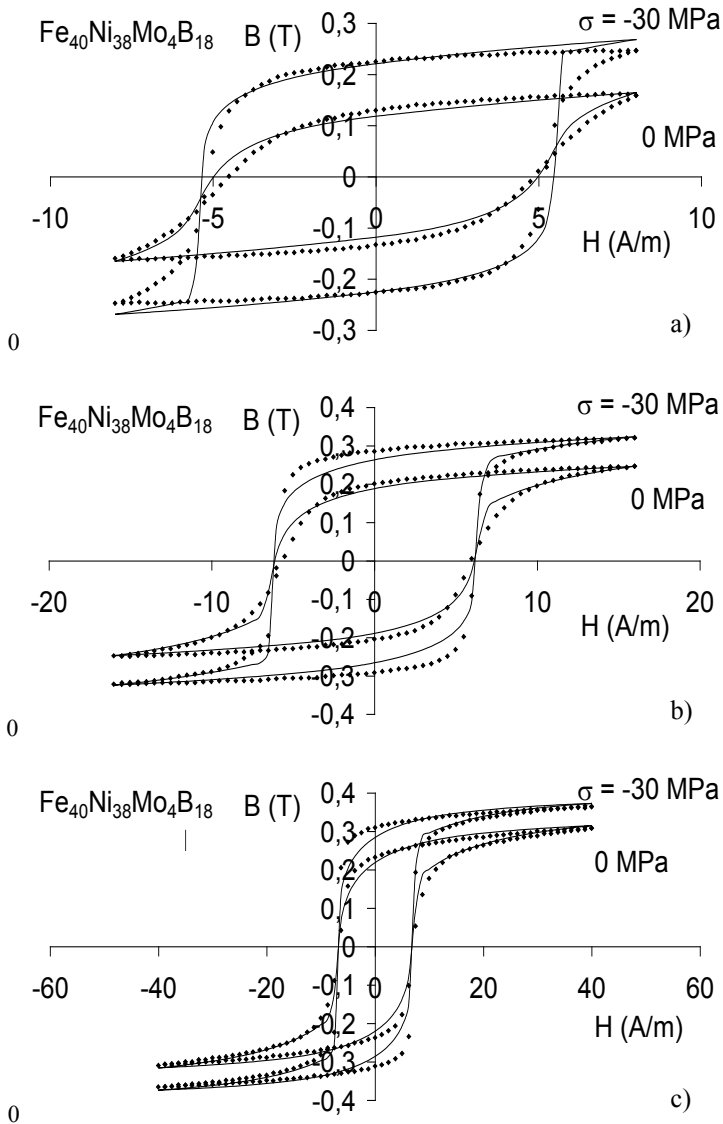
gdzie δ jest grubością ściany domenowej. Dlatego współczynnik c , analogicznie do parametru modelu K_{an} , musi być wyznaczany na podstawie danych pomiarowych.

Zgodnie z zależnością (86) wartość parametru k w modelu jest określona zarówno przez średnią gęstość punktów zaczepienia $\langle p \rangle$ w magnetyku, jak również przez średnią energię stuosiemdziesięciostopniowych ścian domenowych $\langle \varepsilon \pi \rangle$ oraz współczynnik c odwracalności procesu magnesowania. Dlatego także przebieg zmian wartości współczynników k_0 , k_1 i k_2 w funkcji naprężeń powinien być wyznaczony według analizy danych eksperymentalnych.

Z analizy energii swobodnej materiału magnetycznie miękkiego wynika, że w stanie nasycenia w materiale dominuje energia magnetyzacji E_m [170, 173], energia zaś magnetosprężysta ma znikomy udział w całkowitym bilansie energii swobodnej magnetyka. Z tego względu wartość magnetyzacji nasycenia M_s ma stałą wartość, niezależnie od zadawanych do magnetyka naprężeń mechanicznych.

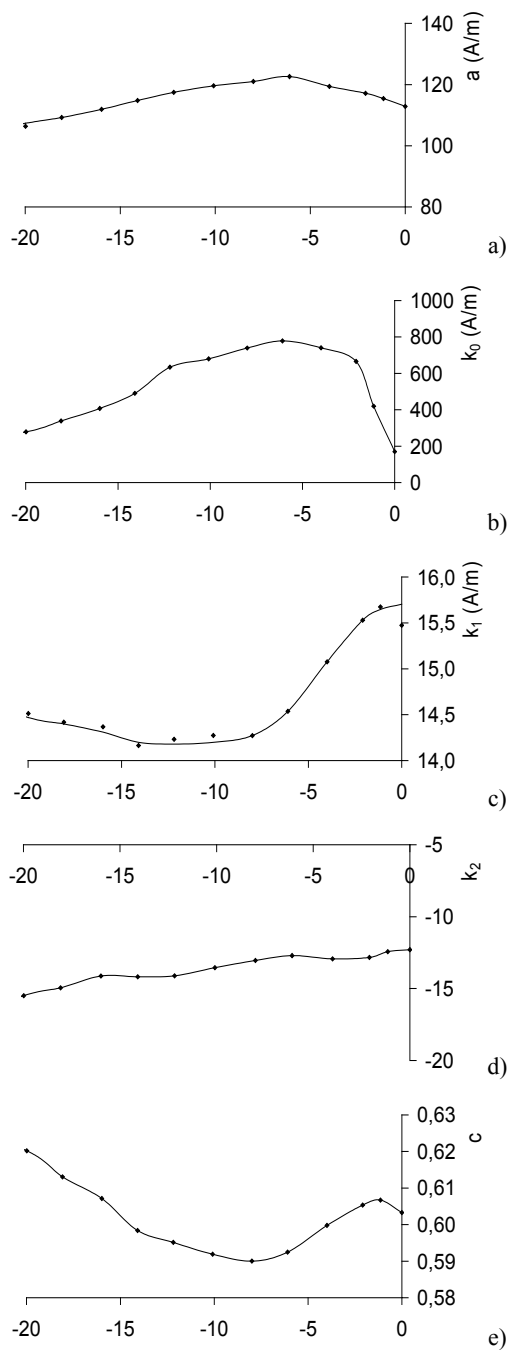
Na rysunkach 27 a, b i c przedstawiono wyniki pomiaru i modelowania wpływu naprężeń ściskających σ zadawanych w kierunku prostopadłym do postawy na pętlę histerezy magnetycznej $B(H)$ rdzenia pierścieniowego wykonanego ze stopu amorficznego o składzie $\text{Fe}_{40}\text{Ni}_{38}\text{Mo}_4\text{B}_{18}$. W procesie produkcji rdzeń wyżarzono w temperaturze 380 °C przez 1 godzinę.

W modelowaniu dla pętli histerezy magnetycznej $B(H)$ uzyskanej dla określonej wartości naprężeń ściskających σ wyznaczono parametry modelu Jilesa–Athertona–Sabluka, które ulegają zmianie pod wpływem naprężeń mechanicznych. Zależności tych parametrów od naprężeń podano na rysunku 28.



Rys. 27. Wpływ naprężeń ściskających σ zadawanych w kierunku prostopadłym do postawy na pętlę histerezy magnetycznej $B(H)$ rdzenia pierścieniowego wykonanego ze stopu amorficznego o składzie $\text{Fe}_{40}\text{Ni}_{38}\text{Mo}_4\text{B}_{18}$ dla zróżnicowanej amplitudy natężenia pola magnesującego: a) 8 A/m, b) 16 A/m, c) 40 A/m. Kropkami (♦) oznaczono wyniki pomiarów, linią ciągłą zaś (-) wyniki modelowania z zastosowaniem modelu Jilesa–Athertona–Sabluka [175]

Fig. 27. Influence of compressive stresses σ on $B(H)$ hysteresis loop of $\text{Fe}_{40}\text{Ni}_{38}\text{Mo}_4\text{B}_{18}$ amorphous alloy for values of the amplitude of magnetizing field: a) 8 A/m, b) 16 A/m, c) 40 A/m. Results of measurements marked as dots (♦) and results of modelling are marked by the solid lines (-) [175]



Rys. 28. Zależność parametrów modelu Jilesa–Athertona–Sablíka od naprężeń ściskających σ [175]

Fig. 28. Compressive stress dependence of Jiles–Atherton–Sablík's model parameters σ [175]

Należy podkreślić, że wartość współczynnika determinacji R^2 , wyznaczonego dla wyników pomiarów i modelowania pętli histerezy magnetycznej $B(H)$, przekracza 0,99 dla wszystkich wartości naprężeń ściskających σ w przedziale 0–10 MPa. Pomimo znacznych wartości względnych odchyżeń maksymalnych, występujących w specyficznych punktach pętli histerezy $B(H)$, wyniki analizy statystycznej potwierdziły dużą zgodność modelu z wynikami pomiarowymi. Zapewniają więc one możliwość zastosowania przedstawionego modelu do modelowania funkcjonalnych właściwości czujników magnetośprężystych stosowanych w warunkach przemysłowych. W odniesieniu do tych czujników wartość mierzonej siły może być, po uwzględnieniu parametrów geometrycznych rdzenia, w prosty sposób przeliczona na wartość naprężeń σ wytwarzanych w rdzeniu.

4.3. MODELOWANIE CHARAKTERYSTYK MAGNETOSTRYKCYJNYCH

Uproszczony model magnetostrykcji wiąże magnetyzację nasycenia M_s materiału, jego magnetostrykcję nasycenia λ_s i magnetyzację chwilową M zależnością kwadratową [144] daną równaniem (64).

Pomimo swojej prostoty model ten umożliwia uzyskanie znacznej zgodności między wynikami modelowania a pomiarem pętli magnetostrykcyjnych $\lambda(H)$ i $\lambda(B)$. Wyniki takiego modelowania w odniesieniu do ferrytu manganowo-cynkowego $Mn_{0,51}Zn_{0,44}Fe_{2,05}O_4$ przedstawiono na rysunku 29.

Jednak model dany równaniem kwadratowym (64) nie opisuje dwóch ważnych w ujęciu fizycznym zjawisk [177]: histerezy magnetomechanicznej w zależności $\lambda(B)$ oraz zjawiska oderwania się pętli histerezy magnetycznej $\lambda(B)$ od osi odciętych (ang. *lift-off phenomenon*). Aby uwzględnić te zjawiska, należy w modelu pętli histerezy magnetostrykcyjnej uwzględnić gęstość energii magnetycznej ϕ_{mag} [177] danej zależnością

$$\phi_{\text{mag}}(M) = \frac{1}{2} \mu_0 (\alpha M^2 + \alpha'' (M_{ah} - M)^2 - 2\alpha' (M_{ah} - M)H) \quad (93)$$

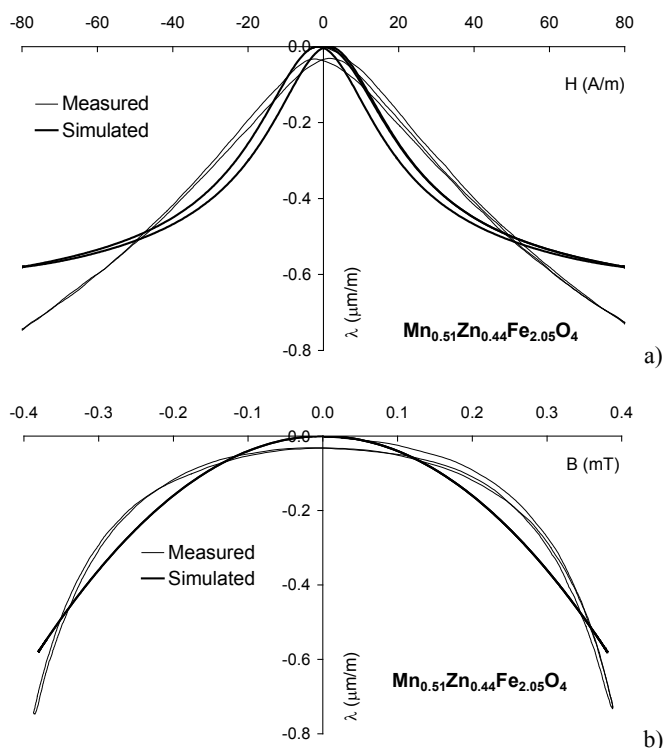
gdzie α' opisuje histerezę w zależności $\lambda(B)$, α'' zaś jest związane ze zjawiskiem oderwania się pętli histerezy magnetostrykcyjnej od osi rzędnych. Ostatecznie pętla histerezy magnetostrykcyjnej opisana jest zależnościami (94)–(96):

$$\lambda(M) = t_1 \left\{ \sqrt{1 + t_2 \phi_{\text{mag}}(M_s)} - (1 + t_2 \sqrt{\phi_{\text{mag}}(M_s) - \phi_{\text{mag}}(M)}) \right\} \quad (94)$$

$$t_1 = -\frac{4b(1+\nu)}{9Y} \quad (95)$$

$$t_2 = \frac{9Y}{2b^2(1+\nu)^2} \quad (96)$$

w których Y to moduł Younga, ν – stała Poissona, natomiast b to stała opisująca izotropowe sprzężenie magneto mechaniczne [177].



Rys. 29. Wyniki modelowania charakterystyk magnetostrykcyjnych ferrytu manganowo-cynkowego $\text{Mn}_{0.51}\text{Zn}_{0.44}\text{Fe}_{2.05}\text{O}_4$ z zastosowaniem zależności kwadratowej: a) pętla $\lambda(H)$, b) pętla $\lambda(B)$ [176]

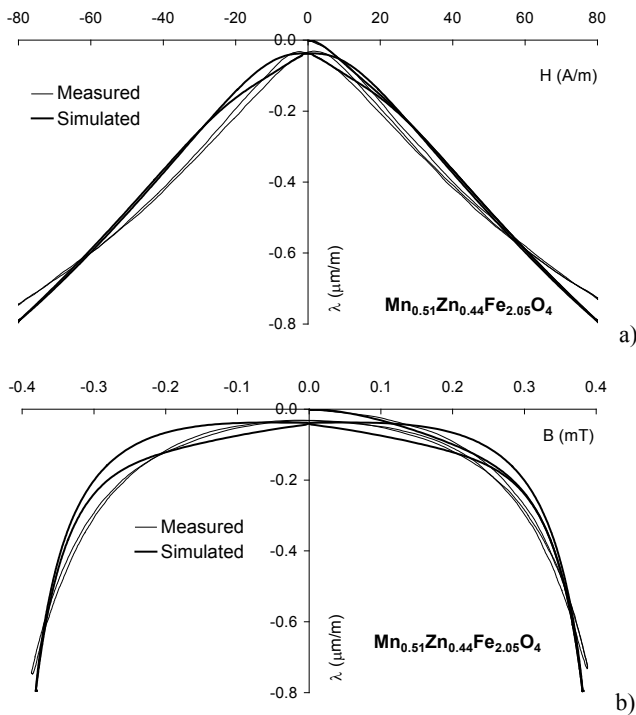
Fig. 29. Results of modelling the magnetostrictive characteristics of $\text{Mn}_{0.51}\text{Zn}_{0.44}\text{Fe}_{2.05}\text{O}_4$ by quadratic polynomial model: a) $\lambda(H)$ loop, b) $\lambda(B)$ loop [176]

W tabeli 1 zebrano parametry pętli histerezy magnetycznej i magnetostrykcyjnej wyznaczone dla ferrytu manganowo-cynkowego $\text{Mn}_{0.51}\text{Zn}_{0.44}\text{Fe}_{2.05}\text{O}_4$. Wyniki modelowania pętli histerezy magnetostrykcyjnej $\lambda(H)$ i $\lambda(B)$ podano natomiast na rysunku 30.

Należy podkreślić, że modelowanie pętli histerezy magnetostrykcyjnej $\lambda(H)$ i $\lambda(B)$ dla ferrytu manganowo-cynkowego $\text{Mn}_{0.51}\text{Zn}_{0.44}\text{Fe}_{2.05}\text{O}_4$ prowadzi do wyników o mniejszej zgodności z wynikami eksperymentalnymi niż modelowanie samej pętli histerezy magnetycznej $B(H)$. Jednak i w przypadku pętli histerezy magnetostrykcyjnej $\lambda(H)$ i $\lambda(B)$ współczynnik determinacji R^2 , wyznaczony dla wyników pomiarów eksperymentalnych i wyników modelowania, przekracza 0.98. Dlatego przedstawiony model pętli histerezy magnetostrykcyjnej może być z powodzeniem stosowany w praktyce.

Tabela 1. Parametry modelu Jilesa–Athertona–Sabluka opisujące pętle histerezy magnetostrykcyjnej $\lambda(H)$ i $\lambda(B)$ dla ferrytu manganowo-cynkowego $\text{Mn}_{0,51}\text{Zn}_{0,44}\text{Fe}_{2,05}\text{O}_4$ [176]
 Table 1. Parameters of Jiles–Atherton model describing magnetostrictive hysteresis loops $\lambda(H)$ and $\lambda(B)$ for $\text{Mn}_{0,51}\text{Zn}_{0,44}\text{Fe}_{2,05}\text{O}_4$ ferrite [176]

Parametr	Wartość	Jednostka
A	9,42	A/m
K	7,45	A/m
C	0,83	–
M_s	$3,41 \cdot 10^5$	A/m
α	$2,08 \cdot 10^{-5}$	–
B	$4,16 \cdot 10^3$	N/m ²
Y	$1,49 \cdot 10^{11}$	N/m ²
ν	0,35	–
α'	$3,01 \cdot 10^{-5}$	–
α''	$2,88 \cdot 10^{-3}$	–



Rys. 30. Wyniki modelowania charakterystyk magnetostrykcyjnych ferrytu manganowo-cynkowego $\text{Mn}_{0,51}\text{Zn}_{0,44}\text{Fe}_{2,05}\text{O}_4$ z zastosowaniem modelu Jilesa–Athertona–Sabluka:
 a) pętla $\lambda(H)$, b) pętla $\lambda(B)$ [176]

Fig. 30. Results of modelling the magnetostrictive characteristics of $\text{Mn}_{0,51}\text{Zn}_{0,44}\text{Fe}_{2,05}\text{O}_4$ by Jiles–Atherton–Sabluk model: a) $\lambda(H)$ loop, b) $\lambda(B)$ loop [176]

5. NOWE MOŻLIWOŚCI WYKORZYSTANIA EFEKTU MAGNETOSPREŻYSTEGO W BUDOWIE SENSORÓW NAPRĘŻEŃ I SIŁ

5.1. WPROWADZENIE

Głównym problemem związanym z budową sensorów magnetosprężystych jest opracowanie odpowiedniej metody zadawania naprężeń w rdzeniu. Metoda ta powinna zapewniać wprowadzanie jednorodnych naprężeń do rdzenia z materiału magnetycznego. Jeżeli wprowadzane naprężenia będą niejednorodne, niektóre obszary rdzenia magnetycznego będą obciążone bardziej niż inne. Może to spowodować uszkodzenie (np. pęknięcie) rdzenia. Ponadto niejednorodny rozkład naprężeń w rdzeniu powoduje zmniejszenie czułości magnetosprężystej rdzenia, ponieważ łączne zmiany parametrów magnetycznych materiału są średnią zmianą stanu magnetycznego obszarów bardziej i mniej obciążonych mechanicznie.

Metoda obciążania rdzenia przetwornika magnetosprężystego powinna umożliwiać obciążanie rdzeni o zamkniętym obwodzie magnetycznym. W przypadku gdy obciążany jest rdzeń o otwartym magnetowodzie, jego czułość magnetosprężysta jest mniejsza niż rdzenia zamkniętego, do którego zostały wprowadzone naprężenia o tej samej wartości. Jest to związane z pojawieniem się energii odmagnesowania w całkowitym bilansie energii swobodnej w magnetyku. Z tego względu wpływ energii magnetosprężystej na stan magnetyczny materiału maleje.

Ze względu na maksymalną czułość magnetosprężystą najkorzystniejsza sytuacja występuje wtedy, gdy naprężenia przyłożone do materiału są równoległe do kierunku pola magnesującego w materiale. Gdy naprężenia σ_0 przyłożone do materiału nie są równoległe do pola magnesującego H , wówczas na właściwości magnetyczne wpływają odpowiadające im naprężenia σ dane wzorem (97).

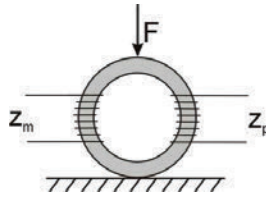
$$\sigma = \sigma_0(\cos \varphi - \nu \sin \varphi) \quad (97)$$

w którym ν – stała Poissona materiału, natomiast φ – kąt między kierunkiem naprężeń σ_0 i kierunkiem pola magnetycznego H .

W opisanych w literaturze badaniach w odniesieniu do magnetyków amorficznych stosowano głównie dwie metody obciążania: obciążano rdzeń pierścieniowy w kierunku średnicy lub próbkę paskową z materiału magnetycznego naklejoną na zginaną niemagnetyczną belkę.

Ideę opracowanej przez Mohriego [59] metody obciążania rdzenia pierścieniowego w kierunku średnicy przedstawiono na rysunku 31. W metodzie tej wykorzystywany jest rdzeń o zamkniętym magnetowodzie, jednak rozkład naprężeń w rdzeniu jest niejednorodny i trudny do wyznaczenia. Ponadto w rdzeniu występują zarówno naprężenia ściskające, jak i rozciągające. W rezultacie czułość ma-

gnetosprężysta rdzenia zostaje zmniejszona w stosunku do stanu, gdy naprężenia byłyby jednorodne.



Rys. 31. Obciążanie pierścieniowego rdzenia magnetycznego w kierunku średnicy [59]

(z_m – uzwojenie magnesujące, z_p – uzwojenie pomiarowe)

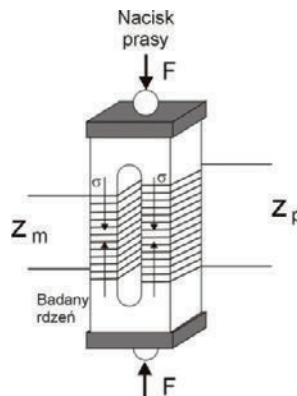
Fig. 31. Ring-shaped core subjected to compressive force applied in the direction of the diameter

(z_m – magnetizing winding, z_p – measuring winding)

Zastosowanie próbek paskowych naklejonych na niemagnetyczną, zginaną belkę umożliwia uzyskanie wyznaczalnego i w przybliżeniu jednorodnego rozkładu naprężeń w obciążanej próbce [61]. Jednak otwarty obwód magnetyczny próbki paskowej zmniejsza czułość magnetosprężystą ze względu na pojawienie się energii odmagnesowania w próbce. Ponadto uzyskiwane w takim sensorze wyniki są uzależnione od wymiarów geometrycznych próbki.

5.2. MAGNETOSPŘĘŻYSTE RDZENIE OKIENNE

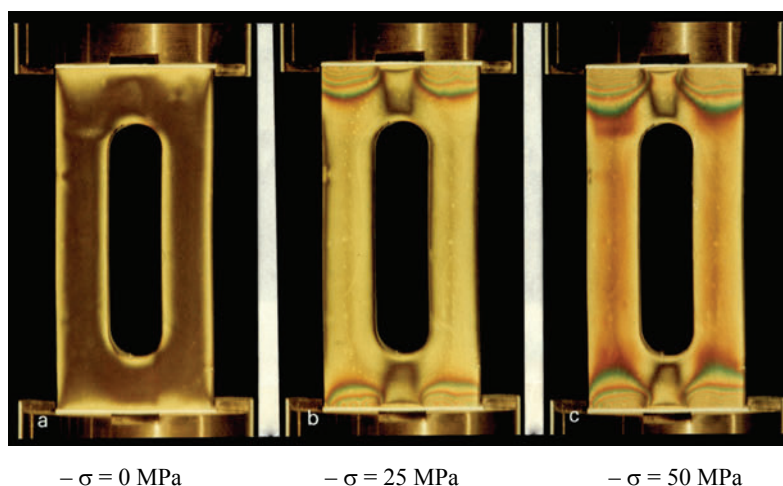
W przypadku sensorów z magnetosprężystymi rdzeniami z litych materiałów magnetycznych (takie jak stale, alferly lub ferryty) możliwe jest zastosowanie metody obciążania rdzeni okiennych według metody przedstawionej w pracy [98]. Ideę metody obciążania rdzeni okiennych przedstawiono na rysunku 32.



Rys. 32. Schematyczne przedstawienie obciążania magnetosprężystego rdzenia okiennego [178]

Fig. 32. Method of applying the stress to the frame-shaped magnetoelastic core [178]

Zaletą tej metody jest uzyskiwanie w rdzeniu okiennym o zamkniętym magneto-
wodzie jednorodnych naprężeń w całej długości kolumn. Z tego względu sensory ma-
gnetosprężyste naprężeń i sił, konstruowane z zastosowaniem rdzeni okiennych, mają
dużą czułość magnetosprężystą. Ponadto ze względu na jednorodny i wyznaczalny
rozkład naprężeń w kolumnach uzyskiwane charakterystyki magnetosprężyste mogą
być interpretowane fizycznie, co ułatwia prowadzenie badań procesów magnesowania
magnetyków w obecności naprężeń. Wadą tej metody są niejednorodności w rozkła-
dzie naprężeń w łącznikach kolumn. Jednak wpływ tego czynnika na uzyskiwaną
czułość magnetosprężystą jest niewielki ze względu na mały udział długości łączni-
ków kolumn w całej długości obwodu magnetycznego. Potwierdzają to badania ela-
stooptyczne rozkładu izochrom w rdzeniu ferrytowym w zależności od naprężeń ści-
skających – σ w kolumnach rdzenia, przedstawione na rysunku 33 [178].



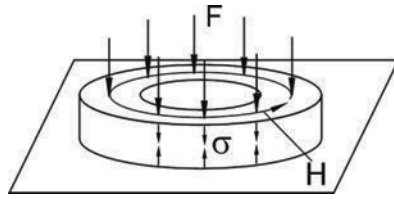
Rys. 33. Rozkład izochrom w rdzeniu ferrytowym
w zależności od naprężeń ściskających $-\sigma$ w kolumnach rdzenia
Fig. 33. Distribution of the isochromatic line in the ferrite frame core
as the function of compressive stresses

Kolejnym ograniczeniem w odniesieniu do magnetosprężystych metody rdzeni
okiennych jest brak możliwości zastosowania ich do badania nowoczesnych materia-
łów magnetycznych (w tym magnetyków amorficznych i nanokrystalicznych), które
są wytwarzane w postaci taśm grubości około 25 μm .

5.3. MAGNETOSPŘEŻYSTE RDZENIE PIERŚCIENIOWE

W opracowanej przez autorów metodzie badania magnetosprężystych właściwości
materiałów magnetycznych jako sensorów naprężeń i sił rdzeń pierścieniowy jest ob-

ciężany w kierunku prostopadłym do kierunku pola magnesującego w sposób przedstawiony na rysunku 34 [179, 180].



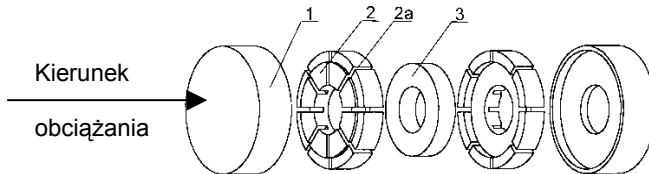
Rys. 34. Idea zadawania naprężeń ściskających do rdzenia pierścieniowego
Fig. 34. Idea of the method of applying compressive stresses to the ring-shaped core

Dla takiego sposobu obciążania rdzenia naprężenia efektywnie oddziałujące na parametry magnetyczne można wyznaczyć z zależności (97). Wynoszą one (dla kąta $\varphi = 90^\circ$):

$$\sigma = -\nu\sigma_0 \quad (98)$$

Głównym problemem dla takiego sposobu obciążania rdzenia pierścieniowego jest jego uzwojenie. Problem ten został rozwiązany za pomocą urządzenia przedstawionego na rysunku 35 [181].

Obciążany rdzeń pierścieniowy (3) jest umieszczony między dwoma cylindrycznymi nakładkami (2) wykonanymi z materiału niemagnetycznego. W nakładkach zostały wycięte rowki umożliwiające uzwojenie badanego rdzenia. Obciążenia ściskające przykładane są do podkładek bazowych (1).



Rys. 35. Urządzenie do zadawania naprężeń do rdzenia pierścieniowego (zgłoszenie patentowe P-345758) [181]: 1 – podkładki bazowe, 2 – cylindryczne nakładki z materiału niemagnetycznego, 2a – rowki na uzwojenie, 3 – badany rdzeń

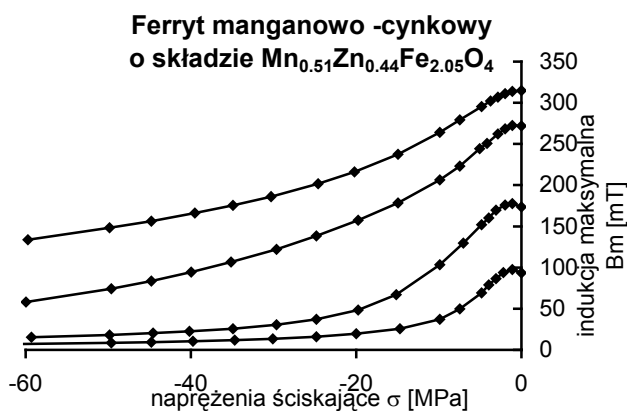
Fig. 35. Device for applying the compressive stresses to the ring-shaped core (patent pending P-345758) [181]: 1 – base backing, 2 – nonmagnetic cylindrical backing, 2a – grooves for windings, 3 – magnetoelastic sensing element

Główną zaletą tej metody jest uzyskanie jednorodnych i zdefiniowanych naprężeń wzdłuż całego, zamkniętego obwodu magnetycznego rdzenia pierścieniowego. Ponadto metoda ta ułatwia badanie zarówno rdzeni pierścieniowych litych, jak i zwijanych z cienkiej taśmy (jak np. w przypadku nowoczesnych magnetyków amorficznych i nanokrystalicznych). Ponieważ rozkład naprężeń jest jednorodny, w rdzeniu nie wy-

stępują obszary mocniej i słabiej obciążone. Dzięki temu rdzeń może być obciążany w szerszym zakresie naprężeń.

5.4. WYNIKI BADAŃ

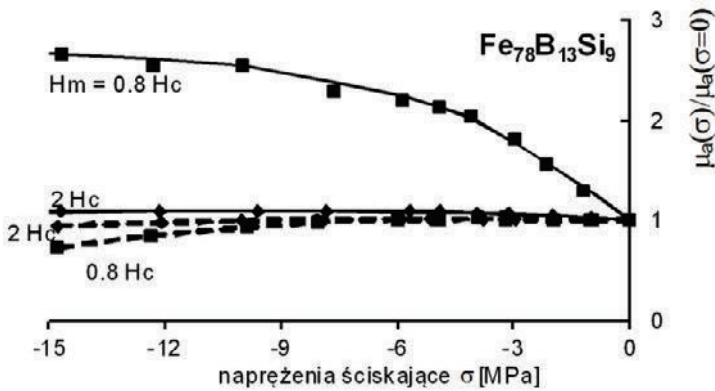
Na rysunku 36 przedstawiono wyniki badania magnetosprężystych właściwości magnetyka ceramicznego – ferrytu manganowo-cynkowego, dla różnych wartości pola magnesującego H_m będącego wielokrotnością pola koercji H_c . Badania przeprowadzono na rdzeniu okiennym. Uzyskane wyniki potwierdziły oczekiwaną wysoką czułość magnetyka na naprężenia mechaniczne od sił zewnętrznych, wyrażaną jako względne zmiany indukcji magnetycznej w rdzeniu (dla stałej wartości pola magnesującego w próbce).



Rys. 36. Zależności $B_m(\sigma)_H$ dla rdzenia ferrytowego $Mn_{0.51}Zn_{0.44}Fe_{2.05}O_4$ (próbka okienna o polu koercji $H_c = 8$ A/m)

Fig. 36. $B_m(\sigma)_H$ relations of the $Mn_{0.51}Zn_{0.44}Fe_{2.05}O_4$ ferrite (frame core of $H_c = 8$ A/m)

Badania przeprowadzono także dla próbek pierścieniowych wykonanych ze stopu amorficznego, o składzie $Fe_{78}B_{13}Si_9$, dla próbek przed i po wyżarzaniu usuwającym naprężenia wewnętrzne wprowadzone do materiału w trakcie procesu produkcji. Do badań zastosowano metodę opisaną w punkcie 5.3. Także w tym przypadku uzyskano duże zmiany parametrów magnetycznych próbki pod wpływem naprężeń od sił zewnętrznych. Na rysunku 37 przedstawiono względne zmiany przenikalności magnetycznej badanego rdzenia. Dla naprężeń 0–15 MPa, w przypadku próbki poddanej wyżarzaniu, zaobserwowano zwiększenie przenikalności magnetycznej próbki o 150%. Potwierdza to oczekiwania dużej czułości magnetosprężystej podczas obciążania rdzeni według opisanej metody



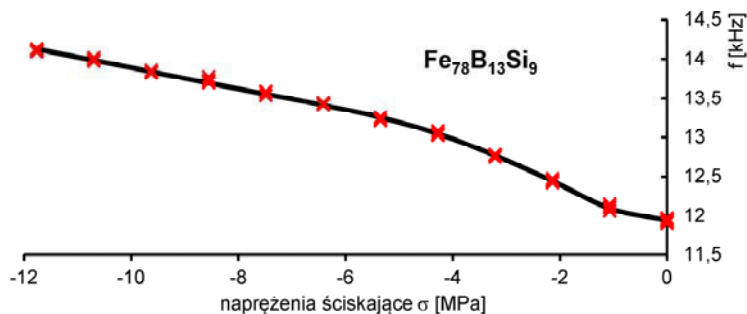
Rys. 37. Względne zmiany przenikalności magnetycznej pod wpływem naprężeń dla próbek pierścieniowych z $\text{Fe}_{78}\text{B}_{13}\text{Si}_9$ (---- próbka przed wyżarzaniem $H_c = 10$ A/m, — próbka po wyżarzaniu $H_c = 5$ A/m) dla różnych wartości pola magnesującego H_m
 Fig. 37. The influence of compressive stresses on the relative permeability of the $\text{Fe}_{78}\text{B}_{13}\text{Si}_9$ ring core, (---- core as quenched $H_c = 10$ A/m, — core after annealed state $H_c = 5$ A/m)

5.5. BUDOWA SENSORA MAGNETOSPREŻYSTEGO

Główną zaletą sensorów magnetosprężystych jest możliwość pominięcia przetwarzania mechanicznego w konstrukcji sensora. W przypadku np. sensorów tensometrycznych konieczne jest użycie elementu sprężystego, którego odkształcenia są przenoszone na tensometr. Natomiast w sensorach magnetosprężystych czuły na naprężenia rdzeń magnetosprężysty jest jednocześnie elementem konstrukcyjnym przenoszącym obciążenie. Ponadto sensory magnetosprężyste budowane na rdzeniach niektórych materiałów amorficznych mogą pracować w zakresie temperatury do 600 °C, czyli znacznie wyższym niż przetworniki tensometryczne.

W opracowanym modelu sensora uzwojony rdzeń pierścieniowy z materiału amorficznego pracuje w równoległym układzie LC generatora rezonansowego, a częstotliwość generatora zmienia się na skutek zmiany przenikalności amplitudowej rdzenia pod wpływem obciążeń mechanicznych. Uzyskiwany sygnał częstotliwościowy jest dogodny do przetwarzania cyfrowego oraz odporny na zakłócenia podczas przesyłania. Charakterystykę przetwarzania magnetosprężystego sensora siły z rdzeniem z materiału $\text{Fe}_{78}\text{B}_{13}\text{Si}_9$ po wyżarzaniu podano na rysunku 38.

W opisywanym sensorze uzyskano zmianę częstotliwości o $\Delta f = 2,2$ kHz dla naprężeń ściskających σ w zakresie 0–12 MPa (częstotliwość f dla $\sigma = 0$ wynosiła około 12 kHz). Estymator odchylenia standardowego rozrzutu wskazań s w punkcie pomiarowym wynosił 0,02 kHz.



Rys. 38. Charakterystyki przetwarzania sensora magnetoelastycznego z rdzeniem pierścieniowym z amorficznego stopu o składzie $\text{Fe}_{78}\text{B}_{13}\text{Si}_9$ po wyżarzaniu

Fig. 38. Characteristic of magnetoelastic sensor with ring-shaped core made of annealed amorphous $\text{Fe}_{78}\text{B}_{13}\text{Si}_9$ alloy

6. WNIOSKI

Przedstawione w artykule wyniki analiz teoretycznych i pomiarów wskazują, na znaczny potencjał rozwoju badań magnetomechanicznych właściwości materiałów magnetycznie miękkich. Badania te mogą być realizowane zarówno w aspekcie badań podstawowych, jak i badań związanych z przemysłowym wykorzystaniem zjawiska magnetoelastycznego w budowie sensorów naprężeń i sił. Podsumowując omawiane zagadnienia można wskazać następujące wnioski:

- Analiza procesów magnesowania i związanych z nimi zjawisk magnetostrykcyjnych i magnetoelastycznych powinna obejmować zarówno analizę procesów w skali mikroskopowej (związanych z procesami w skali atomowej), jak i procesami w skali makroskopowej (związanych z domenową strukturą materiałów magnetycznie miękkich).
- Model Jilesa–Athertona, rozwinięty przez Sablika i Ramesha, umożliwia skuteczne modelowanie zarówno charakterystyki magnesowania izotropowych i anizotropowych materiałów magnetycznie miękkich, jak i charakterystyk magnetoelastycznych i magnetostrykcyjnych tych materiałów.
- Identyfikacja parametrów modelu Jilesa–Athertona wymaga zastosowania nowoczesnych metod z zakresu optymalizacji takich, jak strategie ewolucyjne lub analiza fraktalna [182].
- Przedstawione w pracy metody zadawania jednorodnych naprężeń ściskających i rozciągających do rdzeni pierścieniowych i ramkowych umożliwiają badania zjawisk magnetoelastycznych zarówno w magnetykach ceramicznych (np. ferrytach), jak i w nowoczesnych magnetykach amorficznych i nanokrystalicznych.

- Z wykorzystaniem przedstawionej metodyki zadawania jednorodnych naprężeń do rdzenia magnetycznego istnieje możliwość budowy sensora, w którym uzyskuje się jednorodny i zdefiniowany rozkład naprężeń wzdłuż obwodu magnetycznego rdzenia zamkniętego. W rezultacie uzyskuje się zwiększenie czułości magnetosprężystej sensora.
- Sensory magnetosprężyste naprężeń i sił cechują się dużymi zmianami mierzonych parametrów magnetycznych (np. przenikalność amplitudowa zmienia się do 150% w zakresie naprężeń do 12 MPa).
- Uzyskiwany na wyjściu sensorów magnetosprężystych sygnał częstotliwościowy jest wygodny do przesyłania i przetwarzania cyfrowego.

LITERATURA

- [1] JOULE J.P., *On the effects of magnetism upon the dimensions of iron and steel bars*, Phil. Mag., 1847, 30, 2, 76–87 and 30, 4, 225–241.
- [2] VILLARI E., *Intorno alle modificazioni del momento magnetico di una verga di ferro e di acciaio, prodotte pur la trazione della medesima e pel passaggio di una corrente attraverso la stessa*, Nuovo Cim., 1864, 20, 317–62.
- [3] VILLARI E., *Change of magnetization by tension and by electric current*, Ann. Phys. Chem., 1865, 126, 87–122.
- [4] GUILLEMIN A.V., *Observation relatives an changement qui se produit dans l'elasticite d'un barreaude fer doux sous l'influence de l'electricite*, C.r.hebd. Seanc. Acad. Sci. Paris, 1846, 22, 264–265, 433–443.
- [5] WIEDEMENN G.H., *Über die Berichungen zwischen Magnetismus und Torsion*, Poggendorff Annln, Phys., 1858, 103, 563–77.
- [6] GARSHELIS J., *A study of the inverse Wiedemann effect on circular remanence*, IEEE Trans. Magn., MAG-10, 1974, 344–358.
- [7] BOZORTH R.M., *Ferromagnetism*, D. van Nostrod Co., New York 1956.
- [8] LEE E.W., *Magnetostriction and magnetomechanical effects*, Reports on Prog. in Physics, 1955, 18, 104–229.
- [9] CULLITY B.D., GRAHAM C.D., *Introduction to Magnetic Materials*, Second Edition, IEEE Press, Wiley Publication, New Jersey 2009.
- [10] CHIKAZUMI S., CHARP S.H., *Physics of Magnetism*, Krieger, Malabar, New York 1984.
- [11] AKULOV N.S., *Ferromagnetism*, GITL, Moskwa 1939.
- [12] AKULOV N.S., *Zur Theorie der Feinstruktur der Magnetisierungskurve der Einkristalle*, Phys., 1931, 69, 287.
- [13] BECKER R., *Zur Theorie des Ferromagnetismus*, Z. Phys., 62, 1930, 253–269.
- [14] BROWN, Jr., W.F., *Theory of magnetoelastic effects in ferromagnetism*, J. Appl. Phys., 1965, 36, (3), P.2, 994–1000.
- [15] BIELOW K.P., *Zjawiska w materiałach magnetycznych*, PWN, Warszawa 1962.
- [16] BIELOW K.P., BELANČIKOVA R.Z., LEVITIN R.Z., NIKITIN S.A., *Redkozemelnye ferromagnetycznyje antiferromagnetyki*, Izd. Nauka, Moskwa 1965.
- [17] WONSOWSKI S.W., *Współczesna nauka o magnetyzmie*, PWN, Warszawa 1958.
- [18] VONSOVSKI S.V., *Vlijanie slabych uprugich naprjaženij na načalnuju vosprimčivost ferromagnetyčkov*, Žurnal Eksperimentalnoj i Teoretičeskoj Fiziki, 1947, 17, 12, 1094–1105.
- [19] BECKER R., DÖRING W., *Ferromagnetismus*, J. Springer, Berlin 1939.

- [20] CHIH-WEN CHEN, *Magnetism and Metallurgy of Soft Magnetic Materials*, North-Holland Publishing Company, Amsterdam 1977.
- [21] HECK C., *Magnetic Materials and Their Applications*, Butterworth, London 1974.
- [22] KATZ H.W., *Współczesne elementy magnetyczne i dielektryczne*, WNT, Warszawa 1964.
- [23] SMOLIŃSKI A., *Zjawisko magnetostrykcji i materiały magnetostrykcyjne*, Rozpr. Elektrotech., 1959, 5, 2, 211–237.
- [24] KACZKOWSKI Z., *Applications of the magnetostrictive and piezomagnetic phenomena in radio-electronics and metrology*, VI Vedecka Konferencja Elektrotechnickej Fakulty, Zbornik Prednasok, Fyzika, 1992, 132–155.
- [25] KACZKOWSKI Z., *Współczynniki piezomagnetyczne ferrytów magnetostrykcyjnych i ich histereza magnetyczna*, Prace IF PAN, 39, Warszawa 1972, 1–378.
- [26] KACZKOWSKI Z., *Teoria domenowa magnetostrykcji*, w: *Materiały piezomagnetyczne i ich zastosowanie*, red. Kaczkowski Z., IF PAN–PWN, Warszawa 1978, 158–165.
- [27] WADAS R., *Ferrimagnetyzm*, PWN, Warszawa 1968.
- [28] SZYMCZAK H., ZUBEREK R., *Model of stress-induced anisotropy and magnetostriction in metallic glasses*, IEEE Trans. Magn., 1987, MAG-23, 2551–53.
- [29] SZYMCZAK H., ZUBEREK R., *Field dependence of forced magnetostriction in amorphous alloys*, J. Phys. F: Met. Phys., 1982, 12, 1841–44.
- [30] LACHOWICZ H.K., Szymczak H., *Magnetostriction of amorphous magnetic materials*, J. Magn. Magn. Mater., 1984, 41, 327–334.
- [31] JILES D.C., *Introduction to Magnetism and Magnetic Materials*, Chapman and Hall, London 1998.
- [32] RAMESH A., JILES D.C., RODERIK J., *Generalization of hysteresis modeling to anisotropic materials*, J. Appl. Phys., 1997, 81, 5585.
- [33] SABLİK M.J., BURKHARDT G.L., KWUN H., JILES D.C., *A model for the effect of stress on the low-frequency harmonic content of the magnetic induction in ferromagnetic materials*, J. Appl. Phys., 1988, 63, 3930.
- [34] CHWASTEK K., SZCZYGLÓWSKI J., *The effect of anisotropy in the modified Jiles–Atherton model of static hysteresis*, Archives of Electrical Engineering, 2011, 60, 49–57.
- [35] AUGUSTYNIAK B., DEGAUQUE J., *New approach to hysteresis process investigation using mechanical and magnetic Barkhausen effects*, J. Magn. Magn. Mater., 1995, 140–144, 187–189.
- [36] ANDREI P., CALTUN O., STANCU A., *Differential phenomenological models for the magnetization processes in soft MnZn ferrites*, IEEE Trans. Magn., 1998, 34, 231.
- [37] BECKER R., KERSTEN M., *Die Magnetisierung von Nickeldraht unter starken Zug*, Z. Phys., 1930, 64, 660–681.
- [38] BIENKOWSKI A., *Wpływ naprężeń ściskających na właściwości magnetyczne stali transformatorowej*, Rozprawy Elektrotechniczne, 1977, 23, z. 3, 645–656.
- [39] BIRSS B.R., *Magnetomechanical effects in the Rayleigh region*, IEEE Trans. Magn., 1971, MAG 7, 113–133.
- [40] DUNAEV F., *O magnitnoj teksture ferromagnetikow, podvergnutyh vnešnim nagruzkam*, Izvestija A.N. SSSR, Serija Fiz., 1961, 25, 12, 1502–1507.
- [41] GUEHIS M., SEGALINI S., *Effect of tensile stresses on magnetic properties of strain hardened steel*, Institut de Recherches de la Sid, erurgie 185, rue de President Roosevelt 78105 Saint Germain en Laye, 1987, RE1376, 1–14.
- [42] KERSTEN M., *Über die Abhängigkeit der magnetischen Eigenschaften des Nickels von den elastischen Spannungen*, Zs. f. Phys., 1931, 71, 553–592.
- [43] KIRCHNER H., *Über den Einfluss von Zug, Druck und Torsion auf die Ländsmagnetostraktion*, Ann. Phys., 1936, 5, 27, 49–69.

- [44] LANGMAN R., *The effect of stress on the magnetization of mild steel at moderate field strengths*, IEEE Trans. Magn., 1985, MAG-21, 1314–1320.
- [45] MOSES A., *Effects of stresses on magnetic properties of silicon-iron laminations*, J. Mater. Sci., 1974, 9, 217–222.
- [46] MOSES A.J., THOMAS B., *Stress sensitivity of the A. C. magnetic properties of permendur after various heat treatments*, IEEE Trans. Magn., 1976, MAG-12, 103–106.
- [47] SAITO A., YAMAMOTO K., YAMANE K., *Decrease of magnetostriction in positive magnetostrictive material due to tensile stress*, J. Magn. Magn. Mater., 1992, 112, 17–19.
- [48] SELF W.B., EDWARDS P.L., *Effect of tensile stress on the domain-nucleation field of iron whiskers*, J. Appl. Phys., 1972, 43, 199–202.
- [49] POTOCKY L., KOLLAR P., JURANEK Z., *The role of stresses in surface magnetic properties of some metallic glasses*, Acta Phys. Pol., 1987, A72, 801–805.
- [50] SAITO A., YAMAMOTO K., UKEDA S., *Reversible and irreversible magnetization changes of amorphous ribbons due stress*, J. Magn. Magn. Mater., 1992, 112, 41–43.
- [51] SAITO A., YAMAMOTO K., KUNIMORI O., *Effect of stress on amorphous bend cores*, J. Magn. Magn. Mater., 1992, 112, 278–280.
- [52] SASAKI T., TAKADA S., TOKUTAKE F., SAITO T., *Stress dependence of AC magnetic properties in thick amorphous ribbons*, J. Magn. Magn. Mater., 1992, 112, 356–358.
- [53] VAZQUEZ M., FERNENGEL W., KRONMÜLLER H., *The effect of tensile stresses on the magnetic properties of $\text{Co}_{58}\text{Fe}_5\text{Ni}_{10}\text{Si}_{11}\text{B}_{16}$ amorphous alloys*, Phys. Stat. Sol. (a), 1983, 80, 195–204.
- [54] WARLIMONT H., BOLL B., *Applications of amorphous soft magnetic materials*, J. Magn. Magn. Mater., 1982, 26, 97–105.
- [55] SHIRKOOHI G.H., BOUKHALFA M., *Variation of magnetostrictive properties in grain-oriented and non-oriented electrical steel with texture under applied linear stress*, J. Magn. Magn. Mater., 1992, 112, 222–224.
- [56] TIITTO S.I., *Influence of elastic and plastic strain of the magnetization process in Fe-3,5% Si*, IEEE Trans. Magn., 1976, MAG-12, 855–857.
- [57] WAYNE M., SWIFT M., WOLFE G.F., *Influence of compressive stress on magnetic properties of a (110) [001] single crystal of 3% Si-Fe*, IEEE Trans. Magn., 1976, MAG 12, 244–248.
- [58] CHIRAC H., HRISTOFOROU E., NEAGU M., BARARIU F., DARIE I., *Inverse Wiedemann effect in glass-covered amorphous wires*, Sensor and Actuators, 2000, 81, 147–149.
- [59] MOHRI K., *Application of amorphous alloys for sensors and transducers*, Rapidly Quenched Materials, Vol. II, North-Holland, Amsterdam 1984, 1687–1690.
- [60] KRAUS L., SVEC P., BYDZOVSKY J., *Amorphous CoFeCrSiB ribbon for strain sensing application*, J. Magn. Magn. Mater., 2002, 242–245, 241–243.
- [61] CHIRAC H., PLETEA M., HRISTOFOROU E., *Fe-based amorphous thin film as a magnetoelastic sensor material*, Elsevier. Sensors and Actuators A – Physical, 1 April 2000, Vol. A81, No. 1–3, 166–9.
- [62] MEYDAN T., *Application of amorphous materials to sensors*, J. Magn. Magn. Mater., 1994, 133, 525.
- [63] PFÜTZNER H., BENGTSOON C., SCHÖNHUBER P., *Spontaneous effects of stong stress on domains of transformer steel*, Physica Scripta, 1989, 39, 526–530.
- [64] KACZKOWSKI Z., *Magnetostrykcja i właściwości piezomagnetyczne ferrytów*, [w:] *Materiały piezomagnetyczne i ich zastosowanie*, red. Z. Kaczkowski, IF PAN-PWN, Warszawa 1978, 500–504.
- [65] NEEL L., *Propriétés magnetiques do l'état metallique et energie d'interaction entre atoms magnetique*, ANN. Phys., 1936, 5, 232–279.
- [66] LEE E.W., *Magnetostriction curves of polycrystalline ferromagnets*, Proc. Phys. Soc., 1965, 72, 249–258.
- [67] CALLEN E., GOLDBERG N., *Magnetostriction of polycrystalline aggregates*, J. Appl. Phys., 1965, 36, 976–977.

- [68] CALLEN E., *Magnetostriction*, J. Appl. Phys., 1968, 39 (2), 519–527.
- [69] CULLITY B.D., *Fundamentals of magnetostriction*, J. Metals, 1971, 23, 35–41.
- [70] RINALDI S., TURILLI G., *Theory of linear magnetoelastic effects*, Phys. Rev. B, 1985, 31, No. 5, 3051–3058.
- [71] SABLİK M.J., JILES D.C., *Coupled magnetoelastic theory of magnetic and magnetostrictive hysteresis*, IEEE Trans. Magn., 1993, 29, No. 3, 2113–2123.
- [72] SABLİK M.J., JILES D.C., *A model for hysteresis in magnetostriction*, J. Appl. Phys., 1988, 64 (10), 5402–5404.
- [73] CHIN G.Y., *Processing control of magnetic properties for magnetostrictive transducer applications*, J. Metals, 1971, 23, 42–45.
- [74] DORMAN J.L., Du TREMOLET E., *La magnétostriction des échantillons policristallins sous contrainte*, J. de Phys., 1967, 28, 456–464.
- [75] Du TREMOLET E., MORIN P., Rouchy J., *Magnetic and magnetoelastic properties of a ferromagnetic single crystals with cubic symmetry*, Ann. Phys., 1978, 3, 479–501.
- [76] Du TREMOLET E., *Magnetism IFundamentals, II Materials and Applications*, Springer, 2005.
- [77] Du TREMOLET E., MENDIA MONTERROSO R., *Magnetostriction of iron*, J. Magn. Magn. Mater., 1983, 31–34, 837–838.
- [78] HORNREICH R.M., RUBINSTEIN H., SPAIN R.J., *Magnetostrictive phenomena in metallic materials and some of their device applications*, IEEE Trans. Magn., 1971, MAG 7, 1, 29–48.
- [79] KIKUCHI Y., *Magnetostrictive materials and applications*, IEEE Trans. Magn., 1968, MAG 4, 2, 107–117.
- [80] SABLİK M.J., RUBIN S.W., *Relationship of magnetostrictive hysteresis to the ΔE effect*, J. Magn. Magn. Mater., 1992, 104–107, 392–394.
- [81] SABLİK M.J., KWUN H., *Hysteretic model for Barkhausen noise and the magnetically induced velocity change of ultrasonic waves in ferromagnets under stress*, J. Appl. Phys., 1991, 69, 5791–5793.
- [82] Du TREMOLET E., *Magnetoelastic properties of amorphous alloys*, J. Magn. Magn. Mater., 1982, 25, 251–270.
- [83] Du TREMOLET E., *Pair model of anisotropy and magnetostriction in soft amorphous magnets*, J. Magn. Magn. Mater., 1987, 67, 102–114.
- [84] FÄHNLE M., FURTHMÜLLER J., PAWELLEK R., *Theory and magnetostriction in amorphous ferromagnets*, Physics of Magnetic Materials, eds. W. Gorzkowski, H. Lachowicz, H. Szymczak, World Scientific Publishing Co., Singapore 1989, 228–247.
- [85] FÄHNLE M., FURTHMÜLLER J., PAWELLEK R., BEUERLE T., ELSÄSSER Ch., *Magnetostriction in amorphous ferromagnets: theory and interpretation of experiments*, Physics of Magnetic Materials, eds. W. Gorzkowski, H. Lachowicz, H. Szymczak, World Scientific Publishing Co., Singapore 1990, 204–241.
- [86] FÄHNLE M., EGAMI T., *Fenomenological theory of magnetostriction in amorphous ferromagnets*, J. Appl. Phys., 1982, 53, 2319–2320.
- [87] FÄHNLE M., ELSÄSSER C., FURTHMÜLLER J., PAWELLEK R., BRANDT E.H., *Local magnetic anisotropy and magnetostriction of amorphous ferromagnets*, Physica B, 1989, 161, 225–231.
- [88] KRAUS L., SCHNEIDER J., *Magnetostriction of amorphous $(Fe_{1-x}Ni_x)_{80}P_{10}B_{10}$ alloys*, Phys. Stat. Sol. (a), 1977, 39, K161–K163.
- [89] PAWELLEK R., FÄHNLE M., KRONMÜLLER H., *Theory of forced linear magnetostriction in amorphous ferromagnets*, Physica B, 1989, 161, 237–240.
- [90] TANGE H., KAMIMORI T., GOTO M., FUKAMICHI K., *Saturation and forced volume magnetostrictions of Fe-rich FeZr and FeNiZr amorphous alloys*, J. Magn. Magn. Mater., 1990, 90, 91, 335–336.
- [91] TANGE H., TAMAKA Y., GOTO M., *Saturation and forced volume magnetostriction of Fe-Rich FeZr amorphous alloys*, J. Magn. Magn. Mater., 1989, 81, L243–L246.

- [92] GIRAUD M., MORIN P., ROUCHY J., SCHMITT D., Du TREMOLET E., *Magnetoelastic properties of PrAg*, J. Magn. Magn. Mater., 1983, 37, 83–92.
- [93] MORIN P., ROUCHY J., Du TREMOLET E., *Magnetoelastic properties of RZn equiatomic compounds*, Phys. Rev. B, 1977, 16, 7, 3182–3193.
- [94] ROUCHY J., MORIN P., Du TREMOLET E., *Magnetic and magnetoelastic properties of GdZn single crystals*, J. Magn. Magn. Mater., 1981, 23, 59–68.
- [95] YONGMAO P., DAINING F., FAXIN L., *Giant forced volume magnetostriction in polycrystalline Tb_{0.3}Dy_{0.7}Fe_{1.95} alloys under magnetomechanical loading*, J. Magn. Magn. Mater., 2009, 321, 2783–2787.
- [96] KACZKOWSKI Z., *Piezomagnetyzm materiałów magnetostrykcyjnych*, Problemy fizyki magnetyków, IF UAM i IF PAN, Poznań 1975, 245–267.
- [97] KITTEL C., *Physical theory of ferromagnetic domains*, Rev. Mod. Phys., 1949, 21, 4, 541–583.
- [98] BIEŃKOWSKI A., *Magnetoelastyczne zjawisko Villariego w ferrytach i możliwość jego wykorzystania w budowie przetworników naprężeń i sił*, Oficyna Wydawnicza Politechniki Warszawskiej, Warszawa 1995.
- [99] Van VLECK J.H., *On the anisotropy of ferromagnetic crystals*, Phys. Rev., 1937, 52, 11, 1178–1198.
- [100] POSTUPOLSKI T., OKONIEWSKA-PSZCZÓLKOWSKA E., *Magnetic conditioning – does it always on the neutral state*, Acta Phys. Pol., 1985, A68, 1, 35–39.
- [101] JONES R., *Magnetoelastic interactions in ferrites*, IEEE Transactions on Sonics and Ultrasonic, 1966, SV-13, 3, 86–92.
- [102] BARTEL L.C., *Theory of strain-induced anisotropy and the rotation of magnetization in cubic single crystals*, J. Appl. Phys., 1969, 40, 2, 661–669.
- [103] OHTA K., KOBAYASHI N., *Magnetostriction constants of Mn-Zn-Fe ferrites*, Jap. J. App. Phys., 1964, 3, 10, 576–580.
- [104] MURTHY S.R., RAO T.S., *Magnetostriction of Ni-Zn and Co-Zn ferrites*, Phys. Stat. Sol. (a), 1985, 90, 631–635.
- [105] GREENOUGH R.D., LEE E.W., *The magnetostriction of cobalt-manganese ferrite*, J. Phys. D. Appl. Phys., 1970, 3, 1595–1604.
- [106] BIEŃKOWSKI A., SZEWCZYK R., WIŚNIEWSKA A., *Właściwości magnetostrykcyjne i magnetoelastyczne ferrytów Mn-Zn o wysokiej początkowej przenikalności magnetycznej*, Elektronika, 2004, 5, 29–30.
- [107] BIEŃKOWSKI A., KACZKOWSKI Z., *Major and minor magnetostriction hysteresis loop of Co-Cu-Ni ferrite*, J. Magn. Magn. Mater., 2000, 215–216, 234–236.
- [108] BIEŃKOWSKI A., KACZKOWSKI Z., SZEWCZYK R., *Magnetostrictive properties of Ni-ferrite*, Czechoslovak Journal of Physics, 2002, Vol. 52, Suppl. A, A85–A88.
- [109] KULIKOWSKI J., BIEŃKOWSKI A., *Magnetostriction of Ni-Zn ferrites containing cobalt*, J. Magn. Magn. Mater., 1982, 26, 297–299.
- [110] KULIKOWSKI J., BIEŃKOWSKI A., *Magnetostrictive properties of Co_x(NiZn)_{1-x}Fe₂O₄ ferrites in the case of small changes of iron content*, J. Magn. Magn. Mater., 1984, 41, 63–65.
- [111] KULIKOWSKI J., BIEŃKOWSKI A., *Field, temperature and stress dependence of magnetostriction in Ni-Zn ferrites containing cobalt*, Physica Scripta, 1991, 44, 382–383.
- [112] SZEWCZYK R., BIEŃKOWSKI A., KOLANO-BURIAN A., *Magnetostrictive properties of Fe₄₀Ni₃₈Mo₂B₁₈ alloy*, Materials Science & Engineering A, 2004, 375–377, 1137–1139.
- [113] BROWN, Jr., W.F., *Theory of the approach to magnetic saturation*, Phys. Rev., 58, 1940, 736–743.
- [114] KULIKOWSKI J., *Anizotropia magnetyczna w ferrytach*, [w:] *Materiały piezomagnetyczne i ich zastosowanie*, red. Z. Kaczkowski, IF PAN–PWN, Warszawa 1978, 493–499.
- [115] JILES D.C., *Introduction to Magnetism and Magnetic Materials*, Stanley Thornes Publ., 1998.
- [116] KACZKOWSKI Z., *Magnetostrykcja, termostrykcja i mechanostrykcja*, [w:] *Materiały piezomagnetyczne i ich zastosowanie*, red. Z. Kaczkowski, IF PAN–PWN, Warszawa 1978, 105–114.

- [117] ALLIA P., FERRO-MILONE A., MONTALENTI G., SOARDO G.P., VINAI F., *Theory of negative magnetostriction in grain oriented 3% SiFe for various inductions and applied stresses*, IEEE Trans. Magn., 1978, MAG-14, 362–364.
- [118] Du TREMOLET E., *Change of sign of the magnetostriction of iron under compressive stress*, J. Magn. Magn. Mater., 1979, 13, 307–309.
- [119] HOLT C., ROBEY J., *The AC magnetostriction of 3,25-percent grain-oriented silicon-iron under combined longitudinal and normal compressive stress*, IEEE Trans. Magn., 1969, MAG 5, 384–388.
- [120] KURUZAR M.R., CULLITY B.D., *The magnetostriction of iron under tensile and compressive stress*, Intern. J. Magnetism, 1971, 1, 323–325.
- [121] MOSES A.J., PHILLIPS P.S., *Some effects of stress in goss-oriented silicon-iron*, IEEE Trans. Magn., 1978, MA6-14, 5, 353–355.
- [122] SIMMONS G., THOMPSON E., *Magnetic properties of grain-oriented silicon-iron. Stress sensitivity of the dynamic 50 Hz magnetostriction*, Proc. IEE, 1971, 118, 9, 1302–1306.
- [123] SZYMCZAK H., *Mechanism of stress dependence of saturation magnetostriction in amorphous alloys*, J. Magn. Magn. Mater., 1987, 67, 227–231.
- [124] SZYMCZAK H., LACHOWICZ H., *Stress-induced magnetic phenomena in metallic glasses*, IEEE Trans. Mag., 1988, 24, 1747–1751.
- [125] LACHOWICZ H.K., SZYMCZAK H., *Stress dependence of saturation magnetostriction in metallic glasses*, *Magnetic Properties of Amorphous metals*, A. Hernando, V. Madurga, M.C. Sanchez-Trujillo, M. Vazquez (eds.), Elsevier Science Publishers B.V., 1987, 232–237.
- [126] LACHOWICZ H.K., SIEMKO A., MOSER N., FORKL A., KRONMÜLLER H., *Direct evidence for stress-dependent saturation magnetostriction in metallic glasses*, Phys. Stat. Sol. (a), 1988, 105, 597–600.
- [127] HERNANDO A., GOMEZ-POLO C, PULIDO E., RIVERO G., VAZQUEZ M., *Tensile-stress dependence of magnetostriction in multilayers amorphous ribbons*, Physical Review B, 1990, 42, 6471–6475.
- [128] HERZER G., *Effect of external stresses on the saturation magnetostriction of amorphous Co-based alloys*, Proc. Internat. Conf. Soft Magnetic Materials 7, Wolfson Centre for Magnetic Technology, Cardiff 1985, 355–357.
- [129] SIEMKO A., LACHOWICZ H.K., MOSER N., FORKL A., KRONMÜLLER H., *Stress-induced sign change of λ_s in low-magnetostrictive metallic glass as seen by its domain structure*, Phys. Stat. Sol. (a), 1989, 113, 181–185.
- [130] SIEMKO A., LACHOWICZ H.K., *On indirect measurements of saturation magnetostriction in low-magnetostrictive metallic glasses*, IEEE Trans. Magn., 1987, MAG-23, 2563–2565.
- [131] SIEMKO A., LACHOWICZ H., *Comments on the indirect measurement of magnetostriction in low-magnetostrictive metallic glasses*, J. Magn. Magn. Mater., 1987, 66, 31–36.
- [132] SUZUKI Y., HAIMOVICH J., EGAMI T., *Bond-orientational anisotropy in metallic glasses observed by x-ray diffraction*, Phys. Rev., 1987, 35, 2162–2168.
- [133] CLARK A.E., SPANO M.L., SAVAGE H.T., *Effect of stress on the magnetostriction and magnetization of rare earth- $Re_{1.95}$ alloys*, IEEE Trans. Magn., 1983, MAG-19, 1964–1966.
- [134] CLARK A.E., SAVAGE H.T., SPANO M.L., *Effect of stress on the magnetostriction and magnetization of single crystal $Tb_{27}Dy_{73}Fe_2$* , IEEE Trans. Magn., 1984, MAG-20, 1443–1445.
- [135] SAVAGE H.T., CLARK A.E., SPANO M.L., *Strain-field, relationships in rare-earth iron alloys*, IEEE Trans. Magn., 1984, MAG-20, 1449–1450.
- [136] LIORZOU F., PHELPS B., ATHERTON D. L., *Macroscopic models of magnetization*, IEEE Trans. Magn. 2000, 36, 418.
- [137] STONER E.C., WOHLFART E.P., *A mechanism of magnetic hysteresis in heterogeneous alloy*, Phil. Trans. Roy. Soc., 1948, 240A, 599.
- [138] GLOBUS A., *Some physical consideration about the domain wall size*, Theory of magnetization mechanism, J. de Physique, 1977, C1, C1-1.

- [139] BERTOTTI G., *Dynamic generalization of the scalar Preisach model of hysteresis*, IEEE Trans. Magn., 1992, 28, 2599.
- [140] JILES D.C., ATHERTON D., *Theory of ferromagnetic hysteresis*, J. Magn. Magn. Mater., 1986, 61, 48.
- [141] SABLİK M.J., BURKHARDT G.L., KWUN H., JILES D.C., *A model for the effect of stress on the low-frequency harmonic content of the magnetic induction in ferromagnetic materials*, J. Appl. Phys., 1988, 63, 3930.
- [142] SABLİK M.J., KWUN H., BURKHARDT G.L., JILES D.C., *Model for the effect of tensile and compressive stress on ferromagnetic hysteresis*, J. Appl. Phys., 1987, 61, 3799.
- [143] BOZORTH R.M., *Ferromagnetism*, D. van Nostrand, New York 1956.
- [144] TREMOLET E., *Magnetostriction*, CRC Press, London 1992.
- [145] CHAN J.H., VLADIMIRESCU A., GAO X.C., LIEBMANN P., VALAINIS J., *Nonlinear Transformer Model for Circuit Simulation*, IEEE Transactions on Computer-Aided Design, 1991, 1, 476.
- [146] KITTEL C., *Wstęp do fizyki ciała stałego*, PWN, Warszawa 1999.
- [147] JILES D.C., ATHERTON D., *Theory of ferromagnetic hysteresis*, J. Appl. Phys., 1984, 55, 2115.
- [148] JILES D.C., RAMESH A., SHI Y., FANG X., *Application of the anisotropic extension of the theory of hysteresis to the magnetization curves of crystalline and textured magnetic materials*, IEEE Trans. Magn., 1997, 33, 3961.
- [149] RAMESH A., JILES D.C., BI Y., *Generalization of hysteresis modeling to anisotropic materials*, J. Appl. Phys., 1997, 81, 5585.
- [150] RAMESH A., JILES D., RODERIK J., *A model of anisotropic anhysteretic magnetization*, IEEE Trans. Magn., 1996, 32, 5, 4234–4236.
- [151] SHAMPINE L.F., *Vectorized Adaptive Quadrature in MATLAB*, Journal of Computational and Applied Mathematics, 2008, 211, 131–140.
- [152] CHWASTEK K., SZCZYGLÓWSKI J., *Identification of a hysteresis model parameters with genetic algorithms*, Mathematics and Computers in Simulation, 2006, 71, 206.
- [153] DEANE J.H.B., *Modeling the dynamics of nonlinear inductor circuits*, IEEE Trans. Magn., 1994, 30, 2795.
- [154] VENKATARAMAN R., *Modelling and adaptive control of magnetostrictive actuators*, Ph.D. Thesis, University of Maryland, 1999.
- [155] TAN D., CHEN Z., *On A General Formula of Fourth Order Runge–Kutta Method*, Journal of Mathematical Science and Mathematics Education, 2012, 7.2, 1–10.
- [156] DORMAND J.R., PRINCE P.J., *A family of embedded Runge–Kutta formulae*, J. Comp. Appl. Math., 1980, Vol. 6, 19–26.
- [157] JILES D.C., THOELKE J.B., *Theory of ferromagnetic hysteresis: determination of model parameters from experimental hysteresis loops*, IEEE Trans. Magn., 1989, 25, 3928.
- [158] JILES D.C., THOELKE J.B., DEVINE M.K., *Numerical determination of hysteresis parameters for the modeling of magnetic properties using the theory of ferromagnetic hysteresis*, IEEE Trans. Magn., 1992, 28, 27.
- [159] HERNANDEZ E.D.M., MURANAKA C.S., CARDOSO J.R., *Identification of the Jiles–Atherton model parameters using random and deterministic searches*, Physica B, 2000, 275, 212.
- [160] LEDERER D., IGARASHI H., KOST A., HONMA T., *On the parameter identification and application of the Jiles–Atherton hysteresis model for numerical modelling of measured characteristics*, IEEE Trans. Magn., 1999, 35, 1211.
- [161] WILSON P.R., ROSS J.N., BROWN A.D., *Optimizing the Jiles–Atherton model of hysteresis by a genetic algorithm*, IEEE Trans. Magn., 2001, 37, 989.
- [162] SCHWEFEL H.P., *Evolution and optimum seeking*, Wiley, 1995.
- [163] SABLİK M.J., KWUN H., *Hysteretic model for Barkhausen noise and the magnetically induced velocity change of ultrasonic waves in ferromagnets under stress*, J. Appl. Phys., 1991, 69, 5791.

- [164] SABLİK M.J., KWUN H., BURKHARDT G.L., JILES D.C., *Model for the effect of tensile and compressive stress on ferromagnetic hysteresis*, J. Appl. Phys., 1987, 61, 3799.
- [165] GAUNT P., *Domain wall pinning at random inhomogeneities*, IEEE Trans. Magn., 1983, 19, 2030.
- [166] ROUABHI M., GUYOT M., CAGAN V., KRISHNAN R., *Domain wall pinning and hysteresis losses in amorphous CoNbZr films*, IEEE Trans. Magn., 1993, 29, 3502.
- [167] SZEWCZYK R., *Extension for the model of the magnetic characteristics of anisotropic metallic glasses*, Journal of Physics D – Applied Physics, 2007, 40, 4109.
- [168] CLAYCOMB J.R., BRAZDEIKIS A., LE M., YARBROUGH R.A., GOGOSHIN G., MILLER J.H., *Nondestructive testing of PEM fuel cells*, IEEE Trans. Appl. Supercond., 2003, 13, 211.
- [169] PREOBRAŽENSKIJ A., BIŠARD Je., *Magnitnye materialy i elementy*, Izd. Wysšaja Škola, Moskva 1986.
- [170] CULITY B.D., *Introduction to magnetic materials*, Addison Wesley, 1972.
- [171] OBI Y., *The stress effect on magnetic domain structure of Fe-Co amorphous alloys*, Applied Physics A: Materials Science and Processing, 1979, 18, 119.
- [172] SHISHIDO H., KAN T., IT Y., *The magnetic domain and properties of amorphous ribbons*, IEEE Trans. Magn., 1985, 21, 49.
- [173] JILES D.C., *Introduction to Magnetism and Magnetic Materials*, Chapman and Hall, London 1998.
- [174] SZEWCZYK R., BIEŃKOWSKI A., SALACH J., *Extended Jiles-Atherton model for modelling the magnetic characteristics of isotropic materials*, Journal of Magnetism and Magnetic Materials, 2008, 320, e1049.
- [175] SZEWCZYK R., SALACH J., BIEŃKOWSKI A., *Modelling the magnetoelastic materials for force and torque sensors*, Solid State Phenomena, 2009, 144, 124.
- [176] SZEWCZYK R., *Modelling of the magnetic and magnetostrictive properties of high permeability Mn-Zn ferrites*, Pramana Journal of Physics, 2006, 67, 1165.
- [177] JILES D., *Theory of the magnetomechanical effect*, J. Appl. Phys., 1995, 28, 1537.
- [178] BIEŃKOWSKI A., *Some problems of measurement of magnetostriction in ferrites under stresses*, J. Magn. Magn. Mater., 1992, 112, 143–145.
- [179] BIEŃKOWSKI A., SZEWCZYK R., *Magnetosprężyste sensory naprężeń i sił*, Pomiar Automatyka Robotyka, 2002, 9, 14–17.
- [180] BIEŃKOWSKI A., SZEWCZYK R., *The possibility of utilizing the high permeability magnetic materials in construction of magnetoelastic stress and force sensors*, Sensors and actuators, A, 2004, 113, 270–276.
- [181] BIEŃKOWSKI A., SZEWCZYK R., *Urządzenie do obciążania pierścieniowego przetwornika magnetosprężystego*, Zgłoszenie Patentowe P-345758, 2001.
- [182] MROCZKA J., WOŹNIAK M., ONOFRI R.A., *Algorithms and methods for analysis of the optical structure factor of fractal aggregates*, Metrology and Measuring Systems, 2012, 19, 459–470.

MAGNETOELASTIC PROPERTIES OF SOFT MAGNETIC MATERIALS FOR STRESS AND FORCE SENSORS

Paper presents magnetoelastic properties of both crystalline and amorphous magnetic materials, from the point of view of possible applications in development of force and stress sensors. Theoretical analyses, results of quantitative modeling and experimental results are presented. Within theoretical analyses the role of magnetostriction is considered. On the other hand, new methods of generation of uniform mechanical stresses in ring shaped and frame-shaped samples was developed. These methods enabled measurements of the influence of mechanical stresses on the shape of hysteresis loop of different magnetic materials.

On the base of theoretical analyses Jiles–Atherton–Sablik model of magnetic hysteresis was applied also for magnetomechanical modeling. It should be highlighted, that very good agreement between the results of modeling and experiments was achieved. Finally, conception of robust and sensitive magneto-elastic sensor for industrial applications is presented.

*skaningowa mikroskopia termiczna, pomiary temperatury,
właściwości cieplne mikro- i nanostruktur,
mikroskopia bliskich oddziaływań*

Grzegorz WIELGOSZEWSKI*, Maciej RUDEK*,
Grzegorz JÓZWIAK*, Teodor GOTSZALK*

SKANINGOWA MIKROSKOPIA BLISKIEGO POLA TERMICZNEGO W METROLOGII MIKRO- I NANOSTRUKTUR

Obserwowany w ostatnich latach postęp w dziedzinie mikro- i nanotechnologii, którego skutki obserwujemy na co dzień, wiąże się z postępowaniem w zakresie wytwarzania mikro- i nanostruktur, których przynajmniej jeden z wymiarów jest rzędu dziesiątek nanometrów. To dzięki temu postępowi możliwe staje się powszechne wytwarzanie układów nanoelektronicznych, które są powszechnie integrowane w komputerach osobistych, sprzęcie powszechnego użytku czy urządzeniach gospodarstwa domowego. Warto równocześnie zwrócić uwagę, że wspomniane mikro- i nanostruktury mogą mieć charakter nie tylko przyrządowy (jak w przypadku nano- i mikroelektronicznych układów scalonych), ale również materiałowy (jak w przypadku nowoczesnych leków i diagnostyki biomedycznej, powłok powierzchni itp.). Podzespoły te stają się coraz częściej tak małe, że o ich funkcji zaczynają decydować efekty kwantowe.

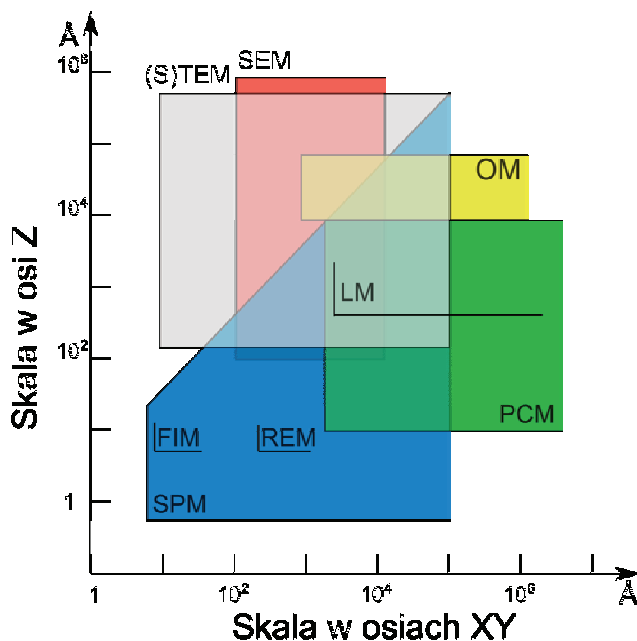
Ten nadzwyczaj intensywny postęp, który umożliwia zwykłemu użytkownikowi sięgać po coraz doskonalsze rozwiązania, jest jednak ściśle powiązany z postępowaniem badań tego typu podzespołów. W wielu przypadkach wystarczające bywają obserwacje prowadzone w makroskali, które ułatwiają niejako uśrednioną lub innymi słowy wypadkową ocenę opracowanego układu zbudowanego faktycznie z niepoliczalnie wielu małych struktur. Dogłębne zrozumienie zjawisk decydujących o istocie reakcji zachodzących w mikro- i nanoskali, które w konsekwencji umożliwiać będzie optymalizację konstrukcji, wymaga jednak coraz częściej użycia metod i technik, za pomocą których obserwacja zjawiska może być prowadzona w obszarach najmniejszych struktur składowych. W tym ujęciu wyróżnić można dwie zasadnicze tendencje towarzyszące postępowi badań mikro- i nanostruktur. Pierwsza z tych tendencji dotyczy

* Zakład Metrologii Mikro- i Nanostruktur, Wydział Elektroniki Mikrosystemów i Fotoniki, Politechnika Wrocławska, ul. Z. Janiszewskiego 11/17, 50-372 Wrocław.

opracowania i zastosowania metod i technik badawczych umożliwiających prowadzenie obserwacji z rozdzielczością ułamków nanometrów. W tak zlokalizowanym obiekcie należy przeprowadzić eksperyment, który powinien dać obraz jego właściwości fizycznych lub chemicznych. Czułość takiego obrazowania powinna stwarzać warunki do obserwacji nawet efektów kwantowych. Druga z tendencji towarzyszących postępowi badań mikro- i nanostruktur związana jest z opracowaniem metod ilościowej (innymi słowy metrologicznej) oceny zjawisk w mikro- i nanostrukturach. O ile oceny jakościowe (wskazujące np. na różnice lub stosunek poszczególnych parametrów) są przeprowadzane stosunkowo często w praktyce laboratoryjnej, o tyle dokonywanie metrologicznej analizy właściwości mikro- i nanostruktur napotyka wiele trudności. Do najważniejszych z nich zaliczyć należy:

- Kłopotliwą interpretację obserwowanych zjawisk. Interpretacja oparta na klasycznej wytrzymałości materiałów, fizyce, chemii i elektrotechnice, która wynika często ze statystycznego uśrednienia, w przypadku struktur tworzonych przez pojedyncze atomy lub molekuly nie znajduje przełożenia wprost. Zagadnienia sprężystości, przewodności elektrycznej i termicznej muszą być zatem analizowane modelami kwantowymi, co bardzo często prowadzi do ogromnie złożonych obliczeń, których wyniki nie są intuicyjnie jasne.
- Brak powszechnie uznanych i stosowanych w praktyce laboratoryjnej technik i metod wzorcowania narzędzi oraz systemów badawczych przeznaczonych do badania mikro- i nanostruktur. Powoduje to, że nadzwyczaj utrudnione staje się porównywanie wyników uzyskiwanych w różnych laboratoriach. Widoczne to jest szczególnie w pomiarach wymiarów geometrycznych mikro- i nanostruktur, w których obserwacje struktur o wielkości poniżej około 100 nm obarczone są niepewnością (ok. 20%).

Warto również zaznaczyć, że w powszechnym zainteresowaniu uniwersyteckich i przemysłowych laboratoriów badawczych zagadnieniami nanotechnologii dostępność do tak zarysowanych technologii powinna być możliwie jak najłatwiejsza. Innymi słowy, metody i techniki pomiaru powinny być zatem na tyle dojrzałe, aby koszty użytkowania aparatury były odpowiednio małe, obsługa urządzeń zaś możliwie najprostsza. Doświadczenia Zakładu Metrologii Mikro- i Nanostruktur (ZMMiN) Wydziału Elektroniki Mikrosystemów i Fotoniki (WEMiF) Politechniki Wrocławskiej (PWr.) wskazują, że metody i techniki, które mogą zapewnić postęp w obserwacjach i pomiarach mikro- i nanostruktur, opierają się na mikroskopii optycznej, elektronowej i bliskich oddziaływań (ang. *Optical, Electron and Scanning Probe Microscopy* – OM, EM, SPM). Na rysunku 1 przedstawiona jest schematycznie charakterystyka poszczególnych technik obrazowania mikro- i nanostruktur; zaznaczone są skale w osiach XY odpowiadające wielkości planarnej obrazowanych mikro- i nanostruktur; oraz w osi Z odpowiadającej wysokości układu. Widać wyraźnie z przedstawionego obrazu, że nie istnieje jedna uniwersalna technologia umożliwiająca obserwacje i ewentualne badanie mikro- i nanostruktur. Szansą wydaje się dopiero integracja kilku technik z precyzyjnymi układami pomiarów makroskopowych.



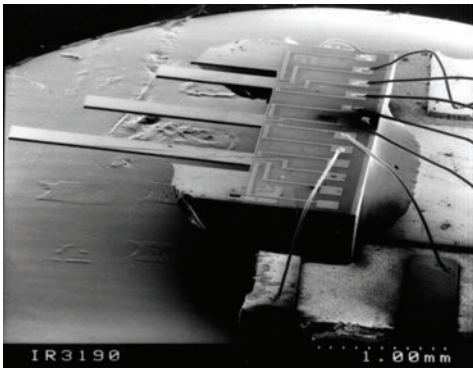
Rys. 1. Różne metody obrazowania mikro- i nanostruktur; LM – mikroskopia optyczna, REM – odbiciowa mikroskopia elektronowa, FIM – jonowa mikroskopia emisyjna, OM – wysokorozdzielcza mikroskopia optyczna, PCM – mikroskopia kontrastu fazowego, SEM – skaningowa mikroskopia elektronowa, TEM – transmisyjna mikroskopia elektronowa, SPM – mikroskopia bliskich oddziaływań (wg: G. Binnig, H. Rohrer, *Scanning tunneling microscopy*, Helvetica Physica Acta, 55, 1982, 726–735)

Fig. 1. Various imaging methods for micro- and nanostructures: LM – light microscopy, REM – reflection electron microscopy, FIM – field ion microscopy, OM – high-resolution optical microscopy, PCM – phase contrast microscopy, SEM – scanning electron microscopy, TEM – transmission electron microscopy, SPM – scanning probe microscopy (based on G. Binnig, H. Rohrer, *Scanning tunneling microscopy*, Helvetica Physica Acta, 55, 1982, 726–735)

Rozwój mikroskopii optycznej datuje się od końca XVII wieku, gdy Robert Hooke i Antoni van Leeuwenhoek skonstruowali pierwsze mikroskopy optyczne. Ich zasada działania może być wyjaśniana zasadami optyki geometrycznej, zgodnie z którą powiększenie jest funkcją parametrów soczewek obiektywu i okularu oraz długości tubusa.

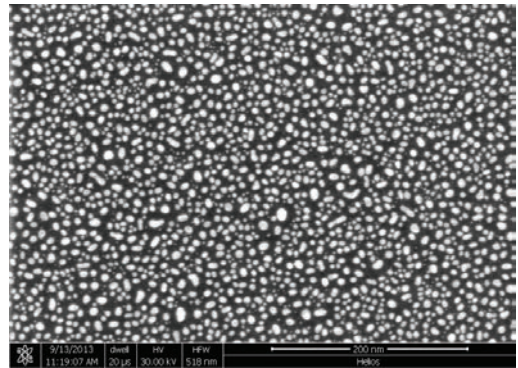
Falowa natura światła powoduje jednak, że zdolność rozdzielcza takiego mikroskopu jest w pierwszej kolejności ograniczona przez ugięcie promieni, co powoduje, że uzyskanie powiększeń obliczanych zgodnie z prostymi równaniami optyki geometrycznej jest niemożliwe. Warto wskazać, że ogromne zasługi dla rozwoju mikroskopii optycznej wniósł niemiecki rzemieślnik Carl Zeiss oraz współpracujący z nim fizyk-optyk Ernst Abbe. Para tych dwóch uczonych i konstruktorów stworzyła podstawy dla współczesnej mikroskopii optycznej. Trudno wymienić choć część technik opartych na osiągnięciach pierwszych konstruktorów mikroskopów optycznych. Urządzenia

te stosowane są w przemyśle, laboratoriach mikrobiologicznych oraz inżynierskich. W połączeniu z komputerowymi technikami analizy i przetwarzania obrazów stały się nadzwyczaj wygodnym narzędziem do obrazowania zachowań mikroświata. Zasadniczymi słabościami tej technologii jest ograniczona zdolność rozdzielcza umożliwiająca w praktyce laboratoryjnej obserwację szczegółów o rozmiarach rzędu setek nanometrów oraz niewielką głębię ostrości, która ułatwia obrazowanie szczegółów jedynie w najbliższym otoczeniu znajdujących się w polu ostrości. Słabości te eliminują w pewnym zakresie mikroskopy elektronowe skonstruowane po raz pierwszy przez Ernsta Ruskę i Maxa Knolla w latach 30. XX w. w Politechnice w Berlinie. Z ich wspomnień wynika, że prowadząc pionierskie eksperymenty uznali oni elektrony za medium „drobniejsze” od światła, dzięki czemu możliwe być powinno oglądanie detali mniejszych aniżeli w przypadku urządzeń optycznych. Ścisłą interpretację tego



Rys. 2. Obraz dźwigni mikromechanicznych zarejestrowany za pomocą skaningowego mikroskopu elektronowego – obraz o małym powiększeniu, widoczne liczne szczegóły wokół preparatu wynikające z dużej głębi ostrości obrazowania [1]

Fig. 2. Scanning electron microscope image of micromechanical cantilevers – small magnification image, many details are visible because of large depth of field [1]



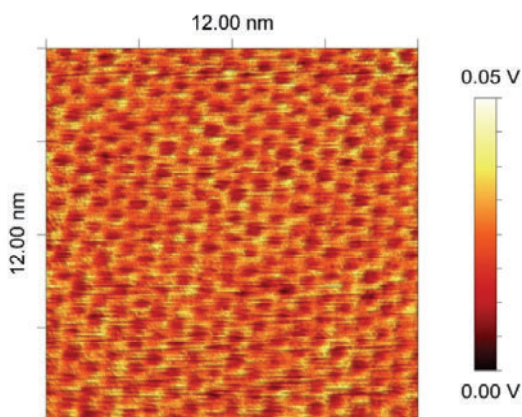
Rys. 3. Obraz struktur złota osadzonych na podłożu węglowym zarejestrowany za pomocą skaningowego mikroskopu elektronowego – obraz o bardzo dużym powiększeniu (1 milion razy)

Fig. 3. Scanning electron microscope image of gold structures deposited on carbon substrate – very large magnification image (1 million times)

przecucia zapewniła teoria korpuskularno-falowa zaproponowana przez Louisa de Broglie’a w latach 20. XX w., zgodnie z którą poruszające się elektrony można opisać falą o ściśle zdefiniowanej długości. Zastosowanie typowych napięć przyspieszających (np. rzędu pojedynczych dziesiątek kilowoltów) sprawia, że w najprostszym przybliżeniu wiązka elektronów kierowana na powierzchnię preparatu ma długość zdecydowanie mniejszą aniżeli światło widzialne zastosowane w układach Carla Zeissa. Ułatwia to obrazowanie szczegółów odpowiednio mniejszych, a co warto szcze-

gólnie podkreślić, technicznie możliwe jest konstruowanie układów ogniskujących elektrony (tzw. soczewek elektronowych) o stosunkowo małych aperturach numerycznych, czego konsekwencją jest ewentualność prowadzenia obserwacji z odpowiednio zwiększoną głębią ostrości. Te dwie cechy stanowią o powszechnym zastosowaniu mikroskopii elektronowej w praktyce laboratoriów uniwersyteckich i przemysłowych. Jednocześnie dzięki automatyzacji pracy i akwizycji obrazów oraz obserwacji preparatów w próżni o różnym ciśnieniu zwiększa się istotnie atrakcyjność tej technologii.

Czas przełomowych odkryć technicznych związanych z obserwacją mikroskopową mikro- i nanostruktur kończy skonstruowanie skaningowego mikroskopu tunelowego przez Gerda Binniga i Heinricha Rohrera w latach 80. XX w. w laboratoriach firmy IBM w Rüschlikon w Szwajcarii [2]. W swoich pierwszych eksperymentach zbliżali oni do powierzchni przewodzącego preparatu zaostroszony drut metalowy. Zgodnie z regułami mechaniki kwantowej prawdopodobieństwo znalezienia elektronów poza objętością litej próbki zwiększa się, jeśli odległość do powierzchni będzie ulegać zmniejszeniu. Jeśli odległość ta będzie rzędu ułamków nanometra, zaostroszony drut metalowy może być kolektorem tzw. prądu tunelowego, który jest mierzalny technicznie. Prąd ten zależy nadzwyczaj silnie od odległości od preparatu, co zapewnia możliwość sterowania wysokością zawieszenia ostrza ponad preparatem. Jeśli dodatkowo proces ten zostanie połączony z odpowiednim przesuwem sondy ponad preparatem, to możliwe będzie uzyskiwanie obrazów powierzchni nawet w skali atomowej (rys. 4).



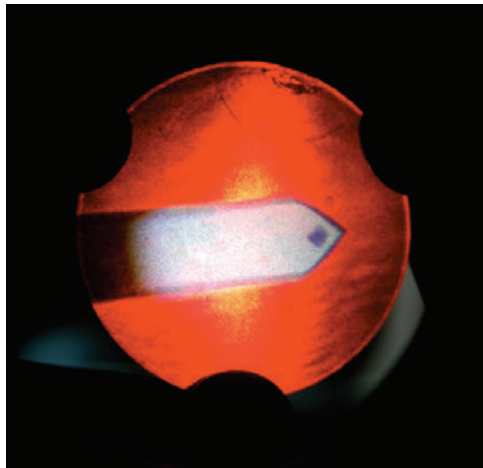
Rys. 4. Obraz efektu moiré rejestrowany na powierzchni grafitu wysokozorientowanego za pomocą skaningowego mikroskopu tunelowego – pole skanowania 12×12 nm, widoczne struktury atomowe grafitu

Fig. 4. Image of moiré pattern recorded on highly oriented graphite surface using scanning tunneling microscope – scan area 12×12 nm, atomic structure of graphite is visible

W 1986 roku Ernst Ruska, Gerd Binnig i Heinrich Rohrer zostali wyróżnieni Nagrodą Nobla w dziedzinie fizyki. Wydaje się, że prace tych badaczy były rzeczywiście

przełomowe dla badań mikro- i nanostruktur. O ich wartości decyduje również to, że dały one początek rozwojowi nowych dziedzin mikro- i nanotechnologii.

Do najważniejszych należą w tym przypadku mikroskopia bliskich oddziaływań, której idea jest obserwacja zjawisk występujących między ostrzową sondą a badanym preparatem. Do najważniejszych jej technik należy m. in. mikroskopia sił atomowych (ang. *Atomic Force Microscopy* – AFM), mikroskopia termiczna (ang. *Scanning Thermal Microscopy* – SThM) czy mikroskopia bliskiego pola optycznego (ang. *Scanning Near-field Optical Microscopy* – SNOM) [3–5]. Prace rozpoczęte przez konstruktorów mikroskopu tunelowego zainicjowały również rozwój technologii tzw. nanonarzędzi (ang. *nanotools*) stosowanych do analizy i syntezy mikro- i nanostruktur. Układy te integrowane razem z układami na podstawie mikroskopii elektronowej i optycznej otwierają absolutnie nowe możliwości obserwacji zachowań mikro- i nanostruktur. Z reguły nanonarzędzia integrują nadzwyczaj małe ostrza, których wymiary sięgają pojedynczych nanometrów. Sondy takie są zbliżane do powierzchni na odległość rzędu pojedynczych stałych krystalograficznych, co zapewnia niejako naturalny dostęp do właściwości badanych mikro- i nanostruktur (rys. 5).



Rys. 5. Widok sondy mikroskopu sił atomowych – struktura sprężysta długości 500 mikrometrów z widocznym na końcówce zarysem nanoostrza;

struktura ta jest przykładem najprostszego nanonarzędzia do syntezy i analizy nanostruktur

Fig. 5. Photograph of an atomic force microscopy probe – a 500- μm long spring structure with the tip outline visible at the end; such a structure is the simplest example of a nanotool for synthesis and analysis of nanostructures

Rozwiązywanie wskazanych wcześniej problemów i postęp w opisanych dziedzinach odbywać się może tylko w ramach interdyscyplinarnej współpracy różnych zespołów badawczych. Obejmuje ona zdefiniowanie problemów, wskazanie i poszukiwa-

nie metod i technik ich rozwiązywania. Wymaga to również wdrożenia odpowiednich programów badawczych i realizacji szeroko zakrojonych inicjatyw badawczych, prowadzonych często w celu umożliwienia komunikacji różnych badaczy.

Opisany na wstępie postęp technologiczny dotyczy również metod i technik obserwacji właściwości cieplnych mikro- i nanostruktur. Konieczność ich zastosowania i wdrażania widoczna jest szczególnie w mikro- i nanoelektronice, gdzie gęstość energii cieplnej rozpraszanej w strukturze osiąga wartości zbliżone do gęstości energii rozpraszanej w reaktorach jądrowych. Konsekwencją postępującej miniaturyzacji i wzrastającego stopnia integracji jest również konieczność określenia zasad transportu ciepła w układach cienkich warstw. Przykładami takich układów są np. systemy wytwarzane w tzw. technologii krzemu na izolatorze (ang. *Silicon On Insulator* – SOI), struktury MOS z ultracienkimi tlenkami bramkowymi o wysokim współczynniku przenikalności dielektrycznej lub struktury połączeń tworzone przez materiały o obniżonym współczynniku dielektrycznym.

Jednocześnie należy zwrócić uwagę, że zjawiska termiczne obserwowane w opisanych systemach są rządzone innymi prawami aniżeli zjawiska termiczne występujące w układach makroskopowych, w których stacjonarny przepływ ciepła jest opisywany prawem Fouriera. W mikro- i nanoukładach przepływ ciepła jest determinowany przez propagowanie i rozpraszanie fononów na granicy ziaren i kryształów, często mając charakter silnie anizotropowy. Opis tego typu zjawisk wymaga uwzględnienia oddziaływania fononów z poszczególnymi molekułami struktury i wymaga ściśle kwantowej interpretacji obserwowanych zjawisk.

Z perspektywy praktyki eksperymentalnej do najbardziej intensywnie rozwijanych metod obserwacji właściwości cieplnych w nanostrukturach należą [6]:

- Pikosekundowa termorefleksja [7], [8];

W metodzie tej impuls laserowy o czasie trwania kilkudziesięciu pikosekund oświetla powierzchnię badanej struktury. Promieniowanie to w miejscu padania wiązki podnosi temperaturę układu i wywołuje propagację fali cieplnej (fononów) w preparacie. Fonony, docierając następnie do kolejnych granic ośrodków, ulegają częściowemu odbiciu, które jest rejestrowane i analizowane za pomocą dodatkowego układu optycznego. Dla większości materiałów stosowanych w technologii mikro- i nanoelektronicznej ich dyfuzyjność D leży w zakresie od 0,005 do 1 cm²/s. W konsekwencji, dla czasu trwania impulsu równego 200 ps, głębokość wnikania fali cieplnej, wyrażana równaniem $l = \sqrt{Dt}$, mieści się w zakresie 10–140 nm. Oświetlanie powierzchni impulsami o dłuższym czasie trwania powoduje zatem, że nie jest możliwe selektywne odwzorowanie przewodności cieplnej poszczególnych warstw tworzących strukturę i obszary między nimi [7]. Ograniczeniem zastosowania tej metody badawczej, jako powszechnej techniki diagnostyki cienkich warstw, są nadzwyczaj duże koszty aparaturowe i złożona struktura systemu. Dodatkowo metoda pikosekundowej termofleksji nie umożliwia dwuwymiarowego odwzorowywania zjawisk transportu ciepła;

- Spektroskopia ramanowska [9–11];
W spektroskopii Ramana wykorzystuje się zjawisko nieelastycznego rozpraszania światła. Próbkę jest w tym przypadku oświetlana monochromatycznym światłem lasera. Część kwantów promieniowania, oddziałując z cząsteczkami próbki, zmienia energię ich oscylacji i przez to zwiększa lub zmniejsza swoją energię. Widma Ramana, czyli zależności intensywności rozproszonego promieniowania od różnicy energii kwantów promieniowania padającego i rozproszonego, zawierają informacje o drganiach charakterystycznych dla badanych cząsteczek. Mogą być wykorzystywane z powodzeniem do identyfikacji i oznaczania zawartości różnych substancji. Nowoczesne spektrometry Ramana umożliwiają wyznaczenie widma próbki przez analizę jej interferogramu, otrzymanego dzięki interferencji promieniowania rozproszonego przez próbkę. Źródłem promieniowania jest zwykle laser Nd:YAG emitujący monochromatyczną wiązkę fotonów o małej energii. Intensywność promieniowania rozproszonego jest o cztery do ośmiu rzędów wielkości mniejsza niż intensywność promieniowania wzbudzającego, aby więc uzyskać mierzalny strumień fotonów rozproszenia ramanowskiego musimy stosować bardzo silne źródła promieniowania do wzbudzania i bardzo czułe układy detekcyjne. Jest to przyczyną wprowadzania do badań laserów, które odznaczają się monochromatycznością i bardzo dużą intensywnością wiązki promieniowania. Istotną wadą spektroskopii Ramana, utrudniającą jej przemysłowe zastosowania, jest duży koszt stosowanych urządzeń i nadzwyczaj złożona struktura aparatury. Warto również wskazać, że spektrometry ramanowskie działają w istocie na zasadzie klasycznych mikroskopów optycznych, co powoduje, że lokalna rozdzielczość obserwacji ograniczona jest klasycznym kryterium Rayleigha. Oznacza to, że w praktyce jest ona rzędu setek nanometrów, co powoduje, że w ujęciu współczesnej technologii nanoelektronicznej ma ona zastosowanie ograniczone.
- Mikroskopia promieniowania podczerwonego [12];
W mikroskopii promieniowania podczerwonego dokonuje się obserwacji promieniowania podczerwonego, jakie jest emitowane z badanej powierzchni. Zgodnie z prawem promieniowania Stefana Boltzmanna, intensywność tego promieniowania jest funkcją współczynnika emisyjności powierzchni i temperatury panującej na niej. Utrudnia to w istotny sposób ilościową analizę rozkładu temperatury panującej na powierzchni preparatu, która jest możliwa po przeprowadzeniu złożonej procedury kalibracji systemu dalekiego pola [12]. Mikroskopia promieniowania podczerwonego, nazywana również mikroskopią termiczną dalekiego pola, jest jednak sprawnym i powszechnie stosowanym narzędziem do diagnostyki typowych układów MEMS. Jej ograniczona zdolność rozdzielcza (rzędu pojedynczych mikrometrów) wyklucza zastosowanie mikroskopii termicznej dalekiego pola do badań nanostruktur. Podobnie jak w przypadku spektroskopii ramanowskiej, lokalna rozdzielczość obserwacji ograniczona jest kryterium Rayleigha.

Mikroskopia promieniowania podczerwonego, która ma ogromne zastosowanie w badaniach struktur makroskopowych, ma zatem w technice mikrosystemów i nanotechnologii znaczenie ograniczone.

- Mikroskopia bliskiego pola termicznego [4, 5];
Mikroskopia bliskiego pola termicznego jest jedną z metod pomiarowych mikroskopii bliskich oddziaływań. Mikroostrze pomiarowe pełni w tym przypadku rolę czujnika temperatury panującej na powierzchni (tryb statyczny mikroskopii termicznej) lub rolę źródła ciepła, które lokalnie ogrzewa strukturę (tryb aktywny mikroskopii termicznej). W odróżnieniu od metod termorefleksyjnej pikosekundowej i spektroskopii ramanowskiej mikroskopia bliskiego pola termicznego ułatwia lokalne obserwacje zjawisk termicznych w mikro- i nanostrukturach. Ocenę właściwości termicznych (cieplnych) w głąb struktury umożliwia natomiast aktywna mikroskopia termiczna. Jak zostanie pokazane w dalszej części pracy, moc konieczna do utrzymania temperatury ostrza na zadanym poziomie odwzorowuje wartość rezystancji cieplnej bezpośrednio pod ostrzem pomiarowym. Zastosowanie w mikroskopii termicznej zmiennoczęstotliwościowej modulacji mocy cieplnej dostarczanej do mikroostrza umożliwi ocenę przewodności cieplnej poszczególnych warstw tworzących badaną strukturę. Modulacja częstotliwości mocy dostarczanej do mikroostrza umożliwia bowiem ustalenie głębokości penetrowania układu przez falę cieplną, której źródłem jest mikroostrze. W metodzie tej, która jest nazywana również metodą 3ω i która była stosowana w badaniach układów cienkowarstwowych o dużej powierzchni, metalowy grzejnik zasilany jest prądem o pulsacji ω [13]. Moc cieplna wydzielana w nim moduluje zatem temperaturę grzejnika z pulsacją 2ω . Temperaturę i właściwości termiczne warstw w otoczeniu bezpośrednio pod grzejnikiem można oceniać, obserwując napięcie elektryczne o częstotliwości kołowej 3ω , które odkłada się na oporniku zasilanym prądem grzewczym. Pomiar zmiennoczęstotliwościowy stosowano w diagnostyce dielektrycznych warstw cienkich (grubości do 300 nm) [14, 15]. Modyfikacja kształtu mikrogrzejnika osadzonego na badanej strukturze umożliwiła również pomiary przewodności cieplnej warstw polimerowych [16], struktur supersieci Si/SiGe i SiGe/SiGe [17] oraz cienkich warstw krzemu w strukturze tzw. krzemu na izolatorze [18]. Zastosowanie metody 3ω w skaningowej mikroskopii termicznej bliskiego pola wiąże się z koniecznością przeprowadzenia pełnego opisu transportu ciepła od ogrzewanej mikrosondy do badanej struktury. W przypadku badań prowadzonych w powietrzu dotyczy to zwłaszcza znajomości zasad wnikania ciepła przez warstwę cieczy (wody) adsorbowanej na obserwowanej powierzchni. W rozważaniach tych należy również uwzględnić ciepło przenoszone przez dźwignię sprężystą w kierunku nasady czujnika. Metoda zmiennoczęstotliwościowej modulacji mocy rozpraszanej w mikroostrzu umożliwiła odwzorowanie właściwości termicznych cienkich warstw polipropylenu osadzonego na podłożu

szafirowym [19]. Na rozdzielczość obserwacji wpływają również właściwości ostrza pomiarowego, które powinno się charakteryzować jak najmniejszymi rozmiarami i małą pojemnością cieplną [20]. Czułość czujnika temperatury, który jest zintegrowany z ostrzem, powinna umożliwiać rejestrację zmian temperatury o wartości około 10 mK. Na możliwości pomiarowe rzutują również właściwości belki sprężystej, której konstrukcja musi ograniczać przewodzenie ciepła od ostrza pomiarowego przez dźwignię do nasady czujnika.

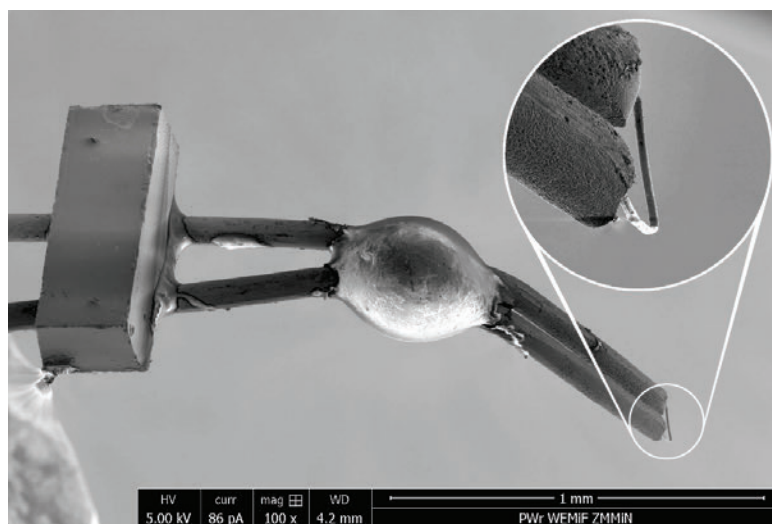
1. OSTRZOWE CZUJNIKI DLA SKANINGOWEJ MIKROSKOPII BLISKIEGO POŁA TERMICZNEGO

W skaningowej mikroskopii termicznej stosowane są czujniki termiczne zintegrowane z ostrzem pomiarowym wykonywane w postaci bądź złączy termoelektrycznych, bądź termorezystorów.

Czujniki wykonywane w postaci złączy termoelektrycznych, pełniące rolę czujników temperatury panującej na powierzchni, mogły być zasadniczo stosowane jedynie w pasywnej mikroskopii termicznej. Wymuszenie przepływu ciepła od ostrza wykonanego w postaci termozłącza do badanej struktury wymaga zastosowania złożonego układu, który oświetla końcówkę włókna sprężystego wiązką promieniowania laserowego [21]. Pierwsze czujniki przepływu ciepła między mikroostrzem a powierzchnią były wykonane w postaci termoelementu osadzonego na końcówce ostrza wolframowego typowego dla skaningowej mikroskopii tunelowej [4]. Wadą takiego rozwiązania była jednak bardzo mała czułość sondy wynosząca około $3 \mu\text{V/K}$ i kłopotliwe wytwarzanie czujników o powtarzalnych parametrach. Zastosowanie technologii mikroelektronicznych do budowy ostrzowych sensorów przepływu ciepła skaningowego mikroskopu termicznego bliskiego pola umożliwiło skonstruowanie systemów o zdecydowanie lepszych właściwościach metrologicznych [22, 23]. W opisywanych rozwiązaniach na ostrzu dźwigni sprężystej, stosowanej w statycznych pomiarach mikroskopii sił atomowych wykonanej z Si_3N_4 , wytworzono strukturę mikrozłącza dwóch metali: Au i Ni. Czułość detekcji zmian temperatury była w tym przypadku większa aniżeli w przypadku konstrukcji włóknowych i wynosiła $14 \mu\text{V/K}$ [24]. Warto w tym miejscu zwrócić uwagę na zastosowanie Si_3N_4 jako materiał ustroju mechanicznego czujnika mikroskopu termicznego. Jest to materiał o bardzo małej przewodności cieplnej, co zapewnia, że ciepło z ostrza nie jest odprowadzane do nasady czujnika. Istotną wadą opisywanych czujników była niepowtarzalność właściwości złącza termoelektrycznego wytworzonego w obszarze ostrza pomiarowego. Do eksperymentów najbardziej spektakularnie pokazujących możliwości badawcze skaningowej mikroskopii termicznej bliskiego pola z cienkowarstwowymi czujnikami temperatury należą pomiary transportu ciepła przez wielościennie nanorurki węglowe [25]. W tym przypadku zaobserwowano

wzdłuż pojedynczej nanorurki długości 4 μm i grubości 50 nm oraz przy zasilaniu struktury prądem 77 μA wzrost temperatury o 20 K [26].

W przypadku czujników wykonywanych w postaci termorezystorów sondą termiczną jest mikropętla oporowa. Sonda taka, stykając się z badaną powierzchnią, może służyć jako czujnik temperatury panującej na powierzchni lub jako mikroźródło ciepła. W ujęciu historycznym pierwszymi sondami termooporowymi były czujniki wykonane z drutu Wollastona [27] i nanopętla platynowe [28, 29]. W przypadku sond Wollastona czujnikiem termicznym jest pętla oporowa wykonana z platynowego włókna średnicy około 10 μm (rys. 6). Włókno to, pokryte dodatkowo otuliną ze srebra, stanowi część sprężystą belki mikroskopu bliskiego pola termicznego. Długość dźwigni sprężystej typowej sondy Wollastona wynosi około 4 mm, stała sprężystości jest równa około 5 N/m, natomiast rezystancja końcówki oporowej wynosi około 3 Ω . W trybie aktywnym sondy Wollastona mogą być rozgrzane do temperatury około 700 K.

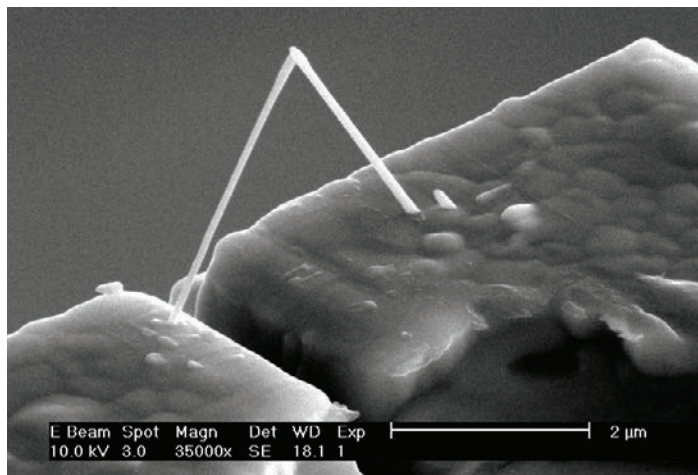


Rys. 6. Skaningowa mikroskopia bliskiego pola termicznego – termooporowa sonda Wollastona [27]

Fig. 6. Near-field scanning thermal microscopy – a thermoresistive Wollaston probe [27]

Zasadniczą wadą opisywanego rozwiązania są mała wartość częstotliwości rezonansowej dźwigni i duże wymiary końcówki mikrosondy termicznej. Ogranicza to w istotny sposób możliwości wysokorozdzielczych pomiarów topografii powierzchni i właściwości termicznych nanostruktur wielkości mniejszej niż 300 nm. Sondy te charakteryzują się, z racji bardzo dużych wymiarów, małymi częstotliwościami rezonansowymi. W konsekwencji mikroskop wyposażony w taką sondę staje się nadzwyczaj wrażliwy na zakłócenia, dodatkowo szybkość skanowania jest niewielka, co powoduje, że czas obra-

zowania jest nadzwyczaj wydłużony. W pewnym sensie przeciwieństwem wręcz makroskopowych układów sond Wollastona są konstrukcje wykonane z kompozytu osadzanego za pomocą wiązki elektronowej (ang. *Electron Beam Induced Deposition* – EBID; rys. 7). W metodzie tej molekuły prekursora, które poruszają się w komorze mikroskopu, rozpadają się na skutek zderzeń z wiązką elektronów. Sterując miejscem ekspozycji i koncentracją prekursora można wytwarzać w ten sposób struktury trójwymiarowe o nanometrycznych rozmiarach, stosowane np. w mikroelektronice próżniowej [30]. Gazem nośnym są zwykle mieszaniny związków metaloorganicznych. Warstwy tą metodą osadzone zawierają wtrącenia krystaliczne wbudowane w węglową macierz amorficzną. Średnica wytworzonego za pomocą osadzania wiązką elektronową włókna nanopętli platynowej wynosi około 70 nm, wysokość czujnika równa jest zaś około 3 μm (rys. 7). Technologię nanopętli platynowych przedstawiono na rys. 7 [29]. Czujniki te osadzano na dźwigniach piezorezystywnych z cienkowarstwowym aluminiowym mikrogrzejnikiem (rys. 8a). Za pomocą zogniskowanej wiązki jonowej (ang. *Focused Ion Beam* – FIB) wytrawiano aluminiową ścieżkę mikrogrzejnika na końcówce belki (rys. 8b). Następnie dzięki osadzeniu za pomocą wiązki elektronów wytworzono czujnik oporowy połączony z obydwojma doprowadzeniami aluminiowymi. Doprowadzenia te umożliwiają spolaryzowanie czujnika i jego podłączenie do elektronicznych układów pomiarowych. Widoczne na rysunku 8c przedłużenie pętli ułatwia kontakt czujnika z powierzchnią badanej struktury. Istotną wadą opisywanej konstrukcji jest praktycznie ręczny sposób jej wytwarzania. Bardzo mała wysokość ostrza powoduje również, że wprowadzanie ostrza w kontakt z badaną powierzchnią jest utrudnione i wymaga dużej wprawy. Rozwinięciem opisywanych konstrukcji są konstrukcje dźwigni z oporowymi grzejnikami platynowymi [31]. Struktury takie są w formie konstrukcji, której ustrój mechaniczny tworzy mikrobeltka wykonana z Si_3N_4 . Zapewnia to odpowiednio dużą rezystancję cieplną platformy mechanicznej. Na uwagę zasługuje konstrukcja samego ostrza termicznego, które tworzy tylko warstwa metalizacji platynowej (w rozwiązaniach komercyjnych ostrze termiczne jest osadzone na wsporniku materiału formującego część sprężystą belki). W ten sposób zmniejsza się masę końcówki ostrza, dzięki czemu zmniejsza się stała czasowa układu oraz energia, jaką pobiera ostrze w kontakcie z mierzoną strukturą. Doprowadzenia do sondy oporowej mogą być wykonane w układzie dwu- lub czteropunktowym, dzięki czemu obserwacja zmian rezystancji końcówki może być prowadzona z większą wiarygodnością. W strukturze platformy mechanicznej zastosowano wycięcia struktury, aby dźwignia mikromechaniczna miała możliwie jak najmniejszą sztywność i jak największą rezystancję termiczną. Warto wskazać, że kompromis taki można osiągnąć w wypadku dźwigni o możliwie małej grubości. Sztywność belek o przekroju prostokątnym i ich rezystancja termiczna są odpowiednio odwrotnie proporcjonalne do grubości struktury w trzeciej i pierwszej potęgze oraz odwrotnie proporcjonalne do jej szerokości. Końcówka ostrza została dodatkowo wycięta za pomocą trawienia jonowego prowadzonego zogniskowaną wiązką jonów. W ten sposób zwiększono rezystancję samej końcówki, tak aby poprawić geometrię końcówki i jednocześnie skupić na niej wydzielanie ciepła.



Rys. 7. Skanینگowa mikroskopia bliskiego pola termicznego – termooporowa nanopętla platynowa
 Fig. 7. Near-field scanning thermal microscopy – a thermoresistive platinum loop

Wychylenie końcówki ostrza dokonywane jest w układzie za pomocą detektora optycznego typowego dla większości układów mikroskopii sił atomowych. Dla poprawy odbicia wiązki laserowej zastosowano metaliczne zwierciadło cienkowarstwowe. Opisywana metalizacja służy również do odprowadzania do nasady belki ciepła wprowadzanego laserowy układ optyczny. W ten sposób zmniejszeniu ulega efekt ogrzewania końcówki sondy.

2. PODSTAWY SKANINGOWEJ MIKROSKOPII BLISKIEGO POLA TERMICZNEGO

Przewodność cieplną K ciała stałego najłatwiej jest zdefiniować w warunkach ustalonego przepływu ciepła przez długi pręt, wzdłuż którego jest obserwowany gradient temperatury dT/dx [32]:

$$\dot{q} = -K \frac{dT}{dx} \quad (1)$$

gdzie \dot{q} jest strumieniem ciepła przepływającym przez jednostkę powierzchni w jednostce czasu. Na podstawie równania (1) można wskazać, że przewodzenie ciepła jest procesem stochastycznym, który zależy od gradientu temperatury (a nie tylko od różnicy temperatury między końcami próbki). Ponadto energia cieplna nie jest przenoszona wzdłuż jednej tylko osi z jednego krańca próbki na drugi. Nośniki ciepła, dyfundując bowiem przez preparat, ulegają licznym zderzeniom. Na wartość przewodno-

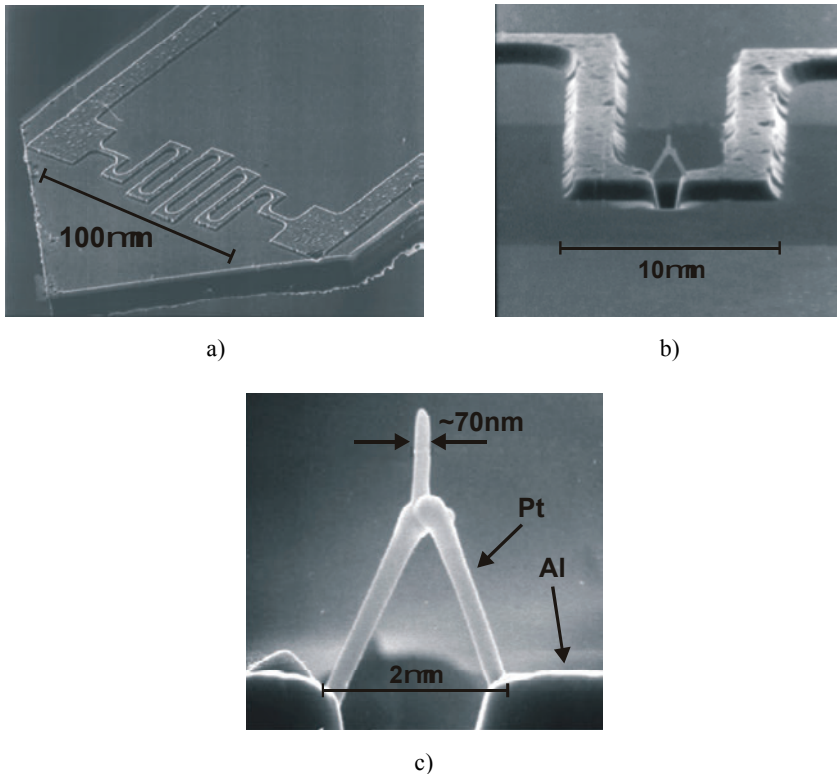
ści cieplnej musi mieć zatem również wpływ średnia droga swobodna nośników energii cieplnej. Na podstawie elementarnej teorii kinetycznej gazów można wykazać, że przewodność cieplna K wyraża się zależnością:

$$K = \frac{1}{3} C v l \quad (2)$$

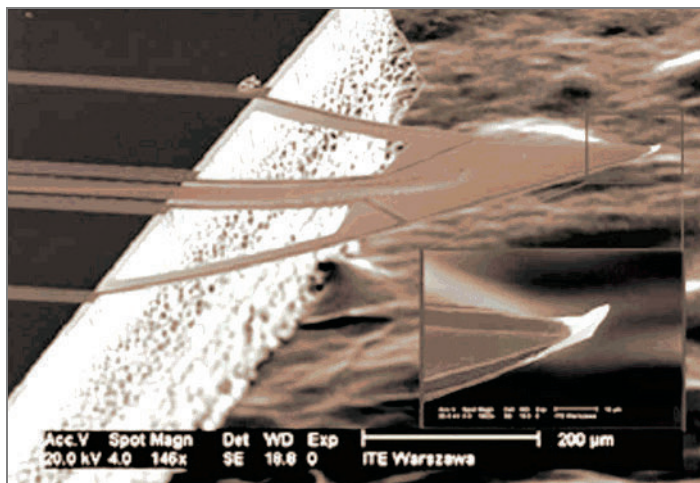
gdzie C jest pojemnością cieplną próbki, v – średnią prędkością cząstki, l – średnią drogą swobodną między zderzeniami.

Należy jednak zauważyć, że lokalna rozdzielczość skaningowej mikroskopii bliskich oddziaływań zależy od trzech zasadniczych czynników:

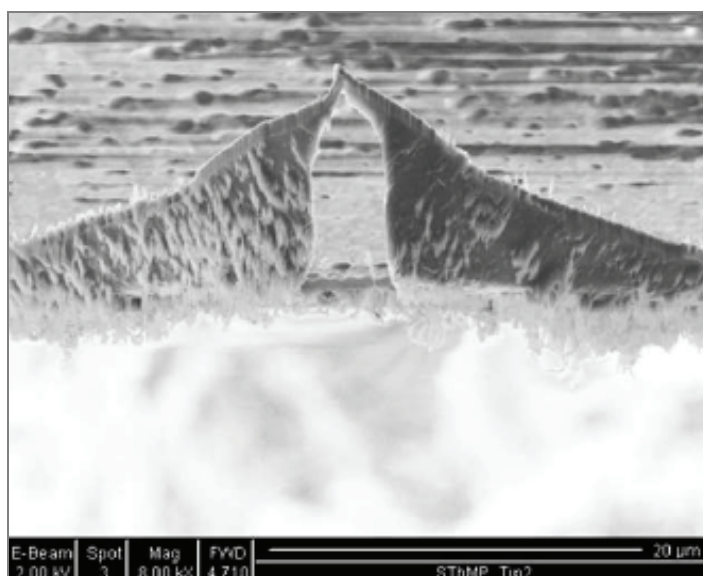
- promienia zakończenia końcówki mikros sondy termicznej,
- właściwości kontaktu termicznego ostrza z badaną powierzchnią,
- właściwości termicznych całej sondy pomiarowej.



Rys. 8. Wytwarzanie nanopętli platynowych tworzących ostrze mikroskopu bliskiego pola termicznego metodą osadzania wspomaganego wiązką elektronową
 Fig. 8. Fabrication of platinum nanoloops as scanning thermal microscopy tips using electron beam induced deposition



Rys. 9. Belka mikroskopu bliskiego pola termicznego o zoptymalizowanej konstrukcji
 Fig. 9. Cantilever of the optimally-designed scanning thermal microscopy probe



Rys. 10. Ostrze sondy mikroskopu bliskiego pola termicznego o zoptymalizowanej konstrukcji
 Fig. 10. Tip of the optimally-designed scanning thermal microscopy probe

Dla typowej konfiguracji pomiarowej spotykanej w skaningowej mikroskopii termicznej, w której obserwowane są właściwości dwuwarstwowego układu, obowiązują następujące równania:

$$T_{w1} - T_{w2} = \dot{q} \frac{\sigma_1}{K_1} \quad (3)$$

oraz

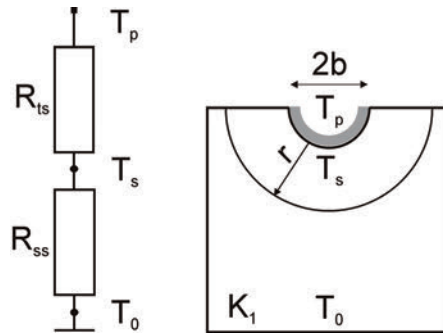
$$T_{w2} - T_{w3} = \dot{q} \frac{\sigma_2}{K_2}, \quad (4)$$

gdzie T_{w1} , T_{w2} , T_{w3} są temperaturami panującymi na poszczególnych przegrodach, natomiast δ_1 , δ_2 i K_1 , K_2 są odpowiednio grubościami i przewodnościami cieplnymi warstw układu.

W aktywnej mikroskopii bliskiego pola termicznego łatwo jest opisywać ostrze pomiarowe jako punktowe źródło ciepła o średnicy $2b$, którego temperatura wynosi T_p (rys. 11). Ciepło z tego źródła jest transportowane do wnętrza próbki przez układ połączonych szeregowo rezystancji cieplnej warstwy kontaktowej między ostrzem a powierzchnią R_{ts} i rezystancji cieplnej, jaką przedstawia sobą badana próbka R_{ss} . Rezystancję R_{ss} można obliczyć rozważając transport ciepła przez wnętrze preparatu. Dla półsfery o promieniu r transport ten jest opisany równaniem

$$\frac{(T_s - T_0)}{R_{ss}} = -2\pi r^2 K_1 \frac{dT}{dr} = P_s \quad (5)$$

gdzie K_1 jest przewodnością cieplną preparatu, P_s – mocą dostarczaną do ostrza, T_0 jest temperaturą odniesienia próbki, natomiast T_s – temperaturą na powierzchni granicznej próbki pod sondą pomiarową.



Rys. 11. Aktywna mikroskopia bliskiego pola termicznego – pomiar przewodności cieplnej preparatu

Fig. 11. Active-mode scanning thermal microscopy
– measurement of the thermal conductivity of the specimen

Po rozdzieleniu zmiennych i scałkowaniu otrzymujemy

$$K_1 [T]_{T_s}^{T_0} = \frac{P_s}{2\pi} \left[\frac{1}{r} \right]_b^\infty \quad (6)$$

Po przekształceniach otrzymuje się zatem

$$P_s = -2\pi K_1 b (T_0 - T_s) \quad (7)$$

Po uwzględnieniu równania (5) i przy założeniu, że $T_s \gg T_0$ otrzymuje się zależność

$$P_s = \frac{(T_s - T_0)}{R_{ss}} = 2\pi K_1 b T_s \quad (8)$$

Rezystancję R_{ss} opisać zatem można zależnością

$$R_{ss} = \frac{1}{2\pi K_1 b} \quad (9)$$

Transport ciepła przez warstwę kontaktową między sondą a powierzchnią do wnętrza próbki i po założeniu, że $T_s \gg T_0$ i po uwzględnieniu równania (9) opisany jest zatem zależnością:

$$P_s = \frac{(T_p - T_0)}{R_{ts} + R_{ss}} = \frac{(T_p - T_0)}{R_{ss}(1 + R_{ts}/R_{ss})} \approx \frac{G_{ss} T_p}{(1 + G_{ss}/G_{ts})} = \frac{2\pi K_1 b T_p}{1 + (2\pi K_1 b / G_{ts})} \quad (10)$$

gdzie G_{ts} jest konduktancją cieplną warstwy kontaktowej między ostrzem a próbką.

Równanie (10) jest podstawowym równaniem aktywnej mikroskopii bliskiego pola termicznego. Analiza równania (10) pokazuje, że:

- moc P_s dostarczana do ostrza jest odwrotnie proporcjonalna do wartości rezystancji cieplnej, jaka jest widziana pod sondą pomiarową;
- moc P_s dostarczana do ostrza jest proporcjonalna do wielkości pola kontaktu mikros sondy termicznej z powierzchnią preparatu i przewodności cieplnej próbki K_1 . Należy jednak zauważyć, że uwaga powyższa jest prawdziwa jedynie w przypadku, gdy $2\pi K_1 b / G_{ts} \ll 1$. W praktyce oznacza to, że wyznaczanie przewodności cieplnej materiału na podstawie obserwacji stałoprądowego rozplywu mocy rozpraszanej w ostrzu mikros sondy termicznej jest możliwe jedynie w przypadku próbek o stosunkowo małej przewodności cieplnej. W celu ilustracji i po założeniu, że średnica kontaktu odpowiada 50 nm, a typowa konduktancja cieplna kontaktu wynosi 10^{-6} W/K, przewodność cieplna, dla której stosunek $2\pi K_1 b / G_{ts}$ wyniesie 0,1 – będzie równa około 30 W/(mK). Jest to wartość przewodności cieplnej odpowiadająca materiałom polimerowym. Charakterystyczna jest w tym przypadku analogia związana z poprawnym pomiarem siły elektromotorycznej źródła napięcia elektrycznego. W idealnym przypadku w oczku prądowym, jaki tworzą źródło napięcia, impedancja wejściowa miernika i rezystancja wewnętrzna źródła napięcia, nie powinien płynąć prąd elektryczny. W realizacji praktycznej takiego pomiaru jego jakość będzie się poprawiać, gdy stosunek impedancji wejściowej woltomierza do rezystancji źródła napięciowego będzie wzrastać.

W przypadku statycznej mikroskopii bliskiego pola termicznego pomiar temperatury za pomocą mikrosondy oporowej może być modelowany układem przedstawionym na rysunku 12. Strumień ciepła przenoszony przez układ szeregowo połączonych rezystancji cieplnych kontaktu powierzchni z sondą, ostrza, dźwigni wynosi

$$j_u = \frac{T_s - T_a}{R_{ts} + R_t + R_c} \quad (11)$$

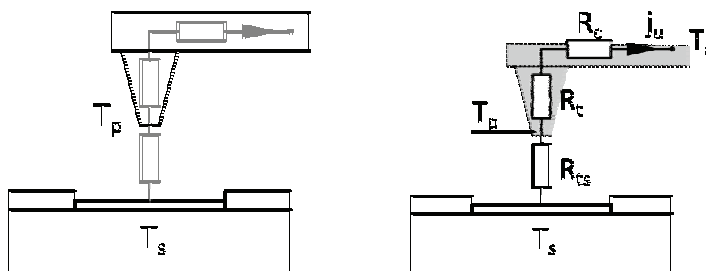
gdzie T_s jest temperaturą powierzchni, T_p – temperaturą sondy, T_a – temperaturą uchwytu sondy, R_t – rezystancją cieplną ostrza pomiarowego, R_c – rezystancją cieplną belki wspornikowej czujnika. Temperaturę ostrza T_p można zatem obliczyć na podstawie równania

$$T_p = T_s - \frac{R_{ts}}{R_{ts} + R_t + R_c} (T_s - T_a) = T_s - \frac{1}{1 + R_t / R_{ts} + R_c / R_{ts}} (T_s - T_a) \quad (12)$$

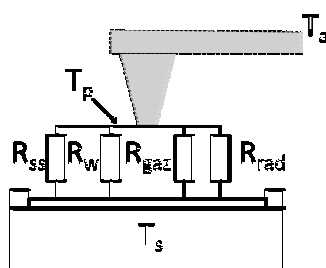
Z zależności (12) wynika, że aby temperatura na ostrzu dźwigni T_p odpowiadała temperaturze próbki T_s (innymi słowy pomiar był prowadzony w sposób poprawny), mianownik równania drugiego członu równania (12) powinien przyjmować możliwie największe wartości. Można to osiągnąć, gdy:

- Stosunek rezystancji R_c do rezystancji R_{ts} jest możliwie jak największy. Składowymi rezystancji cieplnej kontaktu R_{ts} są następujące rezystancje: kontaktu mechanicznego ostrza z powierzchnią R_{ss} , warstwy zaadsorbowanej wody w obszarze sondy R_w , związana z radiacją ciepła między powierzchnią a sondą R_{rad} i cieplna warstwy gazu znajdującego się między ostrzem a powierzchnią R_{gaz} (rys. 13). Z analiz przedstawionych w pracy [22] wynika, że decydujący wpływ na transport ciepła w układzie ostrze–powierzchnia ma warstwa wody zaadsorbowanej na powierzchni preparatu (tab. 1). Modyfikując skład chemiczny atmosfery, w której przeprowadzane są pomiary, można wpływać na wartość tej rezystancji cieplnej i jednocześnie podnieść dokładność pomiaru temperatury panującej na powierzchni próbki.
- Temperatura zamocowania belki T_a powinna być zbliżona lub w idealnym przypadku równa temperaturze T_s panującej na powierzchni. Realizacja tego założenia jest utrudniona, wymaga bowiem integracji z dźwignią mikrogrzejnika, który w układzie regulacji powinien zmieniać swoją temperaturę tak, aby minimalizować przepływ ciepła j_u .
- Rezystancja cieplna dźwigni R_c jest możliwie jak największa. Realizacja tego założenia wymaga wytrawienia w belce wycięć, które przez zmniejszenie przekroju dźwigni działają jak ekrany cieplne. Innym rozwiązaniem jest wykonanie belki z materiału o możliwie małym przewodnictwie cieplnym. Warto wskazać,

że do materiałów najbardziej predystynowanych do tego celu należy azotek krzemu Si_3N_4 , gdyż jego przewodność cieplna jest około dziesięciokrotnie mniejsza niż przewodność cieplna krzemu. Należy jednak zwrócić uwagę, że jednocześnie należy zadbać o odpowiednią charakterystykę mechaniczną sondy, która pełni rolę czujnika sił atomowych.



Rys. 12. Statyczna mikroskopia bliskiego pola termicznego
Fig. 12: Static scanning thermal microscopy



Rys. 13. Statyczna mikroskopia bliskiego pola termicznego
– transport ciepła między sondą termiczną a badaną powierzchnią
Fig. 13. Static scanning thermal microscopy
– heat transfer between the thermal probe and the sample

Tabela 1. Rezystancje cieplne kontaktu ostrza i powierzchni w mikroskopii termicznej bliskiego pola [22]
Table 1. Thermal resistances in tip-sample contact in scanning thermal microscopy [22]

Składowa rezystancji cieplnej między ostrzem a powierzchnią	Rezystancja cieplna [K/W]
R_{ss}	10^6
R_w	10^5
R_{rad}	10^9
R_{gaz}	10^7

3. MODEL TRANSPORTU CIEPŁA MIĘDZY OSTRZEM A POWIERZCHNIĄ W SKANINGOWEJ MIKROSKOPII TERMICZNEJ BLISKIEGO POLA

Przedstawiona w poprzednim rozdziale analiza podstaw pasywnej i aktywnej mikroskopii bliskiego pola termicznego pokazuje, że na metrologiczne (innymi słowy ilościowe) obrazowanie właściwości cieplnych nanostruktur istotny wpływ ma kształt ostrza. Nie jest ono zadaniem trywialnym, uwzględniając szczególnie, że promień końcówki jest rzędu pojedynczych dziesiątków nanometrów. Ocena tak małych szczegółów wymaga zastosowania wysokorozdzielczych mikroskopów elektronowych, w których wiązka elektronów dla zapobieżenia uszkodzeniu końcówki powinna być przyspieszana możliwie niewielkim napięciem (co często stoi w sprzeczności z wymogiem uzyskania odpowiednio dużej zdolności rozdzielczej). Problem staje się jeszcze bardziej złożony, jeśli chodzi o ocenę kształtu kontaktu termicznego między ostrzem pomiarowym a powierzchnią, którego wielkość może odbiegać od kształtu geometrycznego końcówki sondy.

Pomimo że problem modelowania termicznego kontaktu między sondą a badaną powierzchnią jest krytyczny dla ilościowych pomiarów rozkładu przewodności cieplnej, to opublikowano na ten temat zaledwie kilka prac. Gomès i in. [33] badali udział różnych mechanizmów transportu ciepła przy stałym prądzie grzania mikroostrza. Zaobserwowali, że w powietrzu, przy temperaturach do 90 °C, dominuje transport ciepła przez menisk wodny wytworzony przez cienką warstwę wody zawsze obecnej na badanej powierzchni. Powyżej tej temperatury w zależności od przewodności cieplnej próbki dominuje kontakt ciało stałe–ciało stałe (dla próbek o dużej przewodności) lub ciało stałe–gaz–ciało stałe (dla próbek o małej przewodności cieplnej). Problem ilościowego pomiaru rozkładu przewodności cieplnej został poruszony przez Lefèvre i in. w pracy [34]. Pokazali, że sonda pomiarowa nie może być traktowana jako układ izotermiczny i zaproponowali kombinację metod numerycznych z pomiarami SThM do określenia powierzchni kontaktu termicznego między sondą a powierzchnią próbki. Ponadto Lefèvre i Volz w [35] pokazali, że w trybie zmiennoprądowym (mikrogrzejnik jest zasilony prądem przemiennym), zwłaszcza w tzw. trybie 3ω , znacznie zwiększa się czułość pomiarów SThM, a ich wyniki są niezależne od temperatury powietrza. W pracy tej pokazano również, że możliwe jest wyznaczenie promienia krzywizny sondy zanim zostanie ona zbliżona do powierzchni.

Cechą wspólną tych prac jest model kontaktu między sondą a powierzchnią, który upraszcza kształt sondy i topografię powierzchni (zwykle sferyczne lub walcowe ostrze i płaska próbka). Jednak skala rozmiarów współczesnych sond jest tak mała, że nie można uznać, iż badana powierzchnia jest płaska. Powoduje to, że

na właściwości kontaktu cieplnego mają znaczny wpływ kształt sondy i topografii badanej powierzchni. W pracy [36] pokazano, że do analizy przepływu ciepła między wbudowanym mikrogrzejnikiem a badaną próbką można wykorzystać numeryczne modelowanie metodą elementów skończonych. Jednakże niezbędna jest do tego precyzyjna znajomość konstrukcji takiej sondy. Metoda ta zatem może być wykorzystana przez producentów sond SThM, którzy znają dokładnie konstrukcję, parametry procesów technologicznych i parametry wykorzystanych materiałów.

W pracy [37] Józwiaka i in. zaproponowano metodę TThM (ang. *Tip Thermal Mapping* – TThM) mapowania właściwości cieplnych sond SThM. Umożliwia ona obliczenie rozkładu powierzchniowego mocy rozpraszanej przez sondę SThM na podstawie połączonych pomiarów AFM i SThM powierzchni silnie przewodzącego materiału o dużej chropowatości. Taka mapa może służyć nie tylko do oceny jakości właściwości termicznych sondy czy oceny stopnia jej degradacji w trakcie eksploatacji, ale również jest pierwszym krokiem w kierunku usunięcia wpływu topografii powierzchni z wyników pomiarów SThM.

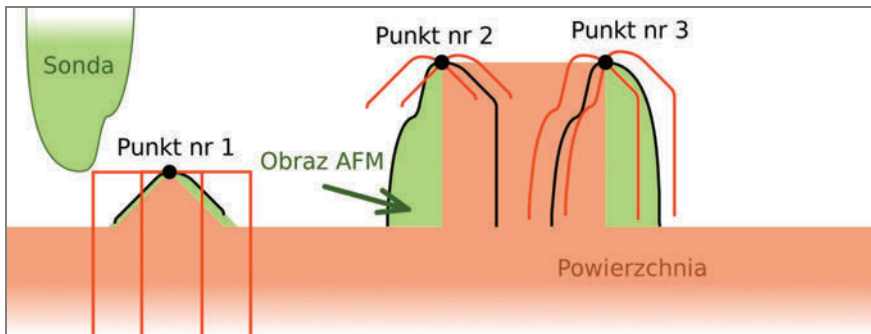
4. MAPOWANIE TERMICZNYCH WŁAŚCIWOŚCI OSTRZA

Metoda TThM wykorzystuje zaawansowane techniki przetwarzania sygnałów do analizy trzech obrazów powierzchni: obrazu AFM otrzymanego za pomocą ostrza o bardzo małym promieniu krzywizny oraz obrazów AFM i A-SThM zarejestrowanych jednocześnie za pomocą badanej sondy SThM. Do obliczenia kształtu sondy SThM wykorzystano metodą RBTR (ang. *Regularized Blind Tip Reconstruction*) zaproponowaną przez Józwiaka i in. w pracy [38]. Kształt sondy i model kontaktu sonda–próbka ułatwiają sformułowanie zagadnienia odwrotnego umożliwiającego obliczenie rozkładu rozpraszania mocy na powierzchni sondy SThM. Dalej przedstawiono szczegółowy opis tego podejścia.

Ślepa rekonstrukcja kształtu ostrza

Metoda ślepej rekonstrukcji ostrza (ang. *Blind Tip Reconstruction* – BTR) umożliwia wyznaczenie kształtu ostrza sondy SPM na podstawie analizy obrazu struktury kalibracyjnej, dla której nie znamy dokładnych wymiarów geometrycznych (kątowności, promieni krzywizny) [39]. Metoda umożliwia obliczenie kształtu największego ostrza, jakie mogło zostać użyte do rejestracji analizowanego obrazu SPM. Jeżeli struktura kalibracyjna składa się z elementów o odpowiednio dużej krzywiznie powierzchni, to takie górne ograniczenie na kształt ostrza będzie zbieżne do jego rzeczywistego kształtu.

Na rysunku 14 zaprezentowano wynik symulacji komputerowej, za pomocą której pokazano zasadę działania metody BTR. Dla przejrzystości prezentacji tego zagadnienia przyjęto, że wynikiem pomiaru AFM jest dwuwymiarowa krzywa (kolor zielony), a nie trójwymiarowa powierzchnia. Taki model można w prosty sposób uogólnić na przypadek trójwymiarowy. Pokazana krzywa AFM została uzyskana w wyniku morfologicznej dylatacji kształtu rzeczywistej dwuwymiarowej krzywej (kolor różowy) powierzchni (w rzeczywistym przypadku powierzchnia ta jest oczywiście obiektem trójwymiarowym) za pomocą krzywej kształtu ostrza (kolor zielony) odbitego względem punktu będącego wierzchołkiem ostrza.



Rys. 14. Ilustracja metody ślepej rekonstrukcji kształtu ostrza AFM [38]

Fig. 14. Illustration of blind AFM tip reconstruction method [38]

W ten sposób zamodelowany jest pomiar za pomocą skanującej sondy AFM. Operacja morfologicznej dylatacji może być porównana do malowania po krzywej pędzlem o zadanym kształcie. W metodzie ślepej rekonstrukcji musimy ten proces odwrócić. Na początku zakładamy pewien określony kształt ostrza (zwykle walec – w naszym przypadku prostokąt), który będzie przez procedurę w kolejnych iteracjach modyfikowany. W trakcie tej modyfikacji przyjmujemy, że wierzchołek ostrza jest położony w ściśle określonym miejscu (zwykle środek krzywej/powierzchni modelu ostrza). Ponieważ jest to wierzchołek ostrza, nigdy nie może on wejść poniżej poziomu krzywej/powierzchni. Oznacza to, że w trakcie otrzymywania obrazu AFM za pomocą morfologicznej dylatacji maksimum krzywej/powierzchni odbitego ostrza nie może znajdować się powyżej krzywej/powierzchni obrazu AFM. Dlatego w każdym punkcie obrazu zastanawiamy się jak ten punkt powstał, „przykładając” bieżący kształt ostrza do tego punktu i odrzucając wszystkie te przypadki, w których wierzchołek znajduje się ponad obrazem AFM. Po każdym takim „przyłożeniu” „odcinamy” z ostrza te części, które wystają ponad obraz, otrzymując w ten sposób serię potencjalnych kształtów odbicia ostrza. Ponieważ nie ma możliwości określenia, który z tych kształtów jest

prawdziwy, wykonuje się na nich operację sumy zbiorów, przez co wyznacza się największe ostrze, które mogło dać analizowany punkt krzywej/powierzchni obrazu AFM. Otrzymany kształt porównuje się z kształtem bieżącym, biorąc ich część wspólną. W następnym kroku przechodzimy do kolejnego punktu obrazu AFM i analiza jest powtarzana z użyciem nowego bieżącego kształtu ostrza. Na rysunku 15 pokazano przebieg takiej analizy w trzech punktach obrazu. Za pomocą linii czerwonych pokazano położenia odbitego ostrza, dla których wierzchołek znajduje się powyżej obrazu AFM (a więc te położenia nie są uwzględniane w analizie). Za pomocą linii czarnych pokazano te położenia, które są wykorzystywane w analizie. Można zauważyć jak kształt obrazu AFM umożliwia stopniowe „odcinanie” fragmentów ostrza, które nie mogą istnieć przy takim, a nie innym kształcie obrazu AFM. Widać jak czarna linia w punkcie trzecim odpowiada kształtem zasymulowanemu ostrzu. Procedurę tę można zapisać matematycznie za pomocą następującego równania

$$P_{i+1} = \bigcap_{\mathbf{x} \in I} [(I - \mathbf{x}) \oplus P_i'(\mathbf{x})] \cap P_i, \text{ podczas gdy } P_i'(\mathbf{x}) = P_i \cap \mathbf{x} - I \quad (13)$$

w którym P_i oznacza model odbicia ostrza w i -tej iteracji, \mathbf{x} jest punktem w przestrzeni trójwymiarowej, I jest modelem obrazu AFM. Operacja $A - \mathbf{d}$ jest translacją zbioru A o wektor \mathbf{d} . Operacje $A \cup B$ i $A \cap B$ są odpowiednio sumą i iloczynem zbiorów A i B . Natomiast

$$A \oplus B = \bigcup_{\mathbf{b} \in B} (A + \mathbf{b}) \quad (14)$$

jest operacją morfologicznej dylatacji zbioru A zbiorem B .

Największą wadą metody BTR jest jej niezwykła wrażliwość na szum pomiarowy w obrazie AFM. Rekonstrukcja kształtu ostrza otrzymana metodą BTR dla obrazu zaszumionego jest obciążona błędem, który sprawia, że sonda jest nadmiernie ostra (patrz rys. 16). W najbardziej popularnej wersji algorytmu zaproponowanej przez Villarubię w pracy [20] zastosowano mechanizm regularyzacji rozwiązania, który wpływa na wrażliwość metody BTR na szum pomiarowy. W rozwiązaniu tym modyfikuje się wrażliwość algorytmu za pomocą dodatkowego parametru t_H zgodnie z równaniem

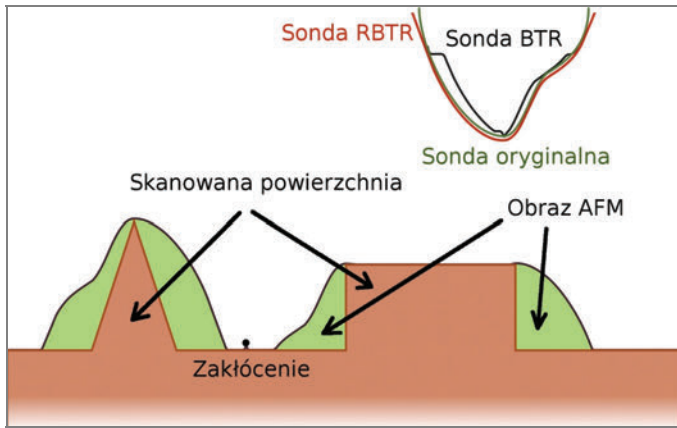
$$P_{i+1} = \bigcap_{\mathbf{x} \in I} [(I + \mathbf{t}_H - \mathbf{x}) \oplus P_i'(\mathbf{x})] \cap P_i, \quad P_i'(\mathbf{x}) = P_i \cap \mathbf{x} - I, \quad \mathbf{t}_H = [00t_H]^T \quad (15)$$

Wartość tego parametru dobierana jest na podstawie sekwencji rekonstrukcji przeprowadzonych dla coraz większych wartości parametru t_H . Jest to procedura czasochłonna, szczególnie dla modeli ostrza, których liczba punktów przekracza 900 (30×30 punktów).

Ogranicza to rozdzielczość uzyskiwanych rekonstrukcji kształtu ostrza. Dlatego w pracy [37] Józwiak i in., wykorzystując mechanizm regularyzacji zaproponowany przez Tiana i in. w pracy [40] i opisany równaniem

$$P_{i+1} = \bigcap_{x \in I} [(I + \mathbf{t}_H - \mathbf{x}) \oplus P_i'(\mathbf{x})] \cap P_i, \quad P_i'(\mathbf{x}) = P_i \cap \mathbf{x} - I - \mathbf{t}_H, \quad \mathbf{t}_H = [00t_H]^T \quad (16)$$

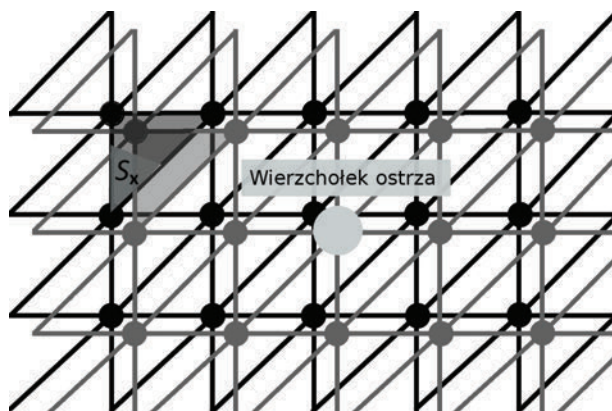
zaproponowali procedurę RBTR, która umożliwia dobór progu t_H na podstawie statystyki szumu pomiarowego. Na rysunku 16 pokazano wpływ szumu na rozwiązanie otrzymane metodą BTR i RBTR. Można na nim zauważyć, że metoda RBTR ułatwiła uzyskanie poprawnego kształtu ostrza, pomimo występującego w obrazie AFM zakłócenia.



Rys. 15. Wpływ szumu na wynik ślepej rekonstrukcji kształtu ostrza [38]
Fig. 15. Influence of image noise on the result of blind tip reconstruction [38]

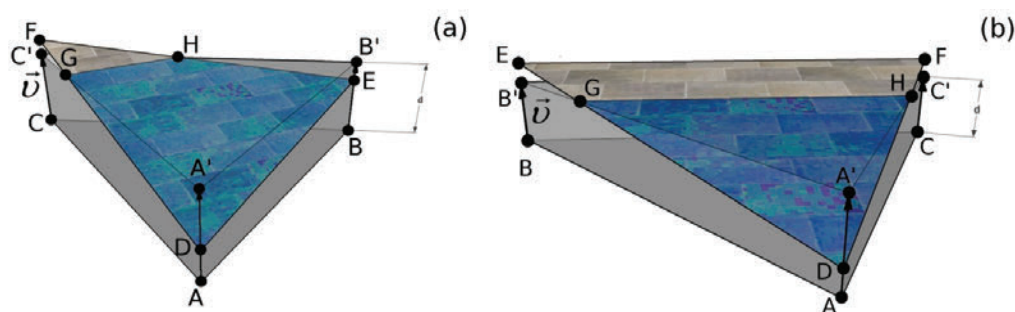
Model kontaktu cieplnego sonda-powierzchnia

Teoretycznie sonda SThM powinna rozpraszać ciepło w sposób jednorodny. Jednakże w przypadku nowoczesnych nanosond SThM to założenie nie jest spełnione ze względu na niepewności wykonania tych niezwykle małych i złożonych urządzeń. Celem procedury TThM jest wyznaczenie rozkładu rozpraszania mocy na powierzchni takiej nanosondy. W metodzie tej sonda modelowana jest za pomocą macierzy liczb, które opisują, jaki procent mocy rozprasza się przez ściśle określony, niewielki fragment powierzchni sondy. Zarówno powierzchnia sondy, jak i powierzchnia badanej próbki jest aproksymowana za pomocą siatki trójkątów z wykorzystaniem danych pomiarowych i wyników ślepej rekonstrukcji kształtu ostrza.



Rys. 16. Model sondy i obrazu powierzchni [37]
 Fig. 16. Model of the probe and the image of the surface [37]

Na rysunku 16 pokazano jak wygląda model sondy i obrazu powierzchni w danym położeniu sondy. Możemy zauważyć, że trójkąty sondy znajdują się dokładnie nad trójkątami obrazu powierzchni próbki za każdym razem, jak wierzchołek sondy znajduje się w jednym z punktów pomiarowych. Dzięki temu można wyznaczyć powierzchnię kontaktu na podstawie wzajemnych wysokości wierzchołków trójkątów znajdujących się bezpośrednio na wprost siebie. Wprowadzono parametr d określający odległość pomiędzy odpowiadającymi sobie wierzchołkami, poniżej której uznaje się, że znajdują się one we wzajemnym kontakcie. Wówczas dla każdej pary trójkątów można wyodrębnić cztery przypadki określające powierzchnię kontaktu między sondą a badaną powierzchnią. Pierwszy najbardziej oczywisty przypadek zachodzi, gdy żaden z wierzchołków nie jest w kontakcie z badaną próbką.



Rys. 17. Metoda obliczania powierzchni kontaktu odpowiadających sobie trójkątów [37]
 Fig. 17. Method for calculating the contact area of corresponding triangles [37]

Wówczas powierzchnia kontaktu dla takiej pary trójkątów wynosi 0. W drugim skrajnym przypadku, kiedy wszystkie wierzchołki pary trójkątów są w kontakcie,

przyjmuje się, że powierzchnia kontaktu jest równa polu trójkąta DEF (rys. 17) i może być policzona zgodnie z równaniem

$$S_x = \frac{1}{2} \left\| \vec{DE} \times \vec{DF} \right\| \quad (17)$$

Na rysunku 17 przedstawione są dwa pozostałe przypadki, w których w kontakcie znajdują się dwa (rys. 17a) wierzchołki pary trójkątów i wówczas pole powierzchni może być obliczone ze wzoru

$$S_x = \frac{1}{2} \left\| \vec{DE} \times \vec{DF} \right\| - \frac{1}{2} \left\| \vec{FG} \times \vec{FH} \right\| \quad (18)$$

oraz jeden wierzchołek (rys. 17b) pary trójkątów. W tym przypadku powierzchnię kontaktu można obliczyć zgodnie z zależnością

$$S_x = \frac{1}{2} \left\| \vec{DG} \times \vec{DH} \right\| \quad (19)$$

W tak zdefiniowanym kontakcie sonda–powierzchnia całkowitą moc rozpraszaną przez sondę w położeniu y można wyrazić wzorem

$$\forall y, p_{\text{tot}}(y) = p_0 + c \sum_{x \in P} S_x(y) p(x) \quad (20)$$

w którym p_0 oznacza część mocy rozpraszanej przez sondę, która nie zmienia się w trakcie procesu skanowania i nie zależy od powierzchni kontaktu, c jest stałą zależną od właściwości środowiska, w jakim przeprowadzany jest pomiar (np. wilgotności czy temperatury sondy), $p(x)$ jest procentem mocy rozpraszanej przez trójkąt x nano-sondy P , $S_x(y)$ oznacza powierzchnię kontaktu dla trójkąta x w położeniu sondy y . Model ten opisuje przepływ ciepła pomiędzy sondą a próbką tylko wtedy, gdy spełnione są specjalne warunki. Pierwszym potencjalnym źródłem problemów może być transfer ciepła przez warstwę powietrza pomiędzy sondą a próbką. W tym przypadku przepływ ten może zachodzić w dwóch reżimach. W pierwszym zakłada się, że podczas transportu ciepła przez powietrze zachodzi wymiana energii cieplnej przez zderzenia między cząsteczkami powietrza. Taki transport ciepła można opisać równaniem Fouriera, w którym parametrem charakterystycznym jest przewodność cieplna wynosząca dla powietrza około 0,024 W/(mK). W tym przypadku przepływ ciepła zależy od szerokości warstwy powietrza. Jeżeli jednak jej szerokość jest mniejsza niż średnia długość drogi swobodnej jego cząsteczek, to można mówić już tylko o transporcie balistycznym, który jest niezależny od szerokości warstwy powietrza. Dla tego przypadku można zdefiniować tylko konduktancję cieplną, która dla powietrza wynosi około $1,3 \cdot 10^5$ W/(m²K) [41]. Dla porównania przewodność cieplna diamentu wynosi około 2200 W/(mK), co na odcinku 400 nm daje konduktancję cieplną na poziomie

$5,5 \cdot 10^9 \text{ W}/(\text{m}^2\text{K})$. Ponieważ diament jest materiałem hydrofobowym, transfer ciepła przez menisk wodny jest mało prawdopodobny, ale nawet wtedy przewodność cieplna wody wynosi $0,5 \text{ W}/(\text{mK})$, co na odcinku 400 nm daje konduktancję cieplną na poziomie $1,25 \cdot 10^6 \text{ W}/(\text{m}^2\text{K})$. Jest to wartość 10 razy większa niż konduktancja cieplna warstwy powietrza. Biorąc podane fakty pod uwagę, w trakcie skanowania powierzchni diamentu moc rozpraszana przez warstwę powietrza powinna pozostawać stała i niezależna od położenia sondy, co jest uwzględnione w modelu w postaci parametru p_0 . Wartość ta jednak nie może być wyznaczona przez pomiar mocy rozpraszanej przez sondę w powietrzu, gdyż wiadomo [42], że jej wartość zmienia się po wejściu w kontakt z badaną powierzchnią.

Do wyznaczenia pola powierzchni obszaru kontaktu $S_x(\mathbf{y})$ konieczna jest znajomość kształtu sondy SThM i topografii powierzchni. W przedstawionym rozwiązaniu do wyznaczenia kształtu sondy SThM wykorzystano procedurę RBTR, natomiast topografia może być zmierzona w osobnym pomiarze za pomocą ostrej sondy AFM, której promień krzywizny jest zwykle ponad 10 razy mniejszy niż promień sondy SThM.

Estymacja parametrów modelu (metoda TThM)

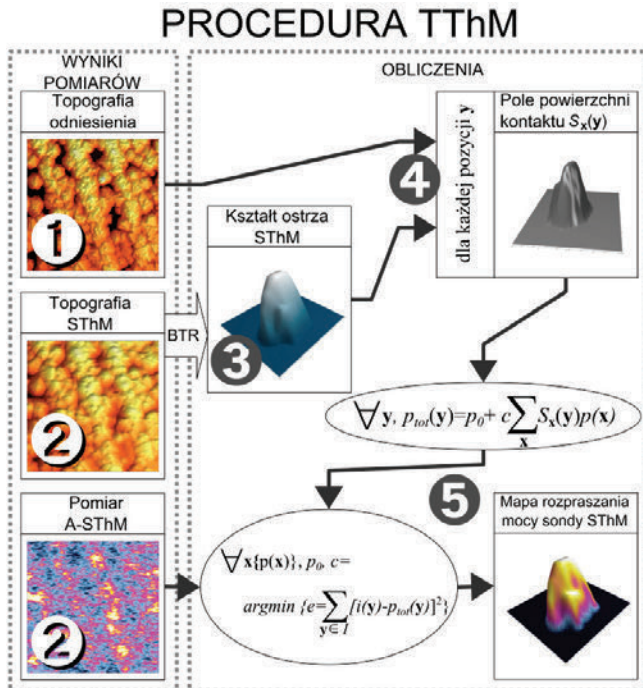
Równanie (20) można wykorzystać do obliczenia rozkładu $p(\mathbf{x})$ mocy rozpraszanej przez każdy trójkąt modelu sondy SThM. Załóżmy, że dysponujemy rzeczywistym kształtem powierzchni materiału o dużej przewodności cieplnej, chropowatości i hydrofobowości (np. diament) oraz parą obrazów AFM i A-SThM jej zarejestrowanych za pomocą badanej sondy SThM. Korzystając z procedury RBTR, możemy wyznaczyć kształt sondy SThM. Wykorzystując kształt sondy i powierzchni, jesteśmy w stanie wyznaczyć pola kontaktu $S_x(\mathbf{y})$ dla każdego trójkąta \mathbf{x} i położenia sondy \mathbf{y} , a korzystając z równania (20) i obrazu A-SThM, możemy sformułować problem optymalizacyjny w postaci

$$p_{0,c}, \forall \mathbf{x} \in P \quad p(\mathbf{x}) = \operatorname{argmin} e = \sum_{\mathbf{y} \in I} [i(\mathbf{y}) - p_{\text{tot}}(\mathbf{y})]^2 \quad (21)$$

w którym $i(\mathbf{y})$ jest mocą rozpraszaną przez sondę SThM w położeniu \mathbf{y} . Zamieniając $cS_x(\mathbf{y})$ na $\hat{S}_x(\mathbf{y})$, otrzymujemy model liniowy, który ze względu na dużą liczbę zmiennych rozwiązywany jest za pomocą nieliniowej metody gradientów sprzężonych [40]. Metoda ta gwarantuje otrzymanie rozwiązania według liczby iteracji równej liczbie zmiennych. Metoda wykorzystuje gradient optymalizowanej funkcji, który może być wyznaczony z równania (21). Szerszego wyjaśnienia wymaga kwestia pomiaru tego samego fragmentu powierzchni za pomocą głowicy AFM o wysokiej rozdzielczości. Żeby to zrobić, należy oddalić się na bezpieczną odległość od powierzchni głowicą A-SThM, wymienić sondę A-SThM na ostrą sondę AFM i ponownie zbliżyć się do powierzchni. Pomiary te można również przeprowadzić w odwrotnej kolejności. Ze względu na niepewność pozycjonowania układu zgrubnego przesuwu po ponownym zbliżeniu sondy AFM system pomiarowy znajduje się w innym punkcie w przestrzeni

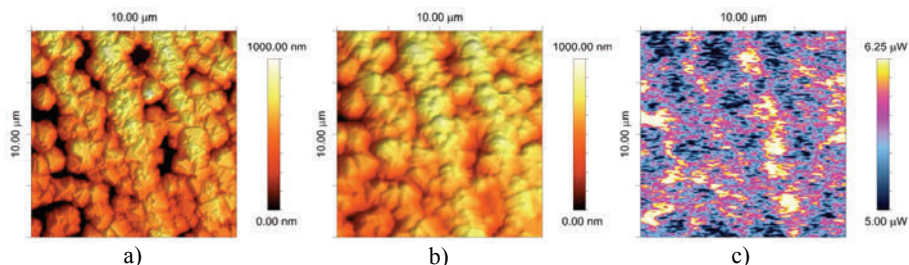
XYZ. Wykonane po takiej operacji skanowanie powierzchni zobrazuje nieco inny obszar niż poprzedni pomiar głowicą A-SThM, co uniemożliwi prawidłowe wyznaczenie powierzchni kontaktu. Dlatego konieczne jest, aby w obu pomiarach zeskanować obszar większy niż zakładany obszar analizy. Wtedy dzięki odpowiedniej procedurze numerycznej będzie możliwe wzajemne ustawienie obrazów topografii uzyskanych ostrzem AFM i SThM. W dziedzinie przetwarzania obrazów problem ten nazywany jest rejestracją obrazów (ang. *Image Registration*).

Rejestracji tej można dokonać, używając funkcji korelacji wzajemnej obrazów jako kryterium umożliwiającego wyznaczenie optymalnej transformacji (przesunięcia i obrotu) jednego z obrazów. Można to przeprowadzić przez iteracyjną translację i obrót o sukcesywnie zmieniany wektor i kąt, maksymalizując wzajemną korelację obrazów. W wyniku opisanej procedury dostajemy wzajemnie odpowiadające obrazy topografii AFM i SThM i obraz rozpraszanej mocy A-SThM. Dalszą procedurę zilustrowano na rysunku 18.



Rys. 18. Ilustracja procedury TThM: 1) pomiar topografii ostrzem AFM, 2) pomiar topografii i mocy rozpraszanej ostrzem SThM, 3) obliczenie kształtu ostrza metodą RBTR, 4) wyznaczenie powierzchni kontaktu, 5) optymalizacja parametrów p_0 , c i $p(x)$ [37]

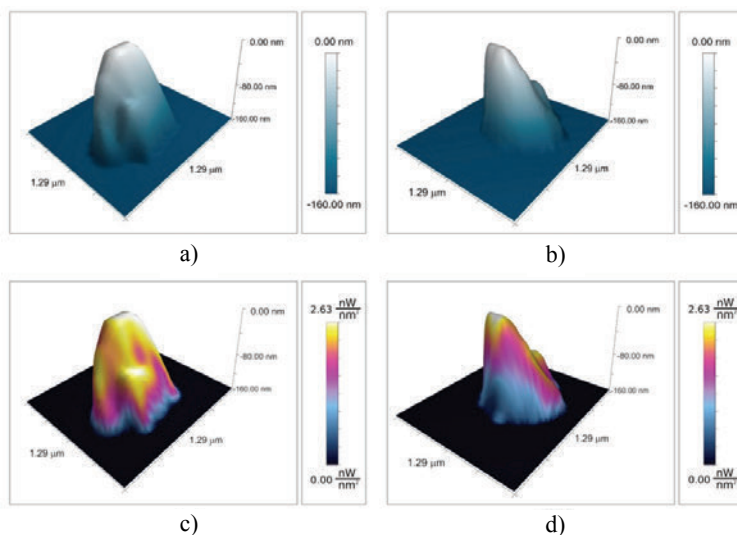
Fig. 18. Illustration of the TThM procedure: 1) standard AFM topography image, 2) SThM topography and dissipated power images, 3) RBTR tip shape calculation, 4) contact area determination, 5) optimization of p_0 , c and $p(x)$ [37]



Rys. 19. Obrazy niezbędne do mapowania właściwości termicznych ostrza; obrazy: a) topografii powierzchni zmierzony ostrzem AFM, b) topografii powierzchni zmierzony ostrzem SThM, c) rozpraszania mocy zarejestrowany w trybie A-SThM [37]

Fig. 19. Images required for tip thermal mapping: a) standard high-resolution AFM topography image, b) SThM topography image, c) active-mode SThM power dissipation image [37]

Na rysunku 19 pokazano przykładowe obrazy uzyskane w wyniku pomiarów AFM i SThM. Na ich podstawie został obliczony kształt ostrza SThM i rozkład rozpraszanej mocy (patrz rys. 20). Dla porównania na rysunku 21 pokazano obraz ze skaningowego mikroskopu elektronowego, na którym można zaobserwować, że obniżenie rozpraszania mocy z jednej strony ostrza jest spowodowane brakiem metalizacji.

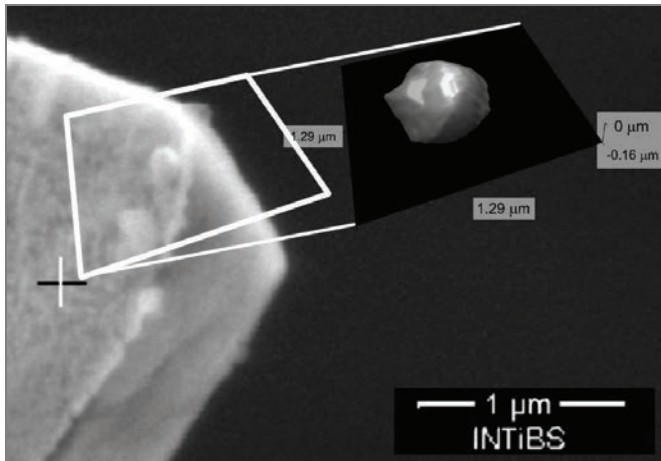


Rys. 20. Wyniki analizy obrazów AFM i SThM:

a, b) kształt sondy SThM, c, d) rozkład rozpraszania mocy przez sondę [37]

Fig. 20. Results of TThM analysis of AFM and SThM images:

a, b) the shape of SThM tip, c, d) power dissipation from the SThM tip [37]



Rys. 21. Obraz SEM sondy SThM z nałożoną rekonstrukcją kształtu ostrza [37]
 Fig. 21. High-resolution SEM image of the SThM tip with the in-scale image
 of the reconstructed tip shape [37]

Obniżenie poziomu rozpraszanej mocy występuje również dookoła kulistej struktury widocznej na obrazie SEM i w zrekonstruowanym kształcie sondy SThM. Spowodowane jest ono warstwą powietrza, która pośredniczy w wymianie ciepła we wklęsłych obszarach dokoła widocznej kulki metalizacji.

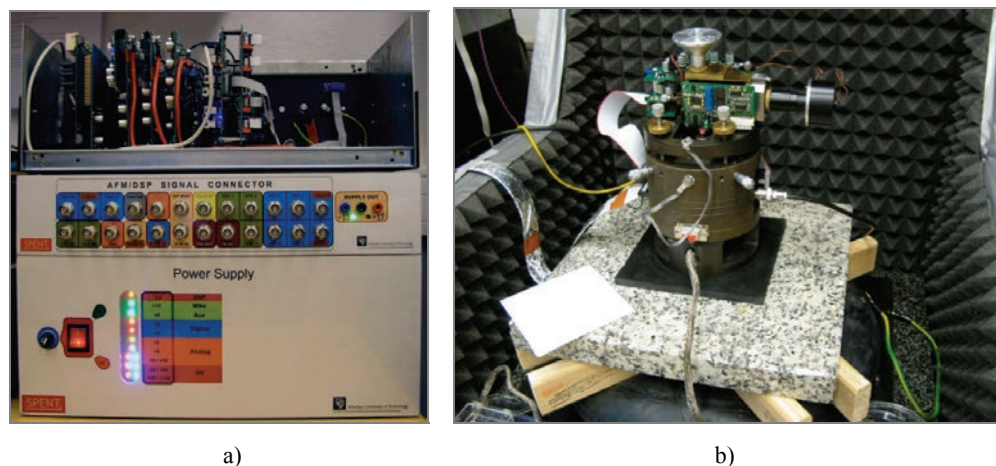
Metoda TThM ułatwia zatem obserwację degradacji właściwości termicznych sond SThM w trakcie ich eksploatacji, podobnie jak metody estymacji kształtu ostrza umożliwiają ocenę stopnia zużycia sond AFM.

5. MIKROSKOP BLISKIEGO POLA – SYSTEM POMIAROWO-STERUJĄCY

Prezentowany system pomiarowy jest sprzętem laboratoryjnym, charakteryzującym się otwartą konstrukcją i architekturą. Został on zaprojektowany, wykonany i uruchomiony przez zespół ZMMiN WEMiF PWr. System ten dzięki swojej wielofunkcyjności, możliwości rozbudowy, dokładności pomiarowej oraz przechowywania i obróbki danych pomiarowych stanowi podstawową bazę doświadczalno-pomiarową w mikroskopii bliskich oddziaływań.

System pokazany na rysunku 22a opiera się na konwencji struktury modułarnej, co ułatwia jego rozwijanie i uruchamianie, ponieważ każdy z jego elementów może działać osobno lub być wykorzystany w innym systemie. Idea modularności opiera się na zastosowaniu jednej, wspólnej magistrali systemowej z rozproszonym zasilaniem i sygnałami sterującymi oraz osobnych kart. Każda z nich przystosowana jest do

realizacji określonego zadania, np. przetwarzania sygnałów analogowych na odpowiadającą im postać cyfrową, pochodzących z głowicy pomiarowej lub generowanie przebiegów analogowych, wykorzystywanych do skanowania. Ten ostatni sygnał podawany jest po wstępnej obróbce na piezoelektryczny skaner XYZ (rys. 22b).



Rys. 22. System pomiarowy SPM WRUT: a) modułowy system kontrolno-pomiarowy wraz z zasilaczem, b) optyczna głowica pomiarowa mikroskopu SPM z zespołem skanowania XYZ oraz z systemem do tłumienia drgań

Fig. 22. WRUT SPM system: a) modular measurement and control system with supply device, b) optical SPM measurement head with XYZ scanning module and vibration suppression system

Prezentowany system można podzielić na trzy podstawowe składniki:

- głowicę pomiarową wraz z zespołem skanującym,
- uniwersalny sterownik systemowy z zasilaczem,
- oprogramowanie sterujące procesem skanowania i umożliwiające obróbkę danych.

W celu przybliżenia zasady działania systemu, każdy z elementów został pokrótce opisany.

Głowica pomiarowa

Głowica pomiarowa jest to zespół mechaniczny, który ułatwia wykonanie pomiaru struktury za pomocą krzemowej mikrobilki i dedykowanego układu detekcji jej odchylenia, jak również umożliwiający aktywną jej wychylenia. Element ten może zostać wzbogacony o dodatkowe moduły, rozszerzające możliwości pomiarowe, np. badanie cienkich warstw dielektryków z przewodzącą sondą bądź obserwacje właściwości termicznych mikro- i nanostruktur za pomocą specjalnie zmodyfikowanej sondy, wyposażoną w pętlę termiczną. Stosowane układy wzbogacają już istniejący system i umożliwiają częściową

obróbkę sygnałów analogowych biegnących z i do głowicy (filtrowanie, wzmocnienie słabych sygnałów, układ do generowania drgań samej belki pomiarowej). W celu uzyskania informacji o tym, jak sama sonda odchyła się podczas pomiaru, można wykorzystać jedną z metod detekcji ugięcia: prądu tunelowego, pojemnościową, piezoelektryczną, optyczną bądź piezorezystywną. Dwie ostatnie zostaną omówione w dalszej części pracy ze względu na ich dominujący udział w mikroskopach SPM.

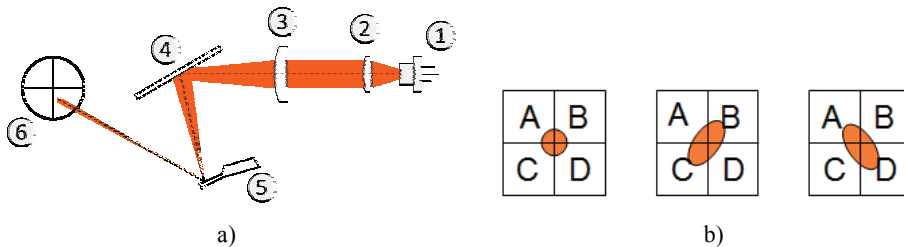
Detekcja natężeniowa ugięcia belki jest najbardziej rozpowszechnioną metodą wśród mikroskopów bliskich oddziaływań (rys. 23a). Charakteryzuje się ona dużą czułością i odpornością na zewnętrzne zakłócenia. Ponadto umożliwia wykorzystanie jej z większością dostępnych dźwigni dostępnych na rynku. Podstawowym założeniem jest wykorzystanie skoligowanej wiązki laserowej, skupionej na powierzchni belki. Plamka lasera zostaje przesunięta na powierzchni fotodetektora pod wpływem ugięcia mikrodźwigni (rys. 23b). Odbity promień od dźwigni pada na detektor czterosekcyjny (ang. *Position Sensing Device* – PSD) przedstawiony na rysunku 24. Wykonując podstawowe operacje matematyczne z sygnałów z każdej sekcji fotodetektora, możliwe jest uzyskanie informacji o położeniu plamki lasera.

$$\text{Suma} = A + B + C + D$$

$$\text{Topografia} = (A + B) - (C + D)$$

$$\text{Tarcie} = (A + C) - (B + D)$$

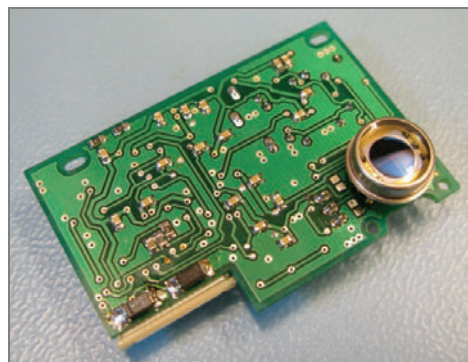
Sygnały topografii i tarcia odnoszą się do ugięcia i skręcenia mikrobilki, dając podstawową informację o jej ruchu – rys. 25 [43].



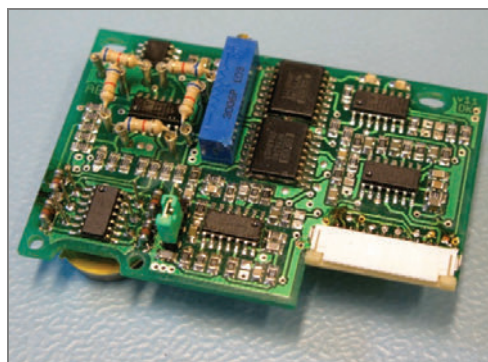
Rys. 23. Detekcja natężeniowa ugięcia belki: a) schemat przebiegu wiązki lasera, b) przykładowe rozkłady plamki lasera na detektorze czterosiecznym

Fig. 23. Intensity-based detection of cantilever deflection: a) laser beam routing, b) some of the possible positions of the laser beam on the position-sensing device

Wzrost rozdzielczości pomiaru można uzyskać przez zwiększenie odległości między detektorem a belką pomiarową. Gdy to nastąpi, zwiększy się również ramię, a tym samym bezwzględne przesunięcie plamki lasera. Jednym z ograniczeń jest powierzchnia detektora, wielkość plamki oraz moc promienia lasera docierającego do fotodetektora.



a)



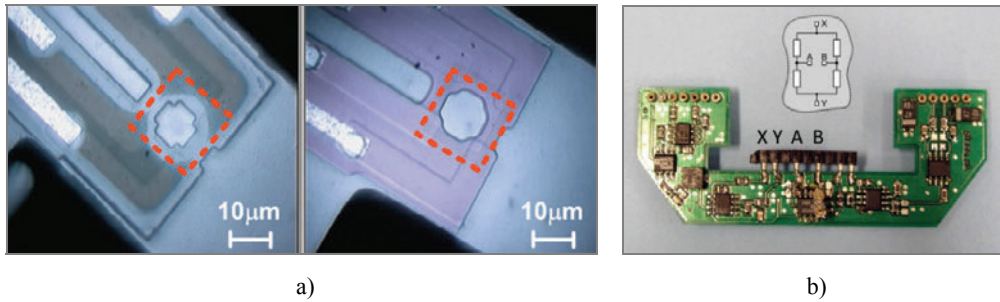
b)

Rys. 24. Układy elektroniczne normalizujące sygnał z fotodetektora czterosekcyjnego:
 a) dolna warstwa płytki elektronicznej wraz z detektorem, b) górna część układu elektronicznego,
 z układami wzmacniającymi, filtrującymi oraz normalizującymi sygnał

Fig. 24. Electronic devices used for normalization of the photodiode signals:

a) bottom layer with the PSD, b) top layer with amplifiers, filters and normalization devices

Opisana metoda natężeniowa charakteryzuje się niezawodnością i bardzo dużą precyzją pomiarową, rozdzielczość detekcji wychylenia belki dochodzi do 10 fm i może być stosowana w cieczech. Wymagana jest również kalibracja – przypisanie wartości napięciowej z fotodetektora odpowiadającej jej wychyleniom w nanometrach, a skomplikowany układ optyczny może stanowić podstawową wadę. Układ optyczny można zastąpić, wykorzystując m.in. tzw. piezorezystywną detekcją wychylenia mikrobelki. Zjawisko piezorezystywne polega na zmianie wartości rezystancji danego materiału pod wpływem naprężeń mechanicznych [44–46]. Zasada działania polega na kontrolowaniu zmiany rezystancji elementu tak samo jak w układach tensometrycznych, z tą różnicą, że piezorezystor nie znajduje się na powierzchni, a wewnątrz struktury półprzewodnikowej w zasięgu największych naprężeń mechanicznych (rys. 25). Najczęściej tworzą one zespół piezorezystorów tworzący mostek Wheatstone’a lub pojedyncze piezorezystory, będąc polaryzowane ze stabilnego źródła napięciowego. Wyjściowy sygnał napięciowy niesie ze sobą informację o ugięciu belki i zawiera się w zakresie kilkunastu mikrowoltów. Detekowany sygnał umożliwia określenie działania sił dla pojedynczych pikoniutonów i wychyleń rzędu pikometrów. Niestety jest on podatny na zakłócenia, a w szczególności na zmiany temperatury, która może być interpretowana jako ugięcie. Ponadto wytworzone piezorezystory mogą charakteryzować się rozrzutem wartości, co powiększa błędy pomiarowe. Główną zaletą opisanej metody jest jej prostota i łatwa integracja z układami elektronicznymi. Kolejnym atutem jest niestosowanie drogich i skomplikowanych układów optycznych do ogniskowania i justowania wiązki laserowej, która nie sprawdzi się w matrycach czujników ze zintegrowanym detektorem ugięcia.



Rys. 25. Piezorezystywna dźwignia pomiarowa – system do detekcji ugięcia belki krzemowej:

a) część belki pomiarowej wraz z zaznaczonymi piezorezystorami, b) układ elektroniczny do polaryzacji mostka Wheatstone'a i odczytem napięcia odpowiedzi z piezorezystorów

Fig. 25. Piezoresistive cantilever – system for detection of deflection:

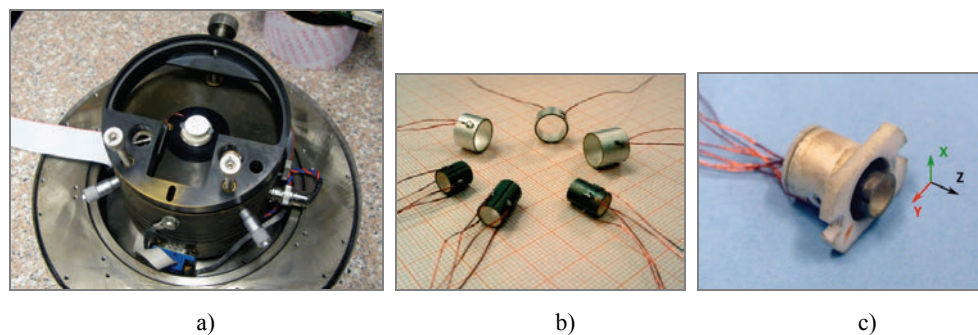
a) location of piezoresistors on the cantilever, b) electronic device for polarization of the Wheatstone bridge and for response signals read-out

Głowica pomiarowa ma również możliwość zadania wychylenia belki krzemowej. Można to zrealizować na kilka sposobów. Pierwszym i podstawowym jest wykorzystanie aktuacji termicznej. Gwałtowne zmiany temperatury są w stanie wprowadzić ją w drgania, efekt ten połączony jest z jej budową. Ponadto niewielka masa belki oraz duża przewodność cieplna między nią a otoczeniem umożliwia uzyskanie stałych czasowych rzędu milisekund. Ogrzewanie dźwigni odbywa się za pomocą mikrogrzejnika umieszczonego na szczycie belki; warto podkreślić też możliwość kontroli wychylenia statycznego belki z rozdzielczością rzędu pojedynczych Angströmów w zakresie pojedynczych mikrometrów [47, 48]. Układ termicznego wychylenia umożliwia wzbudzenie mikrobilki w zakresie jej częstotliwości rezonansowej, dzięki czemu metoda ta wykorzystywana jest do kontroli poszczególnych dźwigni w układach matryc. Nie jest to jedyna metoda aktuacji dźwigni lub ich zespołów. Efekt ten możliwy jest do uzyskania dzięki zastosowaniu układów piezoelektrycznych, magnetycznych lub elektrostatycznych. Dwa ostatnie muszą być wspierane przez zastosowanie zewnętrznych układów elektrycznych sprzężonych ze strukturami wytworzonymi na mikrodźwigni.

Zespół przesuwu próbeki

Zespół ten stanowi nieodzowny element głowicy pomiarowej. Sama podstawa zawiera w sobie element aktywny do przesuwania badanej próbki w trzech osiach – X , Y i Z (poziomo i pionowo; rys. 26a). Najczęściej stosuje się w tym celu specjalnie skonstruowany element piezoelektryczny w postaci piezorurki bądź piezostosu. Elementy te wykonane są z PZT (ceramika z cyrkonianu-tytanianu ołowiu) i wykorzystują efekt piezoelektryczny odwrotny. Polega on na deformacji kształtu materiału pod wpływem działającego zewnętrznego pola elektrycznego. W praktyce wykorzystywane jest napięcie polaryzacji ceramiki rzędu setek woltów (rys. 26b). Piezorurka ma pięć elektrod sterujących, cztery z nich znajdują się na zewnątrz w celu uzyskania wychylenia w osiach X i Y , natomiast na

środkową elektrodę podawany jest sygnał odniesienia, w tym wypadku masa (rys. 26c). Jak już wspomniano, próbka poruszana jest w trzech osiach, dlatego do uzyskania wychyleń w osi Z wykorzystuje się piezoturkę o większej średnicy umieszczoną na zewnątrz poprzednio omówionej. Jej zasada działania jest podobna, lecz ma tylko 2 elektrody polaryzacyjne (wewnętrzną i zewnętrzną). Przedstawiony układ w zależności od długości piezoturki umożliwia uzyskanie wychyleń rzędu pojedynczych mikrometrów, aż do 10–25 μm z rozdzielczością skanowania do dziesiątek nanometrów.

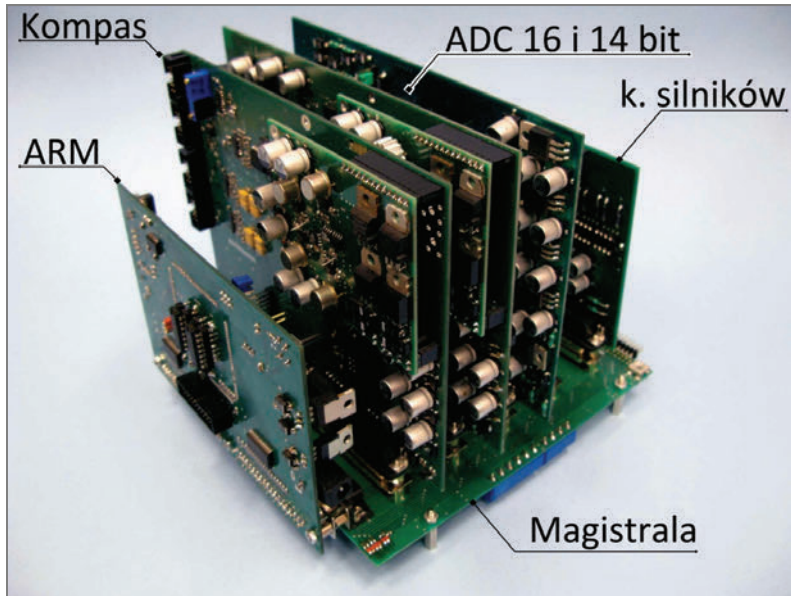


Rys. 26. a) Podstawa do aktuacji próbki z uchwytem na głowicę, b) piezoturki wykorzystywane w skanerach, c) gotowy układ skanujący z ceramicznym uchwytem montującym z zaznaczonymi osiami skanowania

Fig. 26. a) Base with piezoelectric scanner and housing for the measurement head, b) piezotubes used in scanners, c) piezoscanner with ceramic mounting plate with scanning axes indicated

Podstawowa jednostka pomiarowa złożona jest ze stałoprądowego zasilacza o dużej wydajności prądowej oraz sterownika kontrolno-pomiarowego. Oba układy mają strukturę modułową w celu ułatwienia użytkownikom serwisowania oraz rozwijania. Sam sterownik składa się z kilku podstawowych kart montowanych na wspólnej magistrali. Zdjęcie sterownika mikroskopu bliskich oddziaływań wraz z zestawem podstawowych kart przedstawiono na rysunku 27.

Do wspólnej magistrali systemowej możliwe jest podłączenie aż 8 kart peryferyjnych. Komunikacja między uniwersalnym sterownikiem typu ARM a poszczególnymi kartami odbywa się za pomocą 16-bitowej, szeregowej szyny danych z wykorzystaniem dodatkowych linii sterujących. Dzięki magistrali możliwe jest doprowadzanie zasilania do każdej z kart dla podzespołów analogowych i cyfrowych. Takie rozwiązanie zmniejsza znacząco powstawanie zakłóceń w torach analogowych i zwiększa niezawodność systemu pomiarowego. Dedykowana konstrukcja uniwersalnego sterownika ARM z układem LPC2478 firmy NXP, zaprojektowanego specjalnie dla tego typu urządzenia, stanowi alternatywę dla drogich i skomplikowanych procesorów sygnałowych DSP (np. TigerSharc TS101 firmy Analog Devices). Optymalizacja kodu i rozwiązania autorskie zapewniają możliwość ciągłego skanowania, wstępnej obróbki danych oraz komunikacji z komputerem typu PC bez utraty danych pomiarowych.



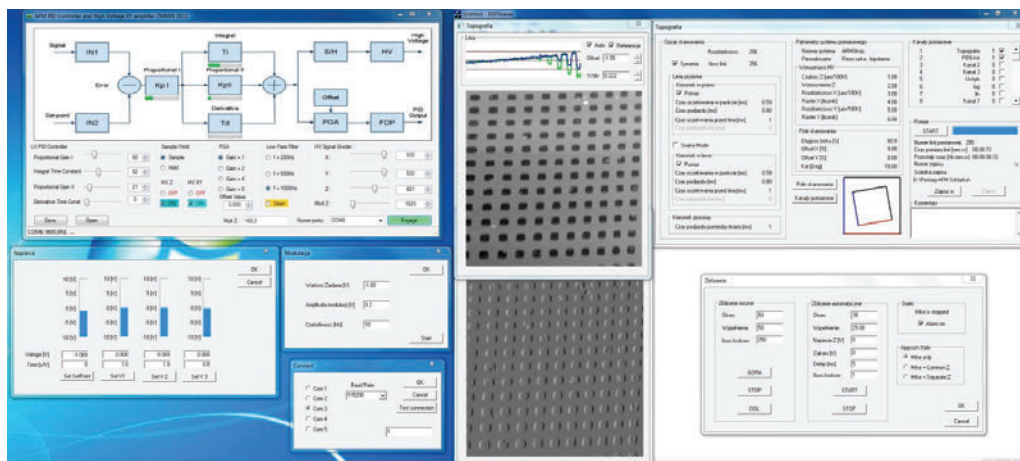
Rys. 27. System kontrolno-pomiarowy zmontowany na magistrali ze sterownikiem typu ARM i rozmieszczeniem poszczególnych kart

Fig. 27. Measurement and control system based on a dedicated bus, with ARM-based controller and specific-purpose extension cards

Kolejnym atutem uniwersalnego sterownika jest możliwość rozwijania konstrukcji i wzbogacania go o nowe tryby pomiarowe z wykorzystaniem już istniejących kart. W celu zapewnienia ciągłej komunikacji między systemem a komputerem wykorzystuje się jedno z wbudowanych peryferii. Standardową techniką wykorzystywaną do komunikacji jest port UART (ang. *Universal Asynchronous Receiver and Transmitter* – UART) z autorskim oprogramowaniem DSP Skaner (rys. 28). Jednak wraz z rozwojem systemu możliwe jest uruchomienie i wykorzystanie innych peryferii, takich jak USB lub Ethernet.

Jak już zostało wspomniane, możliwe jest dołączenie do 8 kart peryferyjnych. Pośród nich są moduły analogowo-cyfrowe do próbkowania sygnałów ciągłych. Sam moduł wyposażony jest w grupę filtrów antyaliasingowych dla ośmiu wejść 14-bitowych przetworników lub dwóch 16-bitowych układów ADC (ang. *Analog to Digital Converter* – ADC). Poziom mierzonych sygnałów zawarty jest w przedziale ± 10 V i próbkowany jest z częstotliwością 100–500 kSPS w zależności od użytej karty. Sam system ma możliwość czasowego gromadzenia danych pomiarowych i ich uśredniania. Należy również zwrócić uwagę, że podczas pomiaru każde z wejść może rejestrować niezależny sygnał analogowy, nie tylko z głowicy pomiarowej, ale również z innych urządzeń – wzmacniacz fazoczuły, konwerter prąd–napięcie i inne. Zaletą tego rozwiązania jest uzyskanie danych pomiarowych, zdigitalizowanych w ściśle

ustalonym czasie pomiarowym oraz możliwość obserwacji i obróbki w przypisanym oprogramowaniu na komputerze klasy PC.



Rys. 28. Programy DSPScanner oraz AFM PID Controller służące do kontrolowania procesu skanowania
Fig. 28. Screenshots of DSPScanner and AFM PID Controller applications used for control of the scanning process

System pomiarowy opracowany w ZMMiN WEMiF PWr. wyposażony jest również w przetworniki cyfrowo-analogowe umiejscowione na osobnej karcie sterowania procesem skanowania zwanej Kompasem. Wykorzystywana jest do generowania sygnałów ciągłych do sterowania parametrami procesu skanowania. Ten skomplikowany układ elektroniczny, wyposażony w przetworniki DAC (ang. *Digital to Analog Converter* – DAC), i potencjometry cyfrowe umożliwiają wygodne pozycjonowanie się na próbce pomiarowej, wybór skanowanego miejsca oraz jego obrót bez utraty rozdzielczości pomiarowej, jak na przykład w niektórych systemach komercyjnych. Maksymalna liczba punktów pomiarowych sygnału sterującego skanowaniem XY wynosi nawet 1024 dla pojedynczej linii, co przekłada się bezpośrednio na uzyskanie zdolności rozdzielczej na poziomie pojedynczych nanometrów (dla pola skanowania $10 \times 10 \mu\text{m}$). Karta Kompas ma cztery niezależne wyjścia napięciowe wykorzystywane np. do polaryzacji podłoża próbki czy innych. Dodatkowymi trybami pomiarowymi, niespotykanymi w wielu systemach pomiarowych, jest spektroskopia sił czy skanowanie po zadanej trajektorii, np. harmoniczej. To ostatnie rozwiązanie zmniejsza wpływ histerezy piezoskanera i przekłada się bezpośrednio na uzyskanie lepszych wyników pomiarowych. Kolejnym rozszerzeniem systemu pomiarowego jest wyposażenie go w kartę odpowiedzialną za ruch głowicy pomiarowej z wykorzystaniem precyzyjnych silników prądu stałego. To rozwiązanie umożliwia zgrubne zbliżenie belki pomiarowej do powierzchni w płaszczyźnie pionowej. Poruszanie głowicą pomiarową

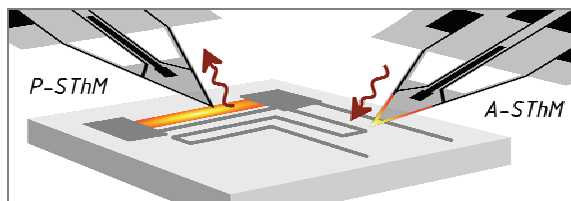
możliwe jest nawet na dystans kilkunastu milimetrów z rozdzielczością do 10 μm . Do precyzyjnego sterowania wysokością zawieszenia ostrza nad powierzchnią wykorzystuje się regulator proporcjonalno-całkująco-różniczkujący (ang. *Proportional-Integral-Derivative Controller* – PID) wraz ze wzmacniaczem wysokiego napięcia. Przedstawiony układ odpowiada za zachowanie stałej odległości między ostrzem a próbką podczas pomiaru oraz za odchyłanie w płaszczyznach *XYZ* piezorurki. Jest to możliwe dzięki połączeniu go z sygnałami wyjściowymi z karty Kompas. Zastosowany regulator PID ma tor analogowy, ale sterowany jest za pomocą potencjometrów cyfrowych. Umożliwia to sprawowanie nad nim kontroli z wykorzystaniem zewnętrznego oprogramowania na komputer klasy PC (AFM PID Controller, rys. 28) za pomocą interfejsu komunikacyjnego UART. Zestaw regulatora i wzmacniacza również jest montowany na wspólnej magistrali systemowej, jednak ma doprowadzone z zewnątrz wysokie zasilanie, rzędu $\pm 150\text{ V}$.

Układy sterująco-pomiarowe skaningowego mikroskopu bliskiego pola termicznego

Zastosowanie sond termooporowych w pasywnej i aktywnej mikroskopii bliskiego pola termicznego wymaga opracowania precyzyjnych elektronicznych układów pomiarowo-sterujących, których zadaniem jest:

- akwizycja sygnałów, które odpowiadają temperaturze końcówki sondy,
- sterowanie temperaturą ostrza i pomiar mocy elektrycznej dostarczanej do sondy, która jest niezbędna do utrzymania temperatury końcówki na zadanym poziomie.

Na rysunku 29 przedstawiono schematycznie układy pomiarowe aktywnego i pasywnego mikroskopu bliskiego pola termicznego ze schematycznie zaznaczonymi współczesnymi sondami oporowymi [49].

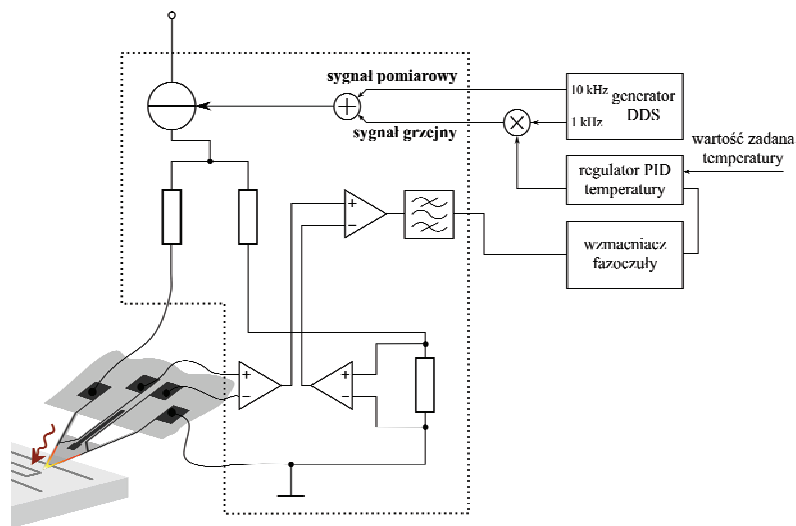


Rys. 29. Pasywna i aktywna mikroskopia bliskiego pola termicznego z zastosowaniem oporowych sond termicznych

Fig. 29. Comparison between passive- and active-mode scanning thermal microscopy with thermoresistive probes

W proponowanych rozwiązaniach sondy oporowe są włączone w układ mostka zmiennoprądowego (rys. 30) [50]. Należy zwrócić uwagę, że sygnały detekowane na przekątnej mostka są w zakresie pojedynczych mikrowoltów, czego konsekwencją jest

konieczność nadzwyczaj starannego projektowania układów elektronicznych. Zastosowanie układów zmiennoprądowych eliminuje w tym przypadku m.in. wpływ pasywnych sił termoelektrycznych w obwodzie i ułatwia eliminację zakłóceń.



Rys. 30. Schemat układu sterująco-pomiarowego mikroskopu bliskiego pola termicznego współpracującego z oporowymi sondami mikroskopu bliskiego pola termicznego

Fig. 30. Schematic of the measurement and control system for the scanning thermal microscope using thermoresistive probes

Układ pomiarowo-sterujący tworzy mostek pomiarowy współpracujący z czteropunktowymi sondami oporowymi. Sygnał odpowiadający zmianom rezystancji sondy jest prowadzony doprowadzeniami napięciowymi. Zwykle mostki pomiarowe pracują w układach dwuczęstotliwościowych.

Ponieważ w trybie pasywnym mikroskopii termicznej bliskiego pola pomiar napięcia niezrównoważenia dokonywany był w układzie wzmacniacza fazoczułego (ang. *lock-in amplifier*), źródło prądowe zasilające mostek było sterowane z precyzyjnego generatora bezpośredniej syntezy sygnału (ang. *Direct Digital Synthesis – DDS*), który charakteryzuje się dużą stabilnością amplitudy i częstotliwości generowanego sygnału. Możliwa jest dzięki temu eliminacja tzw. szumu fazowego, który powstaje, gdy zmienia się różnica faz między sygnałem z przekątnej mostka i ustalonym sygnałem taktującym prostownik wzmacniacza fazoczułego. W torze pomiarowym wzmacniacza fazoczułego zastosowano również filtry pasmowoprzepustowe, których zadaniem jest wstępne odfiltrowanie zakłóceń pomiarowych i ograniczenie szumu. Niestabilność częstotliwości sygnału zasilającego mostek może zatem wpływać na wahania sygnału wyjściowego tych filtrów, szczególnie w przypadku zastosowania układów o dobroci powyżej 5. W eksperymentach, w których nie jest wymagana detekcja sy-

gnałów o bardzo małym poziomie, opisywany układ może być uproszczony przez zastosowanie w torze pomiarowym jedynie filtrów pasmowych.

W przeprowadzanych eksperymentach mostek pomiarowy zasilano prądem o natężeniu 200 μA . Dla takiej wartości prądu zasilającego nie zaobserwowano samonagrzewania się sondy oporowej. W celu kompensacji wpływu zmian temperatury otoczenia na wynik jej pomiaru temperatury panującej na powierzchni próbki w obydwie gałęzie mostka pomiarowego włączano sondy oporowe. Sondę odniesienia mocowano w głowicy pomiarowej możliwie blisko sondy pomiarowej znajdującej się w kontakcie z preparatem. Na podstawie przedstawionych założeń, i przyjmując wartość temperaturowego współczynnika zmian rezystancji platyny 0,00392 1/K, czułość układu mostkowego (rozumianą jako przyrost wartości skutecznej napięcia na przekątnej mostka podczas zmiany temperatury ostrza pomiarowego o 1 K) oszacować można na poziomie 10 $\mu\text{V/K}$. Zdolność rozdzielczą systemu (rozumianą jako zdolność do zaobserwowania najmniejszej zmiany temperatury końcówki oporowej) oblicza się natomiast, odnosząc sumę napięć szumu termicznego rezystorów mostka do oszacowanej wcześniej czułości pomiarowej. Szum termiczny rezystorów tworzących mostek oblicza się według zależności

$$u_{R_i'} = \sqrt{4k_b TR_i' B} \quad (22)$$

i

$$u_{S_i} = \sqrt{4k_B TR_{S_i} B} \quad (23)$$

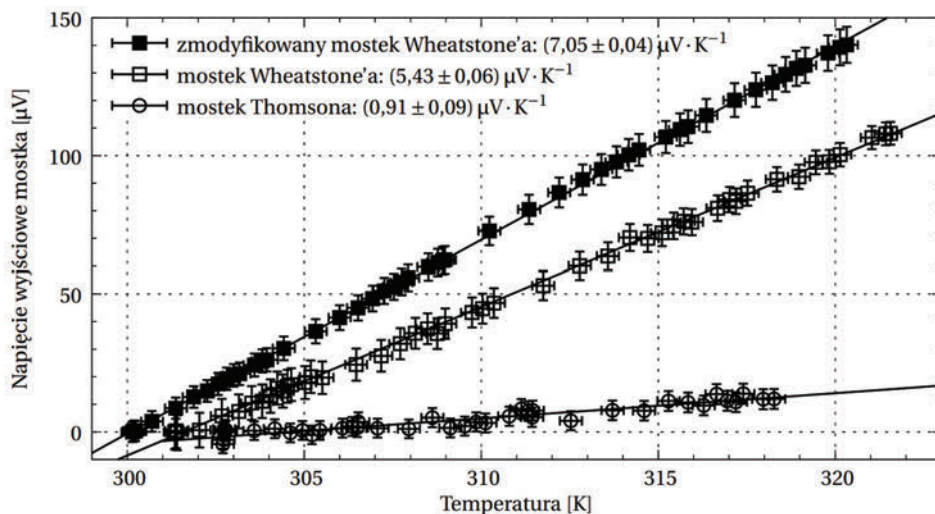
gdzie R_i' jest łączną rezystancją oporników włączanych szeregowo z sondami oporowymi o rezystancji R_{S_i} w gałęziach pomiarowych i gałęziach odniesienia mostka pomiarowego. Napięcie elektryczne związane z tymi źródłami napięciowymi, występujące w przekątnej pomiarowej mostka, wynosi zatem

$$u_{\text{mth}.B} = \sqrt{(u_{R_i'}^2 + u_{S_i}^2)} \quad (24)$$

Po uwzględnieniu wartości rezystancji zastosowanych w opracowanym układzie jest ono w paśmie 10 Hz równe 15 nV, co odpowiada zmianom temperatury rzędu pojedynczych ułamków milikelwina. Należy jednak podkreślić, że podane obliczenia mają charakter jedynie przybliżony i służą wstępnemu zdefiniowaniu wymagań dla elektronicznych układów wejściowych współpracujących z mostkiem. W przypadku praktycznych pomiarów, prowadzonych za pomocą statycznego mikroskopu bliskiego pola termicznego, ogrzewana jest jedynie część sondy oporowej, a nie jej całość, tak jak założono w obliczeniach. Konsekwencją tego zjawiska jest zdecydowanie mniejsza czułość pomiarowa mostka.

Odpowiedź statyczna mikros sondy oporowej jest zwykle wyznaczana eksperymentalnie za pomocą mikrostolika, którego temperatura jest jednocześnie precyzyjnie sterowana za pomocą grzejnika grubowarstwowego i mierzona za pomocą czujnika Pt100. Od konstrukcji takiej oczekuje się zwykle małych rozmiarów, tak aby rozgrze-

wany podest termiczny nie ogrzewał głowicy pomiarowej i mikrobelki. Dodatkowo konstrukcja ta powinna być stabilna mechanicznie, tak aby podwyższanie temperatury nie powodowało rozszerzalności układu, prowadzącej do wzrostu siły nacisku ostrza na końcówkę. Wyniki kalibracji przedstawiono na rysunku 31 [50]. Na wykresie zaznaczono napięcia mierzone w układzie prostokątnym na wyjściu układu fazoczułego. Rejestrację sygnału wyjściowego mostka przeprowadzono dla różnych układów pomiarowych, jakie można było zbudować w zaproponowanej przez nas konstrukcji. Na rysunku 31 pokazano natomiast wyniki kalibracji mostka dla różnych układów połączeń sondy termicznej. Z przedstawionego porównania jednoznacznie wynika, że układ połączenia sondy w konfiguracji zapewnia najlepszą czułość równą około $7 \mu\text{V}$.

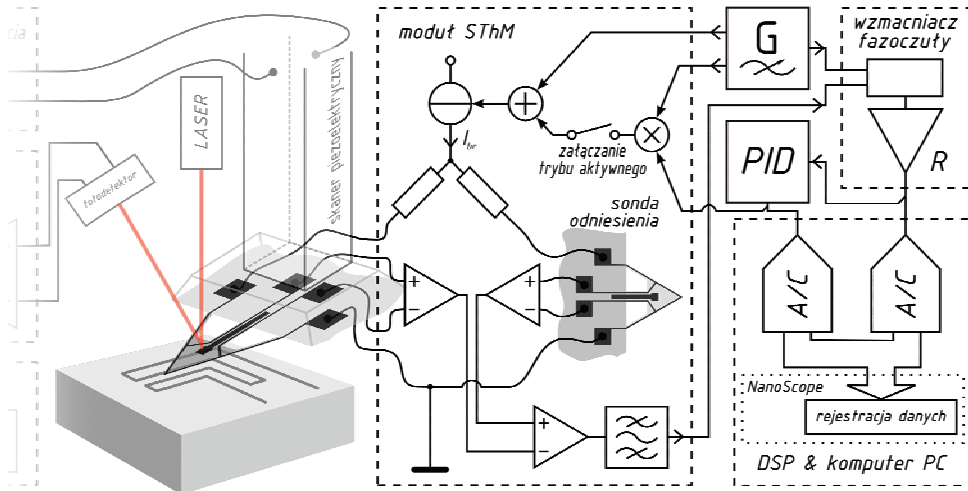


Rys. 31. Kalibracja czułości pomiarowej układu statycznego mikroskopu bliskiego pola termicznego
Fig. 31. Calibration of measurement sensitivity of static scanning thermal microscope

W opracowanych rozwiązaniach układowych bloku pomiaru temperatury mikrosondy zastosowano małoszumowe zmiennoprądowe wzmacniacze różnicowe o gęstości napięciowego szumu termicznego wejścia na poziomie $5 \text{ nV}/\text{Hz}^{0.5}$. Dodatkowo dla poprawy stosunku sygnał–szum w torze pomiarowym zastosowano selektywne filtry aktywne o dobroci 10, dostrojone do częstotliwości nośnej prądu zasilającego mostek. Wyjście filtrów po przeprowadzeniu przez bufory wyjściowe jest łączone z zewnętrznym wzmacniaczem fazoczułym.

W przypadku aktywnej mikroskopii bliskiego pola termicznego sonda oporowa, jak już wspomniano, pełni rolę mikrogrzejnika (aktuatora) i czujnika temperatury (rys. 32). Do sondy doprowadzone są zatem dwa sygnały, z których jeden odpowiada za ogrzewanie końcówki, drugi zaś za określenie jej temperatury. W ujęciu praktyki pomiarowej rozwiązaniem najczęściej stosowanym jest układ z podziałem częstotliwości.

W rozwiązaniu takim sonda jest zasilana sumą dwóch sygnałów o różnych częstotliwościach. Jeden z tych sygnałów, o małym poziomie, jest związany z pomiarem temperatury, drugi zaś – o innej częstotliwości – odpowiada za ogrzanie końcówki. Za pomocą układów filtrów selektywnych (lub wzmacniaczy fazoczułych) obydwa sygnały są rozdzielane i poddawane osobnemu przetwarzaniu.



Rys. 32. Układy sterująco-pomiarowe dla mikroskopu termicznego bliskiego pola – sposób połączenia z nadrzędnym sterownikiem sygnałów

Fig. 32. Measurement and control devices used in scanning thermal microscopy and their connections to external devices

Układem odpowiedzialnym za sterowanie temperaturą końcówki jest regulator PI, mający dwa wejścia. Do pierwszego wejścia podawany jest sygnał $x(t)$ odpowiadający bieżącej temperaturze końcówki pomiarowej. Do drugiego wejścia regulatora podawany jest sygnał wartości zadanej $x^*(t)$, który odpowiada temperaturze, jaka powinna być osiągnięta na mikrosondzie. Różnica tych dwóch sygnałów, nazywana uchybem regulacji $e(t)$, jest poddawana przetwarzaniu proporcjonalno-całkowującemu. W zapisie operatorowym obowiązują zatem następujące równania:

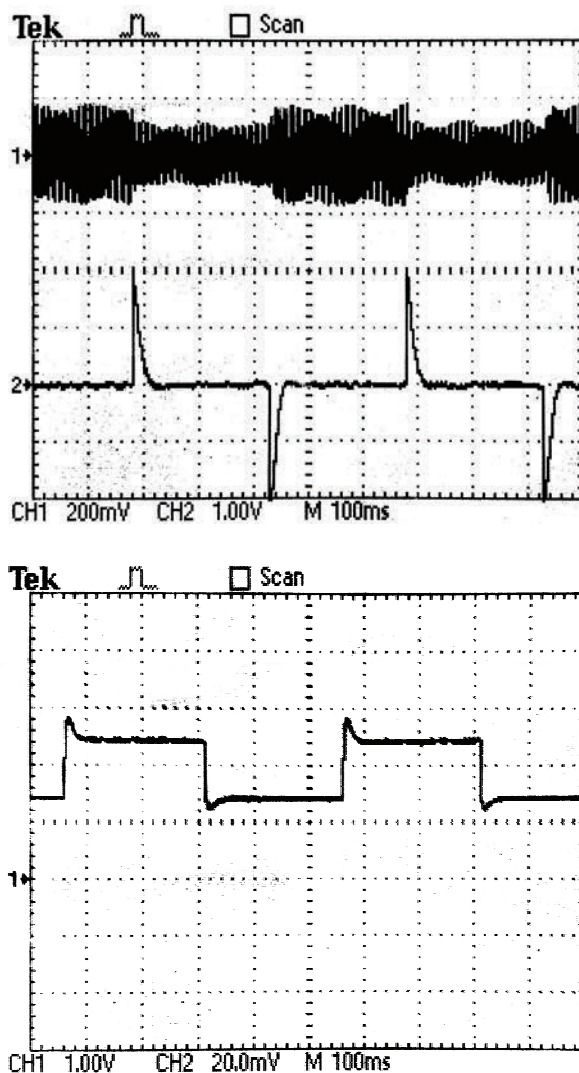
$$E(s) = X(s) - X^*(s) \quad (25)$$

oraz

$$U(s) = K_{pT} \left(1 + \frac{1}{T_{iT} s} \right) E(s) \quad (26)$$

gdzie K_{pT} jest wzmocnieniem proporcjonalnym regulatora, T_{iT} jest natomiast stałą całkowania regulatora. Wyjście regulatora jest połączone z blokiem mnożnika, które-

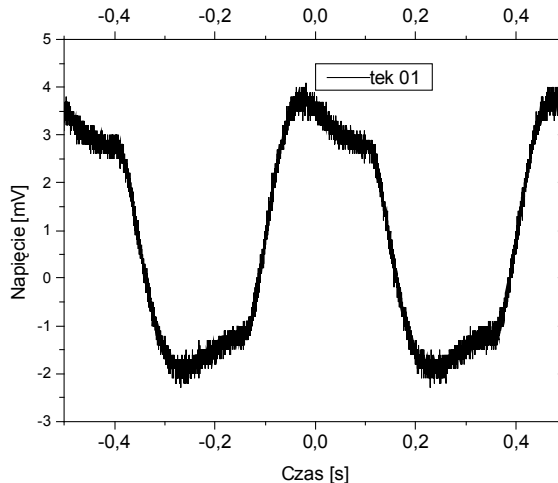
go wejście sygnałowe zasilane jest sygnałem z generatora grzewczego. Sygnał wyjściowy mnożnika doprowadzony jest do źródła prądowego, gdzie obydwa sygnały są sumowane w obwodach wejściowych i następnie poddane przetwarzaniu na prąd zasilania układu mostkowego. Sposób działania układu regulacji temperatury sondy pomiarowej mikroskopu bliskiego pola termicznego przedstawiono na rysunku 33.



Rys. 33. Przebieg regulacji termosondy pomiarowej – sygnał na przekątnej mostka pomiarowego, sygnał uchybu regulacji i sygnał wyjściowy regulatora
 Fig. 33. Controlling the temperature of the thermal probe
 – signals representing the measurement bridge output and PID error and output

W przeprowadzonym eksperymencie ostrze pomiarowe było w kontakcie z powierzchnią litego złota. Pomiar temperatury ostrza był przeprowadzany podczas zasilania mostka prądem o częstotliwości 10 kHz, częstotliwość sygnału grzewczego wynosiła natomiast 1 kHz. Warto jednak zwrócić uwagę, że sygnał grzewczy może być modulowany, dzięki czemu uzyskać można modulację długości fali, jaka jest prowadzona w badanej strukturze. W opisywanym eksperymencie wartość zadaną temperatury sondy modulowano w zakresie 30–40 °C. Na rysunku 33 pokazano przebieg sygnału z wyjścia mostka, w którym można obserwować sygnał grzewczy i korespondujący z temperaturą mikrosondy. Przebieg nr 2 pokazuje natomiast sygnał uchybu regulacji. Z jego analizy wynika, że w fazie przejściowej dochodzi do zaburzenia, które po około 20 ms zostaje zrównoważone.

Na rysunku 33 przedstawiono też sygnał odpowiadający uchybowi regulacji sygnałem wyjściowym, który steruje sygnałem grzewczym. Sygnał ten zmienia warunki zasilania mostka z sondą oporową w taki sposób, aby odpowiedź bloku pomiaru temperatury końcówki była równa wartości zadanej (górna część rys. 33). Na podstawie zarejestrowanego przebiegu wyznaczyć można stałą całkowania regulatora, która jest równa 60 ms. Sygnał wyjściowy regulatora steruje mnożnikiem, który jest z kolei zasilany z generatora sygnału grzewczego (rys. 32).

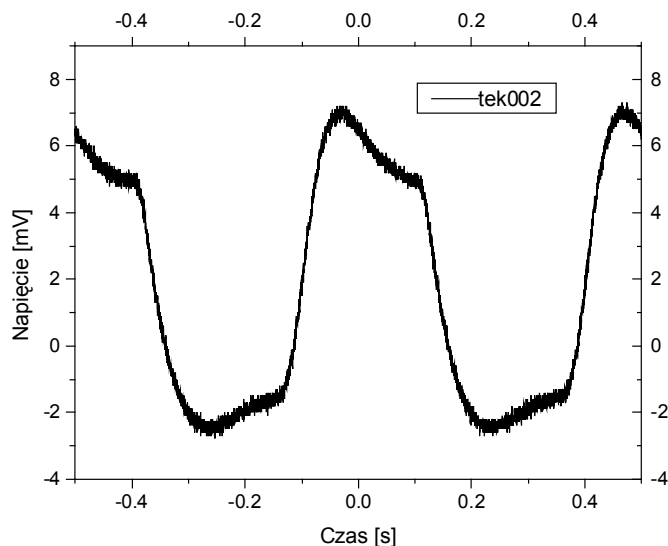


Rys. 34. Zmiany prądu zasilającego sondę Wollastona odpowiadające zmianom temperatury o 5 K

Fig. 34. Changes in Wollaston probe current corresponding to temperature changes of 5 K

Rejestracja w czasie skanowania sygnałów odpowiadających prądowi przepływającemu przez sondę i napięciu odkładanemu na sondzie oporowej umożliwia wyznaczenie mocy, jaka jest potrzebna dla utrzymania temperatury czujnika na zadanym poziomie (rys. 34, 35). W przypadku pomiarów próbek o małej przewodności cieplnej

i po założeniu, że warunki skanowania nie zmieniają się w czasie skanowania powierzchni ze zmian mocy, jaka jest dostarczana do sondy, można wnioskować o zmianach przewodności cieplnej materiałów tworzących tę powierzchnię.



Rys. 35. Zmiany napięcia odkładanego na sondzie Wollastona odpowiadające zmianom temperatury o 5 K

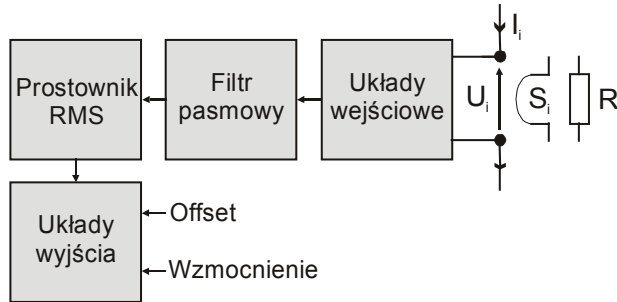
Fig. 35. Changes in Wollaston probe voltage corresponding to temperature changes of 5 K

Struktura układów elektronicznych umożliwiających ocenę mocy dostarczanej do mikrosondy ma dla wiarygodnych badań powierzchni znaczenie bardzo duże. Podstawową strukturę takich układów przedstawiono na rysunku 36. Są one połączeniem wzmacniaczy małoszumowych z zespołami filtrów pasmowoprzepustowych. Oto podstawowy opis konstrukcji tych podzespołów:

- układy wejściowe pracujące w konfiguracji wzmacniaczy różnicowych sprzęganych zmiennoprądowo z sondą oporową lub posobnikiem do pomiaru prądu; układy te konfiguruje się zwykle w celu zapewnienia możliwie dużej wartości współczynnika tłumienia składowej sumacyjnej (ok. 100 dB) i małego prądu polaryzacji (ok. 10 pA);
- filtr środkowoprzepustowy dostrojony do częstotliwości sygnału grzewczego o dobroci ustalonej w zakresie 5–10;
- prostownik sygnału pracujący w paśmie do 2 MHz;
- układy wyjściowe zawierające filtry dolnoprzepustowe, układy odejmowania składowej stałej i wzmacniacze końcowe o regulowanym wzmacnieniu.

Przyjęta struktura bloku umożliwia prostą adaptację systemu do pomiaru sygnałów o różnych częstotliwościach i zastępuje złożone układy fazoczułe wzmacniaczy *lock-in*.

Zdolność rozdzielcza pomiaru mocy w opisywanym układzie wynosi zwykle około 100 pW.



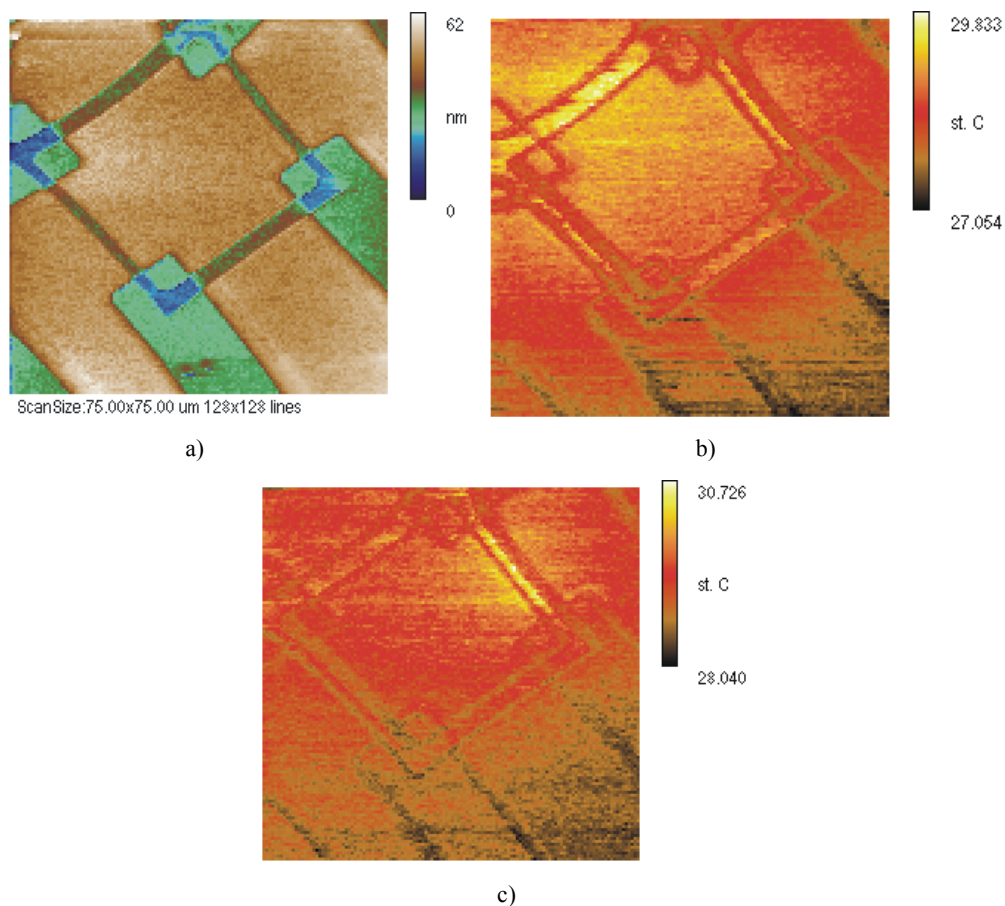
Rys. 36. Struktura toru pomiaru prądu i napięcia aktywnego mikroskopu bliskiego pola termicznego
 Fig. 36. Schematic of the probe current and probe voltage measurement
 in used in active-mode scanning thermal microscopy

6. WYNIKI EKSPERYMENTALNE

Statyczna mikroskopia bliskiego pola termicznego

Statyczny mikroskop bliskiego pola termicznego zastosowano do pomiarów rozkładu temperatury na powierzchni układu mikroelektronicznego tworzonoego przez 4 dyfundowane rezystory [51]. Układ mostka był zasilany ze źródła prądowego. Rezystancja piezorezystorów wynosiła 2,7 k Ω . Na rysunku 37a przedstawiono wynik pomiaru topografii powierzchni mostka 4 rezystorów na podłożu krzemowym (pole skanowania wynosiło w tym przypadku 75 \times 75 μm). Widoczne są mostek rezystorów i obszary tzw. głębokiej dyfuzji, za pomocą których doprowadzane są sygnały elektryczne do struktury mostka piezorezystywnego.

Na rysunku 37b przedstawiono wynik pomiaru temperatury w przypadku, gdy mostek był zasilany symetrycznie, tzn. przez każdy z rezystorów mostka przepływał prąd 5 mA (co korespondowało z mocą rozpraszaną w pojedynczym rezystorze na poziomie 67 mW). Zauważalny jest w tym przypadku równomierny i występujący w obszarach wszystkich piezorezystorów wzrost temperatury o około 1,8 K. Na rysunku 37c przedstawiono wyniki pomiaru temperatury w przypadku, gdy tylko przez jeden piezorezystor przepływał prąd 5 mA, natomiast trzy pozostałe – połączone szeregowo – były zasilane prądem o natężeniu równym 1,6 mA. Zaobserwowano w tym przypadku wzrost temperatury o 1,8 K jedynie w obszarze rezystora zasilanego największym prądem. Przeprowadzone eksperymenty wykazały możliwość zastosowania modularnego mikroskopu bliskiego pola termicznego do diagnostyki zachowań termicznych dyskretnych i scalonych elementów mikroelektronicznych.

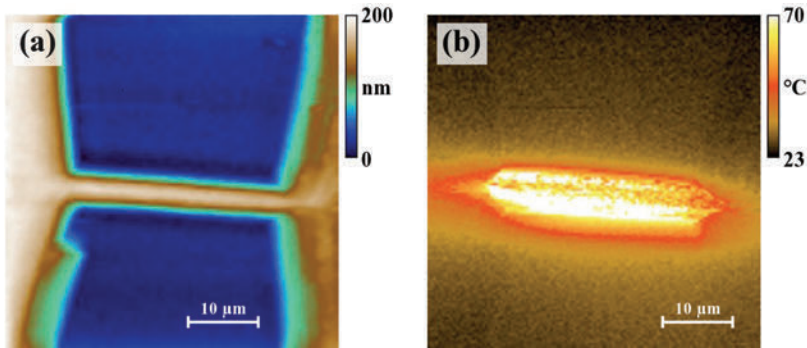


Rys. 37. Pomiar rozkładu temperatury na powierzchni piezorezystywnego detektora ugięcia belki sprężystej: a) topografia struktury, b) rozkład temperatury w przypadku, gdy każdy rezystor jest zasilany prądem 5 mA, c) rozkład temperatury w przypadku, gdy tylko jeden piezorezystor jest zasilany prądem 5 mA, przez trzy pozostałe przepływa natomiast prąd 1,6 mA [52]

Fig. 37. Measurement of the surface temperature of a piezoresistive deflection sensor: a) topography, b) temperature map recorded for equal distribution of current (5 mA for each of the resistors), c) temperature map for non-equal distribution of current (one resistor 5 mA, the other three – 1.6 mA) [52]

Inny przykład pomiarów prowadzonych za pomocą statycznego mikroskopu bliiskiego pola termicznego pokazano na rysunku 38a i b. W badaniach zastosowano sondę pomiarową z rysunku 9. Na rysunku 38a przedstawiono topografię powierzchni bezpiecznika polikrzemowego, stosowanego w układach mikro- i nanoelektronicznych [49]. Struktura polikrzemowa zawiera ścieżkę przewodzącą i dwie elektrody do połączenia z układami zewnętrznymi. Całość jest osadzona na podłożu z dwutlenku krzemu, co zapewnia izolację optyczną i elektryczną układu. Z obrazu topografii wynika, że wysokość struktury wynosi około 200 nm. Przez strukturę mikrobezpiecznika prze-

puszczono prąd roboczy równy 25 mA, który jest zbliżony do maksymalnego prądu, jaki może przenieść ta struktura. W obrazie rozkładu temperatury widać, że temperatura wzrasta na środku struktury bezpiecznika do około 70 °C (rys. 38b). Widoczne jest poszerzenie obrazu w stosunku do topografii, co wynika z rozplywu temperatury wokół przewodnika bezpiecznika. Warto wskazać, że stosunkowo duży kontrast obrazu termicznego wynika z dużego kontrastu między przewodnościami cieplnymi podłoża z dwutlenku krzemu i polikrzemu. Powoduje to jedynie ograniczone rozszerzenia się obszaru o podwyższonej temperaturze, które w przypadku materiałów o większej przewodności cieplnej miałyby większy zakres.



Rys. 38. Statyczna mikroskopia bliskiego pola termicznego – pomiar topografii i temperatury bezpiecznika polikrzemowego na podłożu z dwutlenku krzemu [49]

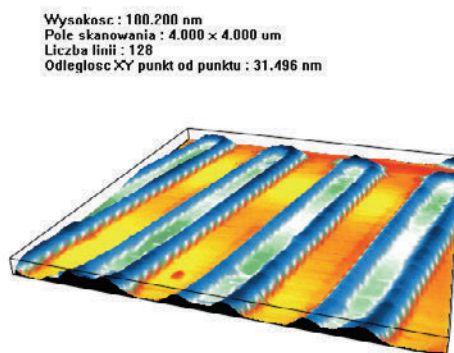
Fig. 38. Static scanning thermal microscopy – topography and temperature of a polysilicon microfuse on silicon dioxide substrate [49]

Aktywna mikroskopia bliskiego pola termicznego

Na rysunku 39 przedstawiono topografię powierzchni struktur platynowych, które są osadzone na utlenionym podłożu krzemowym. Szerokość linii platynowej wynosi 1 µm, natomiast jej wysokość jest równa 80 nm (rys. 40). Podwyższenia struktury, widoczne w okolicy jej ścianek bocznych, są związane z technologią *lift off* osadzania warstw metalicznych i są pozostałością platyny po przeprowadzonym procesie elektronolitografii i trawienia. Opisane struktury zastosowano do testowania zdolności rozdzielczej skaningowego mikroskopu bliskiego pola termicznego.

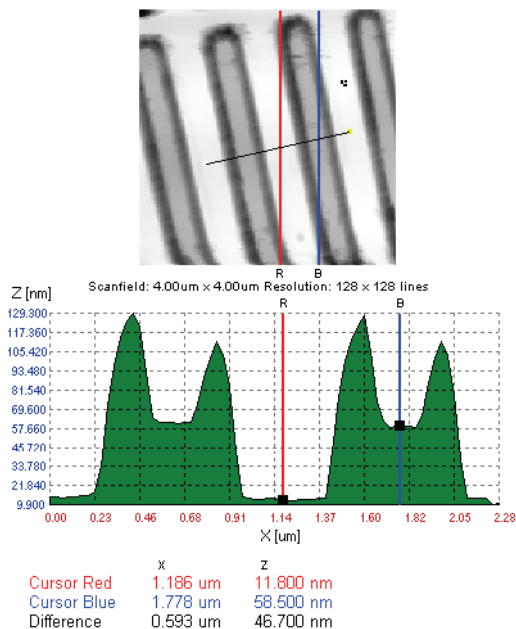
Wynik pomiaru topografii struktur testowych, przeprowadzonego za pomocą mikroskopu termicznego, jest przedstawiony na rys. 41. Na jakość obrazu topografii powierzchni, uzyskaną za pomocą mikroskopu termicznego wpływa w tym przypadku większy promień końcówki ostrza pomiarowego sondy Wollastona. Ostrze to jest bowiem na tyle duże, że nie jest w stanie wniknąć w obszary między liniami i ściankami bocznymi pasków platynowych. Na rysunku 42 zaprezentowano natomiast wy-

niki pomiaru obrazu rezystancji termicznych struktury. Temperatura ostrza była stabilizowana na poziomie 335 K. W obszarach linii platynowych zaobserwowano mniejszą rezystancję cieplną aniżeli w polach podłoża krzemowego (różnica mocy koniecznej do utrzymania tej samej temperatury ostrza wynosiła 67 μW).



Rys. 39. Topografia struktur platynowych osadzonych na powierzchni krzemu (pole skanowania $4 \times 4 \mu\text{m}$)

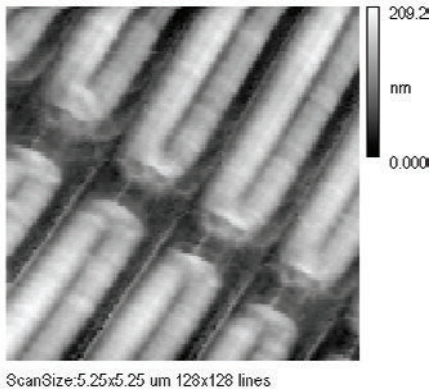
Fig. 39. Topography of platinum structures on silicon substrate (scan area: $4 \times 4 \mu\text{m}$)



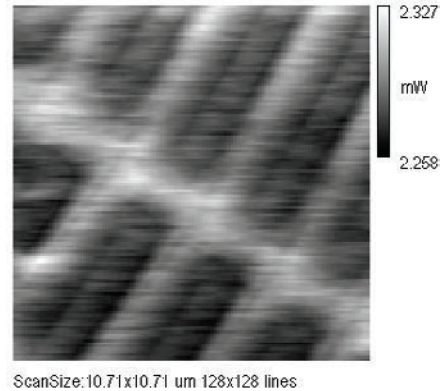
Rys. 40. Wysokość struktur platynowych zmierzona za pomocą skaningowego mikroskopu sił atomowych

Fig. 40. Height of Platinum structures measured using atomic force microscopy

Przeprowadzone pomiary wykazały, że termiczna i lokalna zdolność rozdzielcza skonstruowanego mikroskopu umożliwia odwzorowanie właściwości termicznych układu platynowa linia paskowa–podłoże krzemowe w skali submikrometrowej [53].

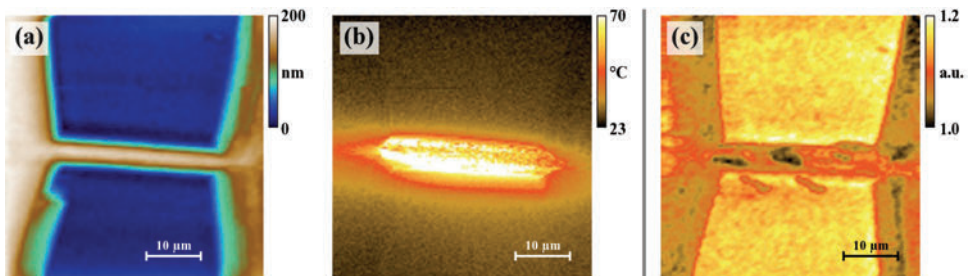


Rys. 41. Topografia powierzchni struktur platynowych osadzonych na podłożu krzemowym zarejestrowana za pomocą skaningowego mikroskopu termicznego bliskiego pola
Fig. 41. Topography of platinum structures on silicon substrate recorded using scanning thermal microscopy



Rys. 42. Obraz rezystancji cieplnych struktur platynowych osadzonych na podłożu krzemowym – obraz zarejestrowany za pomocą skaningowego mikroskopu termicznego bliskiego pola
Fig. 42. Image of thermal resistance of platinum structures on silicon substrate recorded using scanning thermal microscopy

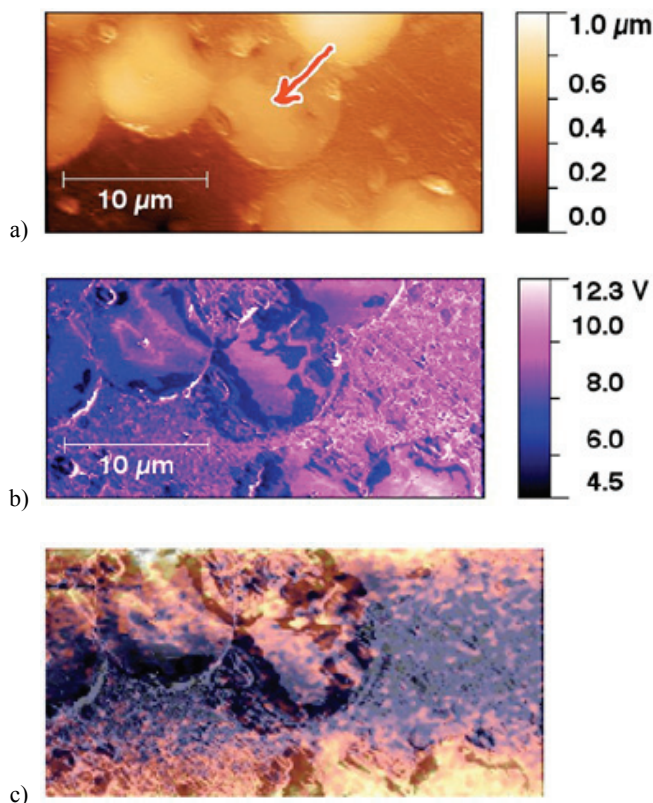
Przedstawione wyniki pomiarów prowadzonych za pomocą aktywnego mikroskopu bliskiego pola termicznego związane były z pomiarami struktur, w których przewodność cieplna metalizacji była mniejsza od przewodności cieplnej podłoża (przewodność cieplna platyny wynosi 71 W/(mK), przewodność cieplna krzemu zaś jest równa 140 W/(mK)).



Rys. 43. Pasywna i aktywna mikroskopia bliskiego pola elektrycznego na strukturach bezpieczników polikrzemowych [49]

Fig. 43. Passive- and active-mode scanning thermal microscopy on the surface of polysilicon microfuses [49]

Aktywna mikroskopia bliskiego pola termicznego została również zastosowana do badania struktur bezpieczników polikrzemowych, które uległy uszkodzeniu po przepuszczeniu przez nie prądu o dużym natężeniu (w praktyce większym niż 35 mA dla struktur z rys. 38). W obrazie rezystancji cieplnej ujawniają się w tym przypadku struktury, których charakterystyka różni się od struktur otaczających je (rys. 43). Należy oczekiwać, że w tym przypadku przepalenie bezpiecznika nastąpiło w miejscu o zmniejszonej przewodności.



Rys. 44. Obrazowanie powierzchni ziaren metalicznych zanurzonych w matrycy z żywicy epoksydowej: a) topografia, b) obraz sił tarcia, c) kontrast termiczny

Fig. 44. Imaging of metallic grains in epoxy resin matrix:

a) topography, b) image of lateral forces, c) thermal contrast

Analizując wyniki uzyskane w trybie aktywnej i statycznej mikroskopii bliskiego pola termicznego warto wskazać na zwiększoną rozdzielczość obrazowania rezystancji cieplnej. Wynika to zasadniczo z dużej dynamiki różnic w przewodności cieplnej materiałów stosowanych w nanotechnologii i technice mikrosystemów. Przewodność cieplna takich materiałów, która bezpośrednio wpływa na rezystancję cieplną struktury

ry, zmienia się od ułamków do pojedynczych tysięcy W/(mK) (odpowiednio dla polimerów i warstw diamentowych oraz grafenowych). Powoduje to, że kontrast obrazów uzyskiwanych w aktywnym trybie mikroskopii bliskiego pola termicznego jest zdecydowanie większy aniżeli w wypadku badań prowadzonych w trybie statycznym.

Bardzo istotnym problemem współczesnej mikroskopii bliskiego pola termicznego są zagadnienia kalibracji układów pomiarowych. W wypadku aktywnej mikroskopii bliskiego pola termicznego do kalibracji odpowiedzi układu stosować można preparaty zawierające w sobie obiekty, których przewodność cieplna będzie się istotnie różnić. Mogą to być np. struktury metaliczne ziarna o wymiarach rzędu setek nanometrów zatopione w żywicy epoksydowej. W idealnym przypadku struktury takie powinny być możliwie płaskie, aby zapewnić w miarę jak najmniejszy wpływ topografii obrazu na wynik pomiaru właściwości termicznych. Na rysunkach 44a, b i c przedstawiono wynik pomiaru ziaren miedzi zanurzonych w matrycy żywicy epoksydowej. W obrazie topografii zaznaczono ziarno metalizacji miedzianej, którego średnica wynosi około 1 μm . Powierzchnia struktury była jedynie wstępnie polerowana (rys. 44a), co jest powodem znacznej chropowatości powierzchni. Badania prowadzono za pomocą sondy oporowej o podniesionej temperaturze. Należy oczekiwać, że ostrze ulegało zatopieniu w miękkim podłożu żywicy, co było źródłem stosunkowo dużego tarcia rejestrowanego w czasie skanowania powierzchni (rys. 44b). Kontrast termiczny obrazujący wzrost energii rozpraszanej w strukturze pokazano na rysunku 44c. Przeprowadzone eksperymenty wykazały, że do uzyskania wiarygodnych obrazów konieczne będzie dodatkowe przeprowadzenie starannego polerowania powierzchni i diagnostyki powierzchni prowadzonej metodami skaningowej mikroskopii elektronowej i spektroskopii z dyspersją energii.

7. PODSUMOWANIE

Przedstawione analizy pokazują, że mikroskopia bliskich oddziaływań jest atrakcyjnym narzędziem do prowadzenia badań mikro- i nanostruktur. Jej ogromnymi zaletami jest duża rozdzielczość obrazowania i możliwość obserwacji efektów kwantowych. Z perspektywy postępu w nanotechnologii i techniki mikrosystemów technologii te będą na pewno coraz powszechniej stosowane. Oceniamy, że w badaniach zachowań termicznych mikro- i nanostruktur mikroskopia bliskiego pola termicznego będzie znajdować szerokie zastosowanie. Dla postępu badań metrologicznych prowadzonych tą technologią konieczny będzie jednak rozwój nowych technik ostrzowych sond pomiarowych oraz technologii eksperymentalnych z nimi powiązanych. Działania te powinny prowadzić do ścisłego opisu kontaktu ostrza z badaną powierzchnią, co umożliwi ilościową ocenę rezystancji termicznej między ostrzem a próbką i przewodnością cieplną preparatu. Zrozumienie tych zjawisk ułatwi również badania właściwości cieplnych mikro- i nanostruktur w głębi struktur. W przypadku układów wielowar-

stwowych otworzy to absolutnie nowe możliwości zastosowania technik mikroskopii bliskiego pola termicznego będącej w istocie metodą badawczą przeznaczoną do badania powierzchni obiektu.

PODZIĘKOWANIA

Przedstawione prace badawcze były m. in. wspierane ze środków Fundacji na rzecz Nauki Polskiej (FNP) w programie Mistrz *Nanometrologia z zastosowaniem zaawansowanych metod mikroskopii bliskich oddziaływań (NanoMetSPM)*.

LITERATURA

- [1] GOTSZALK T., RANGELOW I.W., DUMANIA P., GRABIEC P., *Cantilever with integrated Wheatstone bridge piezoresistive deflection sensor – analysis of vertical/lateral force and displacement sensitivity*, Micromachining and Microfabrication Process Technology II, Proc., 1996, SPIE 2880, 264.
- [2] BINNIG G., ROHRER H., WEIBEL E., *Surface Studies by Scanning Tunneling Microscopy*, Phys. Rev. Lett., 1982, 49.
- [3] BINNIG G., GERBER C., QUATE C., *Atomic force microscope*, Phys. Rev. Lett., 1986, 56, 9, 930.
- [4] WILLIAMS C.C., WICKRAMASINGHE H.K., *Scanning thermal profiler*, Appl. Phys. Lett., 1986, 49, 1587–1589.
- [5] NONNENMACHER M., WICKRAMASINGHE H.K., *Scanning probe microscopy of thermal conductivity and subsurface properties*, Appl. Phys. Lett., 1992, 61, 168–170.
- [6] CAHILL D., GOODSON K., MAJUMDAR A., *Thermometry and thermal transport in micro/nano-scale solid-state devices and structures*, J. of Heat Transfer, 2002, 124, 223.
- [7] PADDOCK C., EESLEY G., *Transient thermoreflectance from thin metal films*, J. Appl. Phys., 1986, 60, 285.
- [8] CLEMENS B., EESLEY G., PADDOCK C., *Time-resolved thermal transport in compositionally modulated metal films*, Phys. Rev. B, 1988, 37, 1085.
- [9] BRUGGER H., EPPERLEIN P., *Mapping of local temperatures on mirrors of GaAs/AlGaAs laser diodes*, Appl. Phys. Lett., 1990, 56, 1049.
- [10] OSTERMEIER R., BRUNNER K., ABSTREITER G., WEBER W., *Temperature distribution in Si-MOSFET's studied by micro Raman spectroscopy*, IEEE Trans. Electron Devices, 1992, 39, 858.
- [11] IWATA K., HAMAGUCHI H., *Microscopic mechanism of Solute solvent energy dissipation probed by picosecond time resolved Raman spectroscopy*, J. Phys. Chem., 1997, 101, 4, 632.
- [12] SZELOCH R.F., GOTSZALK T.P., RADOJEWSKI J., JANUS P., ORAWSKI W., PEĐRAK R., *Thermographic experiments in far – and – near – fields*, Materiały VII Konferencji Naukowej ELTE 2000, 1198.
- [13] CAHILL D., *Thermal conductivity measurement from 30 K to 750 K: The 3 ω method*, Rev. Sci. Instrum., 1994, 61, 802.
- [14] CAHILL D., ALLEN T., *Thermal conductivity of sputtered and evaporated SiO₂ and TiO₂ optical coatings*, Appl. Phys. Lett., 1994, 65, 3, 309.
- [15] LEE S.M., CAHILL D., *Heat transport in thin dielectric films*, J. Appl. Phys., 1997, 81, 6, 2590.
- [16] KURABAYASHI Y., GOODSON K., *Thermal characterization of anisotropic thin dielectric films using harmonic Joule heating*, Thin Solid Films, 1999, 339, 160.
- [17] HUXTABLE S., ABRAMSON A., TIEN Ch., MAJUMDAR A., LaBOUNTY Ch., FAN X., ZENG G., BOWERS J., SHAKOURI A., CROKE E., *Thermal conductivity of Si/SiGe and SiGe/SiGe superlattices*, Appl. Phys. Lett., 2002, 80, 10, 1737.

- [18] JU Y., GOODSON K., *Phonon scattering in silicon films of thickness below 100 nm*, Appl. Phys. Lett., 1999, 74, 20, 3005.
- [19] MOON K., JEONG Y., KWUN S., *The 3 technique for measuring dynamic specific heat and thermal conductivity of a liquid or solid*, Rev. Sci. Instrum., 1996, 67, 29.
- [20] LUO K., SHI Z., VARESI J., MAJUMDAR A., *Sensor nanofabrication performance and conduction mechanism in scanning thermal microscopy*, J. Vac. Sci. Technol., 1997, B15, 349.
- [21] STOPKA M., HADJIISKI L., OESTERSCHULZE E., KASSING R., *Surface investigations by scanning thermal microscopy*, J. Vac. Sci. Technol., 1995, B13, 2153.
- [22] LUO K., SHI Z., VARESI J., MAJUMDAR A., *Sensor nanofabrication performance and conduction mechanism in scanning thermal microscopy*, J. Vac. Sci. Technol., 1997, B15, 349.
- [23] LUO K., SHI Z., LAI J., MAJUMDAR A., *Nanofabrication of sensors on cantolever probe tips for scanning multiprobe microscopy*, Appl. Phys. Lett., 1996, 68, 325.
- [24] LUO K., SHI Z., LAI J., MAJUMDAR A., *Nanofabrication of sensors on cantolever probe tips for scanning multiprobe microscopy*, Appl. Phys. Lett., 1996, 68, 325.
- [25] KIM P., SHI L., MAJUMDAR A., McEUEN P., *Mesoscopic thermal transport and energy dissipation in carbon nanotube*, Physica B, 2002, 323, 67.
- [26] SHI L., PLYASUNOV S., BACHTOLD A., McEUEN P., MAJUMDAR A., *Scanning thermal microscopy of carbon nanotubes using batch fabricated probes*, Appl. Phys. Lett., 2000, 77, 4295.
- [27] RUIZ F., SUN W., POLLAK F., VENKATRAMM C., *Determination of thermal conductivity of diamond-like nanocomposite films using a scanning thermal microscope*, Appl. Phys. Lett., 1998, 73, 1802.
- [28] RANGELOW I.W., GOTSZALK T., ABEDINOV N., GRABIEC P., EDINGER K., *Thermal nano-probe*, Microelectronic Engineering, 2001, 57–58, 737–748.
- [29] EDINGER K., GOTSZALK T., RANGELOW I.W., *Novel high resolution scanning thermal probe*, J. Vac. Sci. Technol., 2001, B 19 (6), 2856.
- [30] KOOPS H.W., SCHLOESSER C., KAYA A., *Conductive dots wires, and supertips for field electron emitters produced by electron-beam induced deposition on samples having increased temperature*, J. Vac. Sci. Technol., 1996, B14, 6, 4105.
- [31] JANUS P., SZMIGIEL D., WEISHEIT M., WIELGOSZEWSKI G., RITZ Y., GRABIEC P., HECKER M., GOTSZALK T., SULECKI P., ZSCHECH E., *Novel SThM nanoprobe for thermal properties investigation of micro- and nanoelectronic devices*, Microelectronic Engineering, 2010, 87, 1370–1374.
- [32] KITTEL Ch., *Wstęp do fizyki ciała stałego*, PWN, Warszawa 1999.
- [32] CAHILL D., GOODSON K., MAJUMDAR A., *Thermometry and thermal transport in micro/nanoscale solid-state devices and structures*, J. of Heat Transfer, 2002, 124, 223.
- [33] GOMES S., TRANNOY N., GROSSEL P., *DC thermal microscopy: study of the thermal exchange between a probe and a sample*, Meas. Sci. Technol., 1999, 10, 805–811.
- [34] LEFEVRE S. et al., *Thermal conductivity calibration for hot wire based dc scanning thermal microscopy*, Rev. Sci. Instrum., 2003, 74, 2014–2023.
- [35] LEFEVRE S., VOLZ S., *3 ω -scanning thermal microscope*, Rev. Sci. Instrum., 2005, 76, 033701.
- [36] WIELGOSZEWSKI G., SULECKI P., GOTSZALK T., JANUS P., SZMIGIEL D., GRABIEC P., ZSCHECH E., *Microfabricated resistive high-sensitivity nanoprobe for scanning thermal microscopy*, J. Vac. Sci. Technol., 2010, B 28, C6N7–C6N11.
- [37] JÓZWIAK G., WIELGOSZEWSKI G., GOTSZALK T., KĘPIŃSKI L., *Thermal mapping of a scanning thermal microscopy tip*, Ultramicroscopy, 2013, 133, 80–87.
- [38] JÓZWIAK G., HENRYKOWSKI A., MASALSKA A., GOTSZALK T., *Regularization mechanism in blind tip reconstruction procedure*, Ultramicroscopy, 2012, 118, 1–10.

- [39] VILLARUBIA J.S., *Algorithms for scanned probe microscope image simulation, surface reconstruction and tip estimation*, J. Res. Natl. Inst. Stand. Technol., 1997, 102, 425–454.
- [40] TIAN F., QIAN X., VILLARUBIA J.S., *Blind estimation of general tip shape in AFM imaging*, Ultramicroscopy, 2008, 109, 44–53.
- [41] JOULAIN K., *Physical processes describing heat transfer at short lengthscale*, 5th European Thermal-Sciences Conference, The Netherlands, 2008.
- [42] KIM K., CHUNG J., HWANG G., KWON O., LEE J.S., *Quantitative measurements by scanning thermal microscope preventing distortion due to the heat transfer through the air*, ACS Nano, 2011, 5, 8700–8709.
- [43] NIERADKA K., GOTSZALK T., SCHROEDER G., *A novel method for simultaneous readout of static bending and multimode resonance-frequency of microcantilever-based biochemical sensors*, Sensors and Actuators B: Chemical, 2012, 170, 172–175.
- [44] GOTSZALK T., GRABIEC P., RANGELOW I.W., *Calibration and examination of piezoresistive Wheatstone bridge cantilevers for scanning probe microscopy*, Ultramicroscopy, 2003, 97, 385.
- [45] RANGELOW I. W., GRABIEC P., GOTSZALK T., EDINGER K., *Piezoresistive SXM Sensors*, Surface and Interface Analysis, 2002, 33, 234.
- [46] GOTSZALK T., GRABIEC P., RANGELOW I., *Piezoresistive sensors for scanning probe microscopy*, Ultramicroscopy, 2000, 82, 39.
- [47] PEDRAK R., IVANOV T.Z.V., IVANOVA K., GOTSZALK T., ABEDINOV N., RANGELOW I.W., EDINGER K., TOMEROV E., SCHENKEL T., HUDEK P., *Micromachined atomic force microscopy sensor with integrated piezoresistive sensor and thermal bimorph actuator for high-speed tapping-mode atomic force microscopy phase-imaging in higher eigenmodes*, J. Vac. Sci. Technol., 2003, B 21, 3102.
- [48] RANGELOW I.W., IVANOV T.Z.V., IVANOVA K., VOLLAND B.E., GRABIEC P., SAROV Y., PERSAUD A., GOTSZALK T., ZAWIERUCHA P., ZIELONY M., DONTZOV D., SCHMIDT B., ZIER M., NIKOLOV N., KOSTIC I., ENGL W., SULZBACH T., MIELCZARSKI J., KOLB S., LATIMIER D.U.P. et al., *Piezoresistive and self-actuated 128-cantilever arrays for nanotechnology applications*, Microelectronic Engineering, May–August 2007, 84, 5–8, 1260–1264.
- [49] WIELGOSZEWSKI G., SULECKI P., GOTSZALK T., JANUS P., GRABIEC P., HECKER M., RITZ Y., ZSCHECH E., *Scanning thermal microscopy: A nanoprobe technique for studying the thermal properties of nanocomponents*, Phys. Status Solidi B, 2011, 248, 2.
- [50] WIELGOSZEWSKI G., SULECKI P., JANUS P., GRABIEC P., ZSCHECH E., GOTSZALK T., *A high-resolution measurement system for novel scanning thermal microscopy resistive nanoprobe*, Measurement Science and Technology, 2011, 22, 9, 094023.
- [51] SZELOCH R.F., GOTSZALK T.P., JANUS P., *Scanning Thermal Microscopy in Microsystem Reliability Analysis*, Microelectronics Reliability, 2002, 42, 1719.
- [52] SZELOCH R., GOTSZALK T., JANUS P., *Thermal characterization of micro-devices with far and near field microscopy*, Opt. Appl., 2003, 33, 4, 669.
- [53] KAMINSKA E., GOLASZEWSKA K., PIOTROWSKA A., KUCHUK A., KRUSZKA R., PAPIS E., SZELOCH R., JANUS P., GOTSZALK T., BARCZ A., *Study of long-term stability of ohmic contacts to GaN*, Physica Status Solidi (c), 2004, 1, 2, 219.

SCANNING THERMAL NEARFIELD MICROSCOPY IN MICRO- AND NANOSTRUCTURE METROLOGY

Scanning thermal nearfield microscopy (SThM) is one of the scanning probe microscopy (SPM) technologies, in which interaction between the nanotip and the investigated surface. In the SThM experiments the thermal nanotip can be used as a thermometer or a local heat source. In this way investigations

of temperature distribution or local thermal resistance are enabled at nanoscale. In this work application of the SThM technology in metrological experiments are described. The aim of the performed research is to describe surface thermal properties in a quantitative way, which from the point of view of the progress in nanotechnology is of crucial importance. This includes description of the technology theoretical and experimental basics, analysis of applied thermal nanotip probes and its characterization. In addition results obtained using advanced SThM methods on micro- and nanostructures fabricated using modern nanoelectrical technologies will be presented.

Sławomir CIĘSZCZYK*

WYBRANE ZAGADNIENIA POMIARU ZAWARTOŚCI GAZÓW W ZMIENIAJĄCYCH SIĘ WARUNKACH

W pracy przedstawiono zagadnienie pomiaru zawartości gazów w zmiennych warunkach środowiskowych. Pojawiają się wówczas interferencje i zmieniają się czynniki wpływające na pomiar. Zagadnienia te przedstawiono w aspekcie układów czujników inteligentnych i wirtualnych, a także ich fuzji. Zaprezentowano metody budowania modeli służących do wyznaczania zawartości gazów, mające właściwość odporności na wpływ warunków pomiaru. W dalszej części rozdziału przedstawiono zagadnienia kalibracji wielowymiarowej, transferu kalibracji oraz korekcji widm dla pomiarów spektrometrycznych dla zmiennej temperatury ścieżki pomiarowej.

1. WPROWADZENIE

Zdobywanie wiedzy o otaczającym nas świecie wymaga coraz doskonalszych instrumentów pomiarowych. W większości przypadków są one przystosowane do pomiarów w laboratorium. Nie wszystkie techniki nadają się do przeniesienia w warunki przemysłowe czy też do analizy środowiska. Ciągłe monitorowanie różnorodnych procesów i środowiska jest coraz częściej wymagane ze względów prawnych. Dzięki temu ogranicza się ryzyko narażenia ludzi na wpływ szkodliwych substancji przez zmniejszenie emisji zanieczyszczeń. Z drugiej strony wymagania ekonomiczne wymuszają polepszenie jakości prowadzenia procesów. Niestety zarówno systemy przemysłowe, jak też środowisko naturalne podlegają wielu interakcjom, co utrudnia wykonanie pomiarów monitorujących i diagnostycznych. Oprócz czysto sprzętowego aspektu pomiarów w procesach technologicznych i atmosferze, kluczową w pozyskiwaniu informacji jest odpowiednia analiza danych pomiarowych i zaawansowane narzędzia matematyczne. Dla danych pomiarowych, otrzymanych z pomiarów środowiska atmosferycznego oraz

* Instytut Elektroniki i Technik Informacyjnych, Politechnika Lubelska, ul. Nadbystrzycka 38D, 20-619 Lublin.

w przypadku pomiarów procesów przemysłowych, klasyczne metody analizy danych nie dają dobrych rezultatów. Wynika to wprost z ograniczonej możliwości zbudowania odpowiedniego modelu matematycznego i wykonania jego kalibracji (wzorcowania). Wiąże się to z trudnością prowadzenia efektywnych i ekonomicznych pomiarów reprezentatywnych próbek kalibracyjnych. W niektórych warunkach taki pomiar kalibracyjny w ogóle nie jest możliwy. Stąd wyznaczanie informacji oparte jest często na modelach symulacyjnych i rozwiązywaniu zagadnienia odwrotnego opartego na tych modelach [1–6]. W tym przypadku wymagane jest przeprowadzenie wcześniejszej analizy rozpatrywanych zjawisk oraz możliwych do zastosowania algorytmów.

W rozdziale drugim przedstawione zostały podstawowe parametry czujników, służące do opisu ich właściwości. Określony został wpływ różnorodnych czynników na powstawanie dryftu i zmiany parametrów czujników. Następnie pokazano metody korekcji dryftu dla pojedynczych czujników oraz macierzy czujników służących do pomiaru zawartości gazów. Do przeprowadzenia korekcji wielkości wyjściowej czujników niezbędne jest przetwarzanie sygnałów oraz kalibracja adaptacyjna. Ułatwia to zbudowanie inteligentnych czujników umożliwiających uwzględnienie aktualnych warunków pomiaru i wartości wielkości wpływających.

Problemy budowy modeli odwrotnych według różnych technik i pomiarów empirycznie przedstawiono w rozdziale trzecim. Dla pomiarów w trudnych warunkach modele takie uwzględniać muszą znaczną liczbę zależności. W niektórych warunkach określenie wartości wielkości procesowych możliwe jest jedynie na podstawie pomiarów innych wielkości. Modele odwrotne czujników wirtualnych również mogą podlegać dryftowi i wymagać adaptacji. Brak selektywności oraz pojawiające się czułości skrośne powodują konieczność zastosowania dodatkowych czujników oraz metod analizy danych, w tym budowy modeli z fuzją danych czujników. W niektórych przypadkach pojedynczy model odwrotny, nawet bardzo skomplikowany, nie może spełnić założonych wymagań. Zagadnienie takie rozwiązać można przez podział obszaru sygnałów wejściowych i zbudowanie dla każdego z nich osobnego modelu. Alternatywnym, bardzo popularnym, podejściem, nazywanym popularnie metodą ensemble, jest budowa zespołu modeli i odpowiednia ich fuzja.

W części czwartej przedstawione zostały najpopularniejsze metody budowania modeli wielowymiarowych w spektroskopii. Metoda regresji cząstkowej PLS jest często używana i modyfikowana do celów budowania modeli odpornych na zmianę warunków pomiarowych. Różnorodne jej modyfikacje stosowane są także do modelowania nieliniowości w danych spektrometrycznych. Rozdział ten jest podbudową do rozdziału piątego, w którym przedstawione zostały modele stosowane głównie w spektroskopii, mające na celu osiągnięcie wymaganej odporności na zmianę warunków zarówno instrumentalnych, jak też fizykochemicznych. W rozdziale tym dokonano podziału takich metod. Uwzględniono szczególnie transfer kalibracji. Jako osobna część opisane zostały metody ortogonalnej korekcji sygnału. Polegają one na usuwaniu z danych (szczególnie widm) sygnałów (części), za które odpowiedzialne są czynniki wpływające.

Ostatnia część opracowania dotyczy dwóch obszarów, w których wykorzystywane są przedstawione w poprzedniej części rozważania. Należy do nich technologia analityczna procesów PAT (ang. *Process Analytical Technology*) oraz pomiary optyczne o otwartej ścieżce pomiarowej.

2. CZUJNIKI, WIELKOŚCI MIERZONE, WYPTYWAJĄCE I INTERFERUJĄCE

Czujniki są urządzeniami, za pomocą których otrzymujemy informacje o otaczającym nas świecie, można je podzielić na następujące grupy:

- Parametryczne (modulujące, aktywne) – pod wpływem wielkości mierzonej następuje zmiana parametru elektrycznego, wymagają dodatkowego źródła energii, właściwości źródła zasilania mogą wpływać na parametry czujnika.
- Generacyjne (pasywne) – same generują sygnał wyjściowy bez dodatkowego źródła zasilania, na ich wyjściu powstaje prąd, ładunek lub napięcie elektryczne.
- Częstotliwościowe – wielkością wyjściową jest częstotliwość.
- Cyfrowe (kodowe) – sygnał wyjściowy jest przedstawiany bezpośrednio w postaci cyfrowej.

Nominalna statyczna charakterystyka przetwarzania jest to zależność wielkości wyjściowej czujnika od wielkości wejściowej przy braku oddziaływań wielkości wpływających oraz interferencji lub w warunkach, w których ich wartość jest ustalona. Model charakterystyki statycznej (przejściowej) może powstać w wyniku analizy zjawisk fizycznych lub też może być modelem empirycznym. W obydwu przypadkach niezbędne jest wykonanie określonej liczby pomiarów kalibracyjnych. Dla modelu stworzonego na podstawie analizy zjawisk fizycznych pomiary takie służą wzorcowaniu, czyli weryfikacji doświadczalnej modelu. W przypadku modelu empirycznego najpierw następuje identyfikacja modelu (rodzaj, struktura modelu), a następnie określenie jego parametrów (np. współczynników funkcji aproksymacyjnej).

Idealna charakterystyka przejściowa czujnika powinna spełniać następujące wymagania:

- liniowość,
- brak czułości skrośnej,
- brak wpływu czynników zakłócających,
- brak wpływu wyjścia na wejście (jeden kierunek oddziaływań),
- jedna wielkość fizyczna odpowiada jednemu wyjściu czujnika,
- statyczność (brak elementów dynamicznych).

Sygnał wyjściowy większości czujników nie zależy jedynie od sygnału wejściowego, który nas interesuje, ale również od innych sygnałów. W tym sensie sygnały oddziałujące na czujnik można podzielić na trzy rodzaje [7]:

- Interesujący nas sygnał zależny od podstawowej wielkości mierzonej.
- Sygnał interferencji (zakłócenia zewnętrzne), na który czujnik jest dodatkowo czuły, sygnał interferujący może być w najprostszym przypadku traktowany jako liniowa superpozycja z sygnałem nas interesującym.
- Sygnał modyfikujący, który może być sygnałem zmieniającym czułość czujnika na wielkość mierzoną oraz sygnałem zakłóceń. Zmienia się zatem zależność między wejściem i wyjściem. W wyniku oddziaływania wielkości modyfikującej pojawia się najczęściej błąd systematyczny.

Czułość definiuje się, jako zależność zmiany wielkości wyjściowej do zmiany wielkości wejściowej czujnika. Czułość jest określana jako nachylenie krzywej kalibracyjnej. Jej wartość jest stała jedynie dla liniowej charakterystyki przetwarzania. Dla elementów nieliniowych czułość zmienia się wraz z wartością wielkości wejściowej przetwornika.

Czułość skrośna oznacza, że przetwornik jest czuły nie tylko na wielkość podstawową, ale również na wielkości dodatkowe (ang. *secondary measurands*). Czułości skrośne w czujnikach aktywnych mogą oddziaływać pośrednio, przez co powodują zmianę czułości na podstawową wielkość mierzoną. Wielkości wpływające mogą powodować znaczący dryft. Dryft można określić jako systematyczną niedokładność o niezerowej wartości [8]. Najczęściej dryft rozumiany jest jako zmiana właściwości czujnika, która pojawia się wraz z upływającym czasem pod wpływem jego zmian strukturalnych. Przy zmiennych warunkach pracy czujnika dryft jest powodowany również zmianą parametrów czujników pod wpływem warunków środowiskowych, w których pracuje czujnik. Warunki środowiskowe są określane przez parametry wpływające na niego. Efekty obydwu przyczyn powstawania dryftu (zmiany parametrów czujnika pod wpływem starzenia się oraz warunków środowiskowych) są często trudne do rozróżnienia.

Offset przejawia się w postaci niezerowego sygnału wyjściowego dla zerowego sygnału wejściowego. Jest to stała wartość wynikająca z niedoskonałości czujnika, powstałych głównie podczas jego wytwarzania. Wielkość ta może być łatwo skorygowana jednorazowo podczas kalibracji wstępnej.

Proces ustalania zależności między fizyczną wielkością mierzoną a sygnałem wyjściowym z czujnika nazywany jest wzorcowaniem (w spektroskopii częściej używa się terminu kalibracja). Odbywa się on przez dostarczanie do czujnika znanych wartości wejściowych z jednoczesnym rejestrowaniem sygnału wyjściowego. Sam pomiar kalibracyjny (wzorcowanie) odbywa się przeważnie w ustalonych warunkach z kontrolą ewentualnych wielkości wpływających. Charakterystyka statyczna określona na podstawie pomiarów wzorcujących i przedstawiona w postaci graficznej jest nazywana krzywą wzorcowania lub kalibracji.

2.1. POMIARY I ICH KOREKCJA

Pomiar jest to określenie cech mierzalnych obiektów oraz zjawisk fizycznych i przedstawienie ich w postaci liczbowej. Wyznaczanie estymaty (odtworzenie) wiel-

kości mierzonej jest oparte na stworzonym modelu pomiaru. Obliczenia są wykonywane na surowych wynikach pomiaru. Podczas odtwarzania mogą jednocześnie być wykonywane zadania, takie jak korekcja i filtracja.

W pomiarach wieloparametrowych odtwarzanie polega na określeniu składowych wektora wielkości mierzonej. Pomiar wieloparametrowy występuje wtedy, gdy część toru pomiarowego jest wspólna dla różnych wielkości mierzonych [9]. W większości przypadków pomiar wieloparametrowego (wielowymiarowy system pomiarowy) wielkości mierzone oddziałują wzajemnie na siebie.

Na wynik pomiaru oddziałują wielkości wpływające oraz zakłócenia. Ich korekcję wykonać można, jeśli określona jest funkcja ich oddziaływania na sygnał wyjściowy czujnika oraz jeśli dysponujemy pomiarem ich wartości. Błąd pomiaru jest to różnica między wynikiem pomiaru a rzeczywistą (prawdziwą) wartością mierzand. Błąd składa się z części systematycznej oraz losowej.

Rodzaje błędów:

- Błędy addytywne – powstają w wyniku oddziaływania wielkości wpływających na wyjściową wielkość niezależnie od podstawowej wielkości wejściowej.
- Błędy multiplikatywne – oddziaływanie wielkości wpływających na wielkość wyjściową w liniowej zależności od wielkości wejściowej.
- Interakcje – powstają w wyniku oddziaływania wzajemnego wielkości wpływających.
- Nieliniowość – odstępstwo charakterystyki statycznej od charakterystyki liniowej.

Błędy pomiarowe mogą być kompensowane lub podlegać korekcji. Kompensacja jest wykonywana w postaci sprzętowej. Wykorzystuje się w niej właściwości elementu, zwanego kompensatorem, znajdującego się w torze pomiarowym. W przypadku kompensacji nie jest wymagany pomiar wielkości wpływających. W najprostszej postaci kompensator oraz czujnik powinny mieć identyczne wrażliwości na wielkości wpływające.

Podstawowymi strukturami kompensacji są [9]:

- Równoległa (w tym jej odmiana różnicowa) – umożliwia kompensację błędów addytywnych.
- Szeregową – umożliwia kompensację zarówno błędów addytywnych, jak i multiplikatywnych.
- Ilorazowa – umożliwia kompensację błędów multiplikatywnych.

Kompensacja układowa może dotyczyć również linearyzacji charakterystyki przetwarzania. Obecnie jest ona jednak stosowana stosunkowo rzadko ze względu na ograniczoną możliwość wytworzenia właściwej charakterystyki elementu kompensacyjnego. Nawet właściwie dopasowany nieliniowy układ elektroniczny będzie wprowadzał błędy związane chociażby ze zmianami temperatury otoczenia. W tego typu sprzętowych układach kondycjonowania trudno jest wprowadzić możliwość zmiany parametrów i dopasowania do ewentualnych zmian czujnika. Dlatego też częściej

stosowane są metody linearyzacji programowej, używane często razem z korekcją błędów, wykonywane w procesie odtwarzania wielkości mierzonej. Podstawowym problemem podczas realizacji korekcji błędów jest konieczność pomiaru wielkości wpływających. Stosunkowo prosto można wykonać korekcję błędów addytywnych niezależnych od wielkości wejściowej. Trudniejsza jest korekcja błędów multiplikatywnych zależnych od wielkości mierzonej.

Idealne systemy pomiarowe powinny być czułe jedynie na podstawową wielkość mierzoną. Minimalizacja wrażliwości na poszczególne wielkości wpływające jest jednym z zadań podczas projektowania systemu pomiarowego. Dla pomiarów wykonywanych w warunkach laboratoryjnych stosuje się często utrzymywanie stałych warunków środowiskowych, w tym wielkości wpływających, takich jak temperatura czy ciśnienie. Podczas konstruowania przyrządów pomiarowych zastosować można stabilizację wielkości wpływających lub też niekiedy separację od wielkości wpływających.

Każdy pomiar podlega zniekształceniom. Błędy systematyczne mogą wynikać ze zmian zachodzących bądź w instrumencie pomiarowym, bądź w mierzonym procesie (obiekcie). Błędy systematyczne w metodach analitycznych można pogrupować na błędy, które mogą zostać skorygowane oraz takie, które skorygowane być nie mogą ze względu na brak możliwości ich określenia [10]. Do błędów niemożliwych do skorygowania zaliczyć można błędy powstające w wyniku kalibracji oraz interferencji nieznanymi składnikami. Systematyczne błędy, które mogą zostać skorygowane, podzielić można na systematyczne stałe oraz systematyczne proporcjonalne do wielkości mierzonej [10].

Dryft wynika z niedokładności produkcyjnych czujników, niejednorodności użytych materiałów, jak też zmienności parametrów materiałowych w związku z warunkami otoczenia oraz starzeniem. Dryft powstaje jako efekt zmian zachodzących w parametrach, które powinny pozostać stałe, takie jak przewodność elektryczna, potencjał, właściwości optyczne. W efekcie odpowiedź czujnika czy też systemu pomiarowego zmienia się pomimo oddziaływania wielkości mierzonej o tej samej wartości. Efekt ten obserwuje się szczególnie, gdy określony czujnik pracuje długi czas. Niektóre rodzaje czujników podlegają tzw. zatrutowaniu, czyli zmianie swoich parametrów pod wpływem ekspozycji na określone czynniki. Narażenie takie jest większe dla warunków środowiskowych lub przemysłowych, gdzie nie jest możliwe kontrolowanie warunków ekspozycji czujnika. Zmiany mogą okazać się nieodwracalne lub nie w pełni odwracalne. Przeciwdziałanie dryftowi osiągnąć można przez inną budowę czujnika lub zastosowanie odpowiednio lepszych materiałów. Jeśli nie jest to możliwe, należy stosować metody kompensacji. Dla oddziaływań liniowych dryft podzielić można na addytywny, multiplikatywny oraz mieszany. Dla oddziaływań nieliniowych podział ten nie jest stosowany.

Dryft powoduje zwiększenie kosztów obsługi urządzeń, ponieważ muszą być one ciągle na nowo kalibrowane, co generuje koszty i zajmuje dużo czasu. Jeśli dryftu nie

da się wyeliminować lub skorygować, ważne jest, aby określić jak często potrzebna jest ponowna kalibracja. Jeśli zmiany parametrów czujników w wyniku starzenia nie są jednostajne, to pożądana jest metoda wskazująca na konieczność kalibracji [11]. Jeśli jednak czujnik pracował zbyt długo narażony na szkodliwe środowisko, konieczna może okazać się jego wymiana.

W warunkach laboratoryjnych wiele czynników zakłócających podlega kontroli. Nie jest to możliwe dla systemów pomiarowych pracujących jako analizatory *in situ*. Podstawowymi problemami mogą być tutaj duże zmiany temperatury, duże zmiany wilgotności, narażenie na korozję, wibracje. Ekstrakcyjne systemy ciągłego monitoringu gazów składają się z dwóch podstawowych części: podsystemu pobierania próbek oraz podsystemy ich analizy. Każdy z tych podsystemów może wprowadzać określony błąd. W samym systemie analizy błędy najczęściej są wynikiem interferencji od innych gazów. Jeśli przed analizą stosowane są systemy usuwania gazów, powodujących interferencje, to wielkość błędu zależy od ich efektywności. Kolejną przyczyną wprowadzania błędów jest konstrukcja podsystemu analizy, a szczególnie jej podatność na różnorodne czynniki wpływające. W przeciwieństwie do pomiarów *in situ* w systemach ekstrakcyjnych sam podsystem analizy może jednak mieć stabilizowane warunki pracy.

Podatność na interferencje oraz wielkości wpływające jest zależna od zastosowanej metody pomiaru. Pewne niepożądane zależności mogą być określone lub wyeliminowane podczas procesu kalibracji. Istotna jest tutaj metoda kalibracji oraz jakość gazów kalibracyjnych. Jeśli nie jest możliwe wyeliminowanie interferującego czynnika przed analizą lub też zbudowanie czujnika, lub zastosowanie techniki czulej na jeden określony składnik, niezbędna okazać się może potrzeba pomiaru składnika interferującego oraz korekcja wskazań.

Jeśli występuje dryft, to budowanie modelu kalibracji powinno się wykonać dla danych z pomiarów wykonanych w określonym czasie i warunkach. Dane pochodzące z pomiarów wykonywanych z występującym dryfem o różnych wartościach nie spowodują zbudowania modelu o dobrej jakości. Z całą pewnością model taki nie mógłby służyć do redukcji dryftu [12]. Jeśli mamy do czynienia z dryfem spowodowanym starzeniem się czujników, to najprostszą metodą eliminacji jego wpływu jest okresowo wykonywany pomiar referencyjny. Pomiar referencyjny wykonywane mogą być dwoma metodami. W pierwszej wykonuje się pomiar technikami alternatywnymi, w drugiej stosuje się wtrącenie składników o znanej zawartości i obserwuje się różnice w widmie przed i po wtrąceniu.

Wyjście przetwornika pomiarowego, uwzględniające dodatkowe czynniki, zapisać można następująco [8]:

$$y = sx + SVB + SSM + y_0 \quad (1)$$

gdzie: x – podstawowa wielkość mierzona, s – czułość na podstawową wielkość mierzoną, SVB – zmienność czułości w wyniku wielkości wyjściowej czujnika, SSM – czułość na dodatkowe wielkości wpływające, y_0 – dryft o stałej wartości.

Podstawową metodą eliminacji dryftu jest okresowe usuwanie wielkości mierzonej z czujnika i pomiar jego wartości wyjściowej, gdy jej brakuje (ang. *chopping*). Wartość sygnału wyjściowego przy braku sygnału wejściowego wynosi

$$y_1 = SVB_1 + SSM(U_1^1, \dots, U_n^1) \quad (2)$$

w obecności sygnału wejściowego odpowiednio

$$y_2 = sx_2 + SVB_2 + SSM(U_1^2, \dots, U_n^2) \quad (3)$$

Porównując wyrażenia na wartości y_2 i y_1 dostrzec można zmianę sygnału zależną od wielkości wpływających. Kluczowanie sygnału (ang. *chopping*) nie jest więc w stanie całkowicie wyeliminować dryftu, może go jedynie znacząco zredukować.

Kolejną metodą minimalizacji dryftu jest metoda zmiennej czułości (ang. *Sensitivity Variation Method* – SVM), w której zmieniana jest czułość przetwornika na wielkość mierzoną. W tym przypadku wielkości y_1 i y_2 wyrazić można zależnościami

$$y_1 = sx + SVB_1 + SSM(U_1^1, \dots, U_n^1) \quad (4)$$

$$y_2 = (s + \Delta s)x + SVB_2 + SSM(U_1^2, \dots, U_n^2) \quad (5)$$

Zmiany od sygnałów wpływających są teraz mniejsze. Większe mogą być jednak różnice w wartościach obciążenia estymacji ze względu na zmianę czułości przetwornika. Problemem może być również nieliniowość czujnika spowodowana zmiennością jego czułości wraz ze zmianą wartości wielkości mierzonej.

Dryft w czujnikach aktywnych może powstawać w wyniku zmiany warunków zasilania czujnika, które ustalają jego punkt pracy. Przykładowo odpowiedź czujników Halla zależy od napięcia zasilania. Przesunięcie z tym związane może być aproksymowane z wykorzystaniem wielomianów drugiego stopnia [13].

Jedną z możliwości korekcji dryftu jest wykorzystanie logiki rozmytej [14]. Dla określonych przedziałów wielkości wpływającej stosowane są modele liniowe, pomiędzy którymi przejście jest wykonywane za pomocą algorytmów logiki rozmytej [15]. Reguła Takagi–Sugeno przyporządkowuje wagi do każdego modelu lokalnego w każdym punkcie wielkości wpływającej. Wyjście modelu y jest więc ważoną sumą określonej liczby modeli lokalnych:

$$y = \frac{\sum_{i=1}^r \lambda_i y_i}{\sum_{i=1}^r \lambda_i} \quad (6)$$

gdzie: y_i – wyjście poszczególnych modeli, λ_i – stopień przynależności danego modelu lokalnego w modelu globalnym dla danej wartości wielkości wejściowej modelu.

Macierze czujników typu elektronicznego nosa mają wiele ograniczeń w stosunku do pracy w trudnych warunkach. Poszczególne macierze różnią się profilem czułości (czułością poszczególnych czujników) i odpowiedzią na wymuszenie, nawet jeśli pochodzą z tej samej serii produkcyjnej. Problem w zastosowaniu macierzy czujników w systemach przemysłowych wynika również z ich słabej selektywności, którą można poprawić przez zastosowanie algorytmów rozpoznawania kształtów. Kolejną wadą jest stabilność parametrów czujników w czasie [16]. Parametry czujników w macierzy, nawet jeśli pracują w tych samych warunkach, zmieniają się w różny sposób [17]. Przebieg i historia oddziaływań również mogą znacząco wpływać na działanie czujników. Ponieważ zmiany są stosunkowo szybkie, to wymagają częstej kalibracji, która niestety jest bardzo kosztowna. Ponownej kalibracji wymaga również wymiana czujnika czy całej macierzy. Główne powody występowania zmian czasowych przyspieszających, i tak dosyć szybkie starzenie się, to [12, 18]:

- wilgotność i jej zmiany,
- degradacja pod wpływem temperatury,
- degradacja pod wpływem oddziaływań mechanicznych,
- adsorpcja różnorodnych czynników,
- zmiany wielkości przepływu mierzonych gazów,
- zmiany ciśnienia.

Aby zwiększyć czasową stabilizację parametrów czujników, można stosować zmiany technologiczne oraz specjalne techniki projektowania czujników. Jeśli jednak mamy już określony czujnik lub matryce czujników, można stosować jedynie odpowiednie techniki pracy czujników lub różnorodne metody przetwarzania sygnałów. Wśród nich wymienić można [18]:

- korekcje jedno- lub wielowymiarowe,
- metody korekcji liniowe lub nieliniowe,
- metody adaptacyjne,
- metody polegające na pomiarze próbek referencyjnych,
- usuwanie zakłócających komponentów sygnałowych.

W macierzy czujników źródła dryftu dzieli się na dwa podstawowe [19]:

- Dryft rzeczywisty lub też pierwszego rzędu (ang. *real-drift*, *first-order drift*) spowodowany zmianami fizycznych właściwości czujnika w wyniku starzenia oraz oddziaływania środowiskowego.
- Dryft pomiarowy lub też drugiego rzędu powodowany zmianami i oddziaływaniami zewnętrznymi, takimi jak zmienne warunki środowiskowe.

Obydwa rodzaje dryftu oddziałują jednocześnie, a ich rozróżnienie jest praktycznie niemożliwe. W sytuacjach praktycznych stosuje się metody przeciwdziałania dryftowi o znanym źródle. Dla macierzy czujników dryft może być kompensowany na poziomie pojedynczego czujnika lub też dla całego ich zbioru. Najczęściej używaną metodą wielowymiarowej korekcji dryftu jest transformacja dopasowująca dane po-

miarowe do aktualnego stanu czujników. W większości przypadków są to transformacje liniowe nie w pełni eliminujące problem. Nieco lepsze rezultaty osiągane są za pomocą metod nieliniowych.

Sam podział metod korekcji dryftu w macierzach czujników przedstawić można następująco [20]:

- Periodyczne powtarzanie kalibracji, dzięki czemu otrzymać można informację o kierunku i wielkości dryftu, który następnie można usunąć (m. in. Multiplikatywna korekcja dryftu MDC oraz wielowymiarowa korekcja wykorzystująca np. PCA).
- Metody attuning, które określają dryft na podstawie danych kalibracyjnych, a następnie usuwają go (m.in. korekcja komponentu CC oraz ortogonalna korekcja sygnałów OSC).
- Filtrowanie i korekcja sygnału (wykonywanie pochodnych sygnału w celu usunięcia linii podstawowej (*baseline*), filtry FIR oraz wykorzystujące transformatę falkową).
- Modele adaptacyjne, które próbują określić zmiany zachodzące w sygnale, najczęściej używa się tutaj sieci neuronowych. Metody takie są przydatne raczej podczas klasyfikacji oraz wolno zmieniającego się dryftu [21].

Dane pomiarowe z macierzy czujników typu elektroniczny nos różnią się od danych ze spektroskopii, ponieważ ich wymiar jest mniejszy oraz nie występuje ciągłość danych między czujnikami, tak jak w widmie w spektroskopii. Poszczególne czujniki mają niezależne od siebie właściwości. W przypadku macierzy czujników korekcja dotyczy może każdego czujnika indywidualnie lub też wektora danych z macierzy jako całości. Najprostszą metodą jest jednowymiarowa korekcja każdego czujnika w macierzy osobno. Zbudować należy model korekcji dla każdego czujnika. Wielowymiarowe metody korekcji mogą korygować nie tylko zmiany poszczególnych czujników. Istnieją nawet metody transformacji macierzowej podczas przenoszenia pomiarowych baz danych między systemami elektronicznego nosa o podobnych właściwościach [22].

Kompensacja metodami jednowymiarowej bezpośredniej standaryzacji UDS (ang. *Univariate Direct Standardization*) oraz PLS2 może być wykonana w przypadku, gdy pomiarowe dane kalibracyjne zawierają dwa lub większą liczbę typów próbek referencyjnych [23]; przy czym próbki takie muszą być mierzone w regularnych odstępach czasu tak, by widoczny stał się kierunek oraz szybkość zmian właściwości czujnika.

Metoda multiplikatywnej korekcji dryftu (ang. *Multiplicative Drift Correction* – MDC) polega na kompensowaniu osobno każdego czujnika w macierzy. Kompensacja dryftu dla każdego czujnika opiera się na następującym modelu regresji [23]:

$$f_{nk} = (ax_n + b)_k \quad (7)$$

W przypadku nieliniowej zależności dryftu stosowane są wielomiany wyższych rzędów, choć często wystarczy funkcja kwadratowa [24, 25]. W elektronicznym nosie stosowane są wykładnicze modele korekcji temperatury [26]. Przykładowo dla kom-

pensacji temperatury oraz wilgotności dla macierzy czujników zastosowano następujące wzory [27]:

$$R_T = R_0 \exp(b_T T) \quad (8)$$

$$R_H = R_0(1 + a \exp(-b_h h)) \quad (9)$$

gdzie: a , b_T , b_h – stałe otrzymane empirycznie, T – temperatura, h – wilgotność względna, R_0 – rezystancja czujnika w temperaturze i wilgotności względnej odniesienia.

Metoda korekcji komponentu (ang. *Component Correction* – CC) oparta jest na PCA [28]. Charakteryzuje się ona tym, że bada całą macierz czujników. Monitorowane są zmiany w kilku pierwszych składowych głównych. Na wstępie wykonywana jest procedura PCA dla danych referencyjnych, a następnie dla całości mierzonych danych. Korekcja dryftu następuje przez odjęcie pierwszej składowej głównej, reprezentującej dryft między poszczególnymi pomiarami. Zachodzą również modyfikacje metody CC [29].

Kompensacja wielu różnych czujników o jednowymiarowym wyjściu jest często wykonywana za pomocą modelu sieci neuronowych. Sieci neuronowe wykazują dobrą efektywność zarówno w kompensacji zmian temperaturowych czujników zawartości gazów [30, 31], jak też czujników innego rodzaju, np. ciśnienia [32, 33], naprężenia i siły według tensometrów [34–36], czujników pojemnościowych ciśnienia [37, 38] oraz dla zmian temperatury i wilgotności dla czujników gazów [39, 40]; kompensacja temperatury dla pojemnościowych czujników wilgotności [41]. W znaczącej większości opisanych przypadków zastosowanie miały sieci typu perceptron wielowarstwowy. Samoorganizujące się mapy Kohonena umożliwiają korekcje dryftu macierzy czujników do celów klasyfikacji [42, 43].

Modele odwrotne oparte na sieciach neuronowych najczęściej wykorzystują fuzję danych z czujnika wrażliwego na badaną wielkość oraz kilka (najczęściej jedną lub dwie) czynników zakłócających. Jest to więc jedno z rozwiązań fuzji danych.

Dryft o stałej wartości może być również kompensowany z wykorzystaniem filtrów Kalmana [44]. Dla niektórych rodzajów czujników do kompensacji wielkości wpływających opracowane zostały specjalne modele. Dla czujników tlenu, opartych na światłowodach oraz intensywności fluorescencji, stosowany jest specyficzny model korekcji temperatury, którego parametry są wyznaczone eksperymentalnie [45].

2.2. CZUJNIKI INTELIGENTNE, LINEARYZACJA

Projekt systemu pomiarowego powinien ograniczać wpływ zarówno sygnałów interferujących, jak i modyfikujących. Same elementy przetworników wielkości (czujnik fizyczny) niekoniecznie mają optymalne właściwości metrologiczne, takie jak czułość, selektywność czy dokładność. Do budowy systemu pomiarowego najlepiej

jest użyć czujników o znikomym lub zupełnym braku wpływu czynników dodatkowych. Jeśli nie jest to możliwe, stosowane są techniki utrzymywania odpowiednich warunków pomiarowych. W pozostałych przypadkach, w tym w warunkach przemysłowych, stosowane muszą być odpowiednie techniki kompensacji lub korekcji. Podstawową metodą rozwiązania tego problemu we współczesnych systemach pomiarowych jest wykorzystanie metod przetwarzania sygnałów, które służą do otrzymania dodatkowych właściwości systemu lub czujnika. Dzięki rozwojowi technik przetwarzania informacji obserwuje się zmianę podejścia do wielu rodzajów systemów pomiarowych w kierunku stosowania większej liczby tanich czujników, zamiast niewielkiej liczby drogich urządzeń.

W torze przetwarzania sygnału w systemie pomiarowym po czujniku występuje element kondycjonujący sygnał. Umożliwia on dopasowanie sygnału do celów zapisu, transmisji lub przedstawienia operatorowi. Złożony układ kondycjonowania sygnału może wykonywać kilka funkcji, takich jak: linearyzacja, korekcja offsetu, korekcja czynników wpływających, korekcja histerezy, autokalibracja, samotestowanie, detekcja awarii. Czujnik wraz z odpowiednim dedykowanym systemem przetwarzania sygnału jest określany jako czujnik inteligentny (ang. *intelligent, smart sensor*). Układ taki może być zintegrowany w jednym układzie scalonym. W znaczeniu ogólnym składnik inteligentny (ang. *smart, intelligence*) ma za zadanie polepszyć właściwości czujnika lub sygnału wyjściowego z czujnika. Polepszenie właściwości może dotyczyć np. kompensacji wpływu wielkości niepożądanych lub kompensacji charakterystyki statycznej czujnika. Polepszenie jakości sygnału może dotyczyć filtracji oraz eliminacji szumów.

Jedną z cech czujników inteligentnych jest również możliwość przeprowadzania autokalibracji, jeśli tylko zajdzie taka potrzeba. Procedura może być wykonana w przypadku, gdy dokładność czujnika zostanie zakwestionowana w wyniku cyklicznej procedury walidacyjnej.

Głównym zastosowaniem czujników inteligentnych są pomiary środowiskowe oraz procesowe, pomiary niezbędne do systemów kontroli automatycznej oraz w urządzeniach mechatronicznych [46–49]. Informacja z czujników inteligentnych jest podawana do systemów automatyki w formie niewymagającej jej dodatkowego przetwarzania, co uwalnia system automatyki od wykonywania dodatkowych operacji. Inteligentny system pomiarowy może wykorzystywać kombinację sygnałów wyjściowych z wielu czujników (integruje sygnały) lub rodzajów czujników do pozyskania wiedzy o obiekcie lub procesie

Zastosowanie inteligencji w systemach pomiarowych jest szczególnie istotne, gdy pracują one w warunkach przemysłowych lub środowiskowych. Główne powody stosowania rozwiązań inteligentnych w warunkach pozalaboratoryjnych, to [50]:

- poprawa właściwości metrologicznych (np. zwiększenie dokładności pomiaru, linearyzacja, kompensacja),
- zwiększenie możliwości operacyjnych oraz obsługi,
- polepszenie możliwości komunikacyjnych,
- diagnostyka systemu pomiarowego,

- pośrednia obserwacja za pomocą czujników wirtualnych wielkości niemożliwych do obserwacji bezpośredniej.

Samo słowo inteligentny użyte w stosunku do czujnika czy systemu pomiarowego nie zawsze odpowiada jego właściwościom, a często jest tylko zabiegiem marketingowym. Niektórzy autorzy przyjmują za czujnik inteligentny, zintegrowany z obwodem przetwarzania sygnału w jednym układzie scalonym. Inni wymagają, by czujnik miał zintegrowane wyjście cyfrowe. W niektórych przypadkach samo zastosowanie może mieć charakter inteligencji, natomiast sam czujnik takiego charakteru nie ma. Stąd też pojawia się potrzeba podziału stopnia inteligencji czujników rozumiana jako posiadanie przez nie nowych możliwości. Wiązą się one również z kolejnymi etapami ich rozwoju. Czujniki o najmniejszym stopniu inteligencji, określane z języka angielskiego jako „smart”, są ograniczone do:

- przetwarzania sygnałów,
- możliwości komunikacji z innymi systemami,
- integracji w jednym układzie scalonym.

Kolejny poziom inteligencji stanowią czujniki inteligentne mające możliwość wbudowanej kompensacji nieliniowości, czułości skrośnych oraz dryftu [51]. Najwyższym poziomem rozwoju charakteryzują się czujniki myślące (ang. *cogent*), które potrafią wyeliminować niepotrzebne dane i przedstawić sygnały w postaci informacji potrzebnej do danego zastosowania [52].

Statyczna charakterystyka czujnika jest w większości przypadków nieliniowa. Linearyzacja jest procesem wykonywanym najczęściej wraz z kalibracją. Zarówno linearyzacja, jak też korekcja wybranych czynników wpływających może być wykonywana z wykorzystaniem analogowych układów elektronicznych [53]. Metody analogowe są jednak stosowane w przypadkach szczególnych. W związku z możliwościami obliczeniowymi mikroprocesorów stosuje się linearyzację po stronie cyfrowej. Umożliwia ona korekcję błędów nieliniowości całego toru pomiarowego.

W zależności od przyjętej metody linearyzacji wymagana może być różna liczba punktów kalibracyjnych oraz różne ich rozłożenie w przestrzeni. Jednym z kryteriów wyboru metody linearyzacji powiązanej z kalibracją może być jej koszt związany bezpośrednio z liczbą niezbędnych punktów kalibracyjnych.

Metody kalibracji z linearyzacją [54]:

- tabela pomiarów w określonych punktach (ang. *look-up table*),
- funkcje przedziałowo-liniowe,
- sieci neuronowe,
- logika rozmyta,
- dopasowywanie krzywych (ang. *curve fitting*),
- kalibracja wielomianami różnego rodzaju.

Dopasowanie krzywych może być wykonane wielomianami interpolującymi lub aproksymującymi. Jeśli w danych występują zakłócenia (szum), lepszą metodą jest aproksymacja.

Aproksymacja funkcjami przedziałowo-liniowa jest odpowiednikiem metody tabeli pomiarowej, co najwyżej może zawierać mniejszą liczbę punktów. Dokładność obydwu metod zależy od liczby punktów oraz rodzaju nieliniowości aproksymowanej funkcji.

Aproksymacja metodami logiki rozmytej polega na wykorzystaniu zbioru reguł wykorzystujących określoną wiedzę. Logika rozmyta dzieli aproksymowaną funkcję na określone przedziały, w których oddziałują dane reguły. Dla funkcji nieliniowych podział na poszczególne zbiory jest wykonany nierównomiernie. W miejscach, gdzie charakterystyka jest bardziej nieliniowa reguła jest odpowiednio więcej.

Wykorzystanie sztucznych sieci neuronowych oraz logiki rozmytej w tanich mikrokontrolerach nie przynosi dobrych rezultatów ze względu na opóźnienia w obliczaniu wyniku oraz rozmiaru programu [55], chociaż najmniejsze błędy dają metody sztucznych sieci neuronowych oraz mają najlepsze właściwości modelowania silnie nieliniowych charakterystyk [56–58].

Do kompensacji nieliniowości wraz z występującą dużą histerezą skutecznie użyte zostały metody wektorów wspierających (ang. *Support Vector Machines* – SVM). Jednym z przykładów jest czujnik względnej wilgotności [59].

Najpopularniejszą metodą linearyzacji są wielomiany. Przykładem może być ich zastosowanie do precyzyjnych pomiarów temperatury z wykorzystaniem termopar. Pomimo że termopary nie są elementami o dużej nieliniowości, to w stosunku do modelu liniowego można osiągnąć w zależności od zakresu temperatury kilkunastokrotne (17-krotne dla zakresu 200 °C) oraz 40-krotne dla zakresu 436–1200 °C, co w jednostkach bezwzględnych oznacza zmniejszenie maksymalnego błędu z –18 °C do –0,4 °C [60]. Wynik ten jest podobny do modelowania za pomocą perceptronu wielowarstwowego, gdzie uzyskano maksymalny błąd względny mniejszy od 0,5% [61].

Do kompensacji czynników środowiskowych w inteligentnych czujnikach zaproponowano również metody łączenia wielomianów z sieciami neuronowymi. Metoda (ang. *Legendre Functional-Link Neural Network* – L-FLANN) wykorzystuje architekturę jednowarstwowego perceptronu, na którego wejście podawane są sygnały przetworzone przez wielomiany Legendre’a [38].

Podstawowym problemem w klasycznych metodach linearyzacji jest konieczność wykonywania procedury kalibracji za każdym razem, gdy zmieni się chociażby jeden punkt kalibracyjny. Rozwiązaniem w takim przypadku jest korekcja progresywna oraz progresywna metoda kalibracji z wykorzystaniem uniwersalnego obwodu kondycjonowania [62].

Wielomiany progresywne umożliwiają uniknięcie tego problemu, a dodatkowo ułatwiają kalibrację adaptacyjną; wymagają mniejszej liczby operacji matematycznych. Wielomiany progresywne (ang. *progressive polynomial*) są tworzone zgodnie z zasadą, dla której każdy punkt kalibracyjny tworzy jeden współczynnik wielomianu. Pierwszy punkt kalibracyjny służy do korekcji offsetu, drugi do korekcji czułości czujnika. Pozostałe punkty pomiarowe (od trzeciego) służą korekcji nieliniowości.

Algorytm do obliczania współczynników jest algorytmem rekurencyjnym, w którym każdy kolejny współczynnik powstaje z wykorzystaniem poprzedniego oraz dodatkowego punktu pomiarowego. Z każdym kolejnym pomiarem kalibracyjnym tworzona jest skorygowana funkcja charakterystyki statycznej. Wielomiany progresywne wykazują dobre właściwości dla nieliniowości osiagającej 30%, w wersji ulepszonej dla nieliniowości osiagającej 45% [63].

Algorytm tworzenia wielomianu, dla danych znormalizowanych, można przedstawić następująco [51, 64–66]:

- Dla N punktów pomiarowych określone zostały następujące znormalizowane wektory:
 - sygnał wejściowy czujnika $x(n) = [x_1, x_2, x_3, \dots, x_N]$,
 - sygnał wyjściowy czujnika $y(n) = [y_1, y_2, y_3, \dots, y_N]$,
 - zlinearyzowany (idealny) sygnał wejściowy po kalibracji czujnika $t(n) = [t_1, t_2, t_3, \dots, t_N]$.

Normalizacja została wykonana następująco: dla wielkości wejściowej i wyjściowej czujnika

$$x_n = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (10)$$

$$y_n = \frac{y - y_{\min}}{y_{\max} - y_{\min}} \quad (11)$$

- Obliczamy pierwszy współczynnik odpowiedzialny za korekcję offsetu

$$k_1 = t_1 - y_1 \quad (12)$$

co odpowiada następującej funkcji korekcji

$$f_1 = y + k_1 \quad (13)$$

- Obliczamy drugi współczynnik odpowiedzialny za czułość czujnika (po założeniu jego liniowości)

$$k_2 = \frac{(t_2 - f_1(x_2))}{f_1(x_2) - t_1} \quad (14)$$

Otrzymujemy następującą funkcję korekcji

$$f_2 = f_1 + k_2(f_1 - t_1). \quad (15)$$

- Trzeci współczynnik obliczony będzie następująco:

$$k_3 = \frac{(t_3 - f_2(x_3))}{(f_1(x_3) - t_1)(f_2(x_3) - t_2)} \quad (16)$$

W rezultacie funkcja korekcji będzie następująca:

$$f_3 = f_2 + k_3(f_1 - t_1)(f_2 - t_2) \quad (17)$$

- Każdy kolejny współczynnik obliczany jest następująco:

$$k_n = \frac{(t_n - f_{n-1}(x_n))}{\prod_{j=1}^{n-1} (f_j(x_n) - t_j)} \quad (18)$$

Funkcja korekcji przyjmuje postać

$$f_n = f_{n-1} + k_n \prod_{j=1}^{n-1} (f_j - t_j) \quad (19)$$

Podane dwa wzory są prawdziwe dla $n = 2, \dots, N$.

Sprawdzenie dokładności zastosowanego algorytmu może być wykonane za pomocą średniego błędu kwadratowego (ang. *Mean Square Error* – MSE)

$$MSE(k) = \frac{1}{NCP} \left(\sum_n^{NCP} \right) \quad (20)$$

Podstawową metodą wyboru punktów kalibracji jest ich równomierne rozłożenie w całym zakresie zastosowania. Niestety metoda ta rzadko daje optymalne rezultaty.

Metoda PDF-Based [66] opiera się na prostym spostrzeżeniu, że czujnik, szczególnie dedykowany do pomiarów procesowych, nie wykonuje równomiernej liczby pomiarów w całym zakresie pomiarowym. Klasyczne metody planowania eksperymentu starają się przedstawić całą wariację zmiennych wejściowych. Jeśli jednak model wykorzystywany będzie głównie w określonym zakresie, punkty kalibracji powinny to odzwierciedlać [67]. Dobrze dobrany czujnik powinien pracować w okolicach 2/3 zakresu pomiarowego, przy czym pamiętać należy, aby rozrzut wartości wielkości mierzonej nie przekraczał wartości maksymalnej. Funkcję gęstości prawdopodobieństwa (ang. *Probability Density Function* – PDF) oblicza się z danych historycznych, jako stosunek liczby pomiarów w danym przedziale pomiarowym do wszystkich wykonanych pomiarów w całym zakresie pomiarowym czujnika [66]

$$pdf(k) = \frac{f(k)}{\sum_{k=0}^{N_c-1} f(k)} \quad (21)$$

gdzie: $f(k)$ – liczba pomiarów w danym przedziale, N_C – liczba przedziałów, na które został podzielony zakres pomiarowy.

Funkcja gęstości prawdopodobieństwa jest histogramem o rozdzielczości będącej jednocześnie długością każdego z przedziałów

$$\Delta x = \frac{x_{\max} - x_{\min}}{N_C} \quad (22)$$

Liczba punktów kalibracyjnych (ang. *Number of Calibration Point* – NCP) w każdym przedziale wyznacza się następująco [66]

$$NCP(k) = pdf(k)NCP \quad (23)$$

Inteligentna metoda kalibracji [68] polega na wyborze najlepszych punktów kalibracji z wszystkich dostępnych punktów danych. W metodzie tej wymagane jest użycie minimalnej oraz maksymalnej wartości zakresu pomiarowego, jako dwóch pierwszych punktów kalibracji. Działanie inteligentnej metody kalibracji można opisać następująco [68]:

- Wykonanie określonej ilości pomiarów kalibracyjnych.
- Wykonanie procedury tworzenia wielomianu progresywnego dla dwóch pierwszych punktów (minimum i maksimum zakresu pomiarowego).
- Obliczenie wartości funkcji wielomianowej dla wszystkich punktów pomiarowych stanowiących dane kalibracyjne.
- Znalezienie punktu o największej rozbieżności między pomiarem a obliczeniami.
- Wykonanie kolejnego punktu tworzenia wielomianu progresywnego dla punktu danych z poprzedniego kroku kalibracji inteligentnej.
- Dla każdego z przedziałów wielkości wejściowej, które tworzą użyte dotychczas punkty kalibracyjne następuje znalezienie punktów o największej rozbieżności, a następnie użycie ich do procedury wyznaczania parametrów wielomianu.
- Procedura kończy się wraz ze znalezieniem i użyciem wszystkich punktów kalibracji lub też liczby punktów odpowiadających założonemu wcześniej stopniowi wielomianu.

Wielomiany progresywne mogą być zaimplementowane w postaci analogowych układów elektronicznych o małym poborze mocy jako odpowiednich struktur układowych [69].

3. MODELE ODWROTNE W POMIARACH

W przypadku pomiarów obiektów złożonych pojawia się zagadnienie pomiarów pośrednich i związany z tym problem budowy modeli obiektów. Model samego systemu pomiarowego jest zwykle przedstawiany w postaci matematycznej jako:

- Liniowa lub nieliniowa funkcja wielowejściowa o jednym wyjściu – jednowymiarowa wielkość mierzona z oddziaływaniem wielkości wpływających.
- Równania macierzowe – liniowy model wielowymiarowej wielkości mierzonej.
- Nieliniowa funkcja wielowejściowa i wielowyjściowa – nieliniowy model wielowymiarowej wielkości mierzonej.
- Równania różniczkowe – pomiar dynamiczny.
- Równania stanu – wielowymiarowy pomiar dynamiczny.

Model jest pewnym narzędziem opisującym rzeczywistość. Dla obiektów i procesów, które opisane są skomplikowanymi zależnościami, możliwe jest zbudowanie modeli matematycznych, a następnie badanie tych modeli w sposób symulacyjny. Zmieniając różnorodne parametry takich modeli, przeprowadzić można eksperyment na modelu matematycznym w kontrolowanych warunkach odnośnie do sygnałów wejściowych, wpływających i zakłócających [70]. Budowanie modelu mechanistycznego nie zawsze jest możliwe lub też nie w każdym przypadku jest to metoda praktyczna. W takich przypadkach stosuje się modele empiryczne. Modele takie wykorzystują zmienność oraz korelacje występujące w danych. Są szczególnie przydatne, gdy ściśle modele matematyczne są zbyt skomplikowane lub nie mogą być określone. Modele empiryczne mogą być budowane trzema głównymi metodami:

- statystycznymi (wyróżnić można metody liniowe i nieliniowe),
- uczenia maszynowego,
- heurystyczne (systemy ekspertowe).

Możliwe jest oczywiście użycie modelu zbudowanego z połączenia wymienionych metod. Wykorzystywane są wtedy unikatowe cechy wybranych metod, co prowadzi do budowy modeli o lepszych właściwościach. Wybór odpowiedniego modelu empirycznego wymaga dobrej znajomości poszczególnych technik. Nie istnieją metody jednoznacznie wskazujące, jaki rodzaj modelu powinien zostać wybrany dla konkretnego zastosowania. Modele empiryczne oparte są na danych eksperymentalnych. Aby stworzyć dobrej jakości model, niezbędne jest zebranie reprezentatywnego zbioru danych. Zbiór taki otrzymać można, wykorzystując techniki planowania eksperymentu tak, by w stosunkowo niewielkiej liczbie pomiarów i otrzymanych z nich danych reprezentowana była cała potrzebna informacja. Następnie dane należy poddać wizualizacji, aby móc ocenić ich przydatność. Po wybraniu struktury modelu w postaci opisujących go równań następuje faza kalibracji modelu, polegająca na określeniu jego parametrów na podstawie pomiarów oraz ewentualnie wiedzy *a priori* (zakresy parametrów, ograniczenia). Model jest wtedy pewną aproksymacją badanego systemu. Błędy zbudowanego modelu są wynikiem różnych czynników zakłócających. Czynniki takie podzielić można na znane i nieznanne podczas przeprowadzania eksperymentu.

Podstawową wadą metod statystycznych jest konieczność znajomości modeli regresji przed ich zastosowaniem. Modele statystyczne nie umożliwiają w prosty sposób wykonywać ich korekcji i adaptacji do zmieniających się warunków i nowych danych.

Jednakże liniowym modelom statystycznym jako jedynym można nadać interpretację fizyczną. W przypadku pozostałych metod budowania modeli empirycznych jest to raczej niewykonalne. Nieliniowe modele statystyczne są najczęściej rozszerzeniem modeli liniowych. W niektórych przypadkach można się w nich doszukiwać interpretacji fizycznej. Mają one również podobnie jak sieci neuronowe możliwość generalizacji i uniwersalnej aproksymacji. Sieci neuronowe zazwyczaj wymagają znacznie większej liczby danych treningowych do zbudowania modelu w porównaniu do nieliniowych metod statystycznych.

Budowa modelu według sieci neuronowych składa się z dwóch głównych kroków. W pierwszym należy dokonać wyboru odpowiedniej architektury sieci. Następnie wykonuje się uczenie i testowanie sieci. Sieć uczy się na przykładach ze zbioru danych zawierających wielkości wejściowe oraz wyjściowe modelu. Wiedza jest w nich zgromadzona w połączeniach między neuronami oraz wagami tych połączeń. Sieci neuronowe różnią się metodami połączeń między węzłami, funkcją aktywacji oraz metodami uczenia. Zalety sieci neuronowych, to:

- łatwa adaptacja do nowych danych,
- dobre właściwości generalizacyjne,
- mała liczba współczynników w stosunku do złożoności realizowanej funkcji.

Największym problemem z sieciami neuronowymi jest brak wyraźnych wskazań odnośnie do wyboru architektury sieci. Stosuje się więc głównie intuicję oraz doświadczenie zdobyte przez eksperymentatora.

Wszystkie techniki modelowania empirycznego mogą być przedstawione w postaci ważonej sumy funkcji bazowych. Porównanie metod może być wtedy dokonane przez porównanie następujących elementów modelu [71]:

- Rodzaj wejściowej transformacji danych (projekcja liniowa i nieliniowa, z wagami, selekcja zmiennych wejściowych).
- Rodzaj i kształt funkcji aktywacji, stałe (liniowe, sigmoidalne, gaussowskie) lub adaptujące się do danych (funkcje sklepane, sieci neuronowe z wsteczną propagacją).
- Rodzaj i kryteria optymalizacji modelu dla transformacji wejścia modelu. Informacja może być uzyskiwana jedynie z danych wejściowych (np. PCA) lub też z danych wejściowych i wyjściowych (np. PLS).

Modelowanie mieszane (ang. *composite modelling*) polega na wykorzystaniu połączenia modeli empirycznych oraz modeli fizycznych [72]. Modele mieszane używają informacji z dwóch źródeł, czyli zarówno z pomiarów, jak teorii fizycznych. Modele fizyczne w czystej postaci oznaczają posiadanie pełnej wstępnej informacji o badanym systemie, podczas gdy modele empiryczne oznaczają brak takich informacji. W praktyce modele fizyczne nie są pełne i pewne informacje empiryczne są niezbędne. Modele mieszane aproksymują zależność między odpowiedzią a zmiennymi wejściowymi. Wykorzystanie obydwu otrzymanych funkcji opisujących obiekt może być wykonana w różny sposób, np. addytywnie lub multiplikatywnie.

3.1. CZUJNIKI WIRTUALNE

Rozwój różnorodnych procesów przemysłowych wymaga coraz lepszych metod pomiaru i monitorowania. Czujnik programowy jest to model, jaki powstał w wyniku obserwacji procesu (obiektu) oraz wiedzy o procesie pozyskanej z różnych źródeł. Jeśli zmienna istotna dla określonego procesu nie jest mierzalna w sposób bezpośredni, to wyznaczenie jej wartości może nastąpić na podstawie informacji pomiarowych o innych zmiennych oraz parametrach. Czujniki softwarowe (wirtualne) używane są głównie w przypadku procesów i obiektów o złożonych zależnościach między zmiennymi. Czujniki programowe zastępczo nazywane są często modelami predykcyjnymi (ang. *predictive models*), czujnikami wnioskującymi (ang. *inferential sensors*) lub wirtualnymi analizatorami on-line [73]. Rozwój tego typu czujników powiązać można z rozwojem możliwości przetwarzania danych, a więc zarówno mocy obliczeniowej komputerów, jak też opracowania i wdrożenia różnego rodzaju algorytmów. Wirtualne wielowymiarowe modele czujników wywodzą się głównie z pomiarów procesów chemicznych, w których liczba zmiennych do monitorowania jest zawsze bardzo duża. Czujniki i pomiary wirtualne należą więc do grupy metod fuzji danych, w których na podstawie wielu źródeł informacji otrzymuje się informację niedostępną w inny sposób.

Predykcja w czujnikach wirtualnych dotyczy często pewnych jakościowych właściwości procesu lub wytwarzanych produktów. Oprócz predykcji czujniki takie mogą również służyć do określania stanu procesu (diagnostyka) oraz wykrywania różnego rodzaju awarii (np. czujników sprzętowych) lub anomalii. Wykrycie defektu za pomocą metod wirtualnych może nastąpić wcześniej niż po wykorzystaniu metod alternatywnych. Wykrywanie problemów często polega na porównaniu określonych parametrów uzyskanych dwoma metodami. Awaria jest wykrywana, gdy dane różnią się między sobą. Mogą to być dwa osobne modele wirtualne wykorzystujące różne sygnały wejściowe. W efekcie możliwe jest bardziej efektywne i bezpieczne prowadzenie procesu.

Przyczyny braku dostępności pomiarów niektórych wielkości procesowych mogą być różne. Koszt takiego rozwiązania może być zbyt duży, nie istnieją odpowiednie czujniki sprzętowe lub nie można ich zastosować w danych warunkach. W części przypadków ważne w ujęciu procesu parametrów mogą być mierzone jedynie w trybie off-line w warunkach laboratoryjnych. Ze względu na koszty określenia parametrów nie można wykonywać wystarczająco często lub też pomiar taki jest wykonany ze znaczącym opóźnieniem. Czujnik softwareowy może być jedyną alternatywą, jeśli wymagane jest szybkie dostarczenie informacji bez opóźnień lub nawet w czasie rzeczywistym. Dane o zmiennej wyjściowej modelu mogą pochodzić z pomiarów laboratoryjnych off-line wprowadzających często znaczne opóźnienia. Informacje w czasie rzeczywistym są niezbędne dla systemów sterowania i automatyki.

Czujnik wirtualny może być również czujnikiem zapasowym w stosunku do jego sprzętowego odpowiednika. Jest to wykorzystywane, jeśli ten ostatni jest narażony na

zagrożające mu oddziaływania, przez co jest zawodny. Czujnik wirtualny może też zmniejszać wymagania w stosunku do liczby oraz jakości czujników sprzętowych. Czujnik fizyczny pracujący w trudnych warunkach może wymagać częstej obsługi oraz wielu przerw w pracy. Rozwiązanie wirtualne może być wtedy znacznie tańsze od konwencjonalnego.

Wyróżnić można dwa podstawowe rodzaje tego typu czujników [74]: bazujące na modelu lub na danych. Czujniki softwareowe oparte na modelu (ang. *model-driven*) nazywane są modelami białej skrzynki ze względu na wykorzystanie pełnej wiedzy o procesie podczas jego budowy. Są one w wielu przypadkach dosyć skomplikowane oraz wymagają znacznych mocy obliczeniowych, co stwarza problemy w ich zastosowaniu w trybie on-line. Czujniki wykorzystujące model opierają się najczęściej na fizycznych lub chemicznych zależnościach dla procesu. Odpowiednik angielskojęzyczny to (ang. *First Principle Models* – FPM), czyli modele oparte na zasadach podstawowych. Modele FPM są podstawą działania urządzeń lub przebiegu procesów. Skupiają się one na prawidłowych, często optymalnych i statycznych warunkach. Tworzenie czujnika na podstawie złożonego modelu procesu może okazać się zbyt skomplikowane, dlatego najczęściej stosuje się modele empiryczne z wykorzystaniem zgromadzonych danych pomiarowych. Innym powodem przeciwko stosowaniu metody białej skrzynki jest brak wystarczająco dokładnej wiedzy o procesie przed zaprojektowaniem czujnika. Czujniki oparte na danych reprezentujących stan procesu są bliższe rzeczywistości ze względu na wykorzystanie danych pomiarowych. Modele oparte na danych nazywane są modelami czarnej skrzynki, ponieważ nie uwzględniają one zjawiska, a jedynie dane dotyczące wielkości wejściowych oraz wyjściowych. Możliwe jest również wykorzystanie modeli hybrydowych (mieszanych), które nazywane są zastępczo modelami szarej skrzynki.

Tworzenie czujników programowych opartych na danych polegających głównie na znalezieniu istotnych zależności między wielowymiarowymi danymi a informacjami o procesie. W wyniku takiego wnioskowania powstają empiryczne modele procesu. W tym przypadku model czujnika wirtualnego opiera się na zgromadzonych danych historycznych, które muszą być reprezentatywne i obejmować całą przestrzeń warunków mogących wystąpić w procesie. Odnosi się to również do zmian parametrów materiałów czy zmian konfiguracyjnych. Może to w niektórych przypadkach prowadzić do modelu o skomplikowanej strukturze. Zgromadzone dane służą zarówno do identyfikacji modelu, jak też do jego kalibracji.

Po zbudowaniu czujnik wirtualny może zostać użyty do symulowania zachowania się procesu w różnorodnych warunkach, jak też do analizy sytuacji nietypowych lub wielu przypadków w znacznie krótszym czasie i mniejszym kosztem niż wykonanie eksperymentów rzeczywistych.

Głównym celem stosowania czujników softwareowych jest predykcja wielkości występujących w procesie, a niedostępnych przez innego rodzaju pomiary. Należy więc zbudować następujący model matematyczny:

$$y = f(\mathbf{x}, \boldsymbol{\theta}) \quad (24)$$

gdzie: y – interesująca nas wielkość opisująca proces, \mathbf{x} – wektor wielkości mierzalnych, $\boldsymbol{\theta}$ – wektor parametrów modelu.

Zadaniem jest więc stworzenie właściwego modelu na podstawie zgromadzonych danych, który następnie służy do określania bieżących wartości interesującej nas wielkości

$$\hat{y}(t) = f(\mathbf{x}(t)) \quad (25)$$

Parametry modelu θ muszą zostać określone na podstawie zbioru dostępnych danych. Wyjście takiego modelu dla danych testujących powinno dawać wynik możliwie zgodny z danymi. Jakość budowanych modeli zależy od jakości danych używanych do ich tworzenia. Jest kilka problemów odnośnie do danych empirycznych pochodzących z procesów: brakujące wartości, szum, nadmiarowość, małą informacyjność, znaczne zróżnicowanie odnośnie do dokładności, istotności, częstotliwości. Pierwszym krokiem jest redukcja wymiaru danych wejściowych oraz eliminacja danych nie wnoszących istotnych informacji. W warunkach procesowych częściej bowiem zdarzają się znaczne rozbieżności oraz odchyłki w danych. Brakujące wartości oznaczają, że wartości określonej zmiennej wejściowej nie odpowiadają jej rzeczywistym wartościom. Zmienna może przyjmować wartości skrajne lub np. wartość stałą, mimo zmieniających się warunków. Najczęstszą przyczyną tego zjawiska jest awaria czujnika sprzętowego lub toru transmisji, zapisu i przetwarzania sygnału. W zależności od długości czasu trwania takie brakujące wartości można zastąpić średnią z wartości przed i po wystąpieniu zaniku. Dla dłuższych niedostępności strategia taka raczej nie przynosi oczekiwanych rezultatów. Można więc próbować pominąć daną zmienną podczas budowania modelu lub stworzenia modelu odtwarzającego tę zmienną na podstawie pozostałych. Brakujące wartości stosunkowo łatwo można zauważyć. Trudniej jest wyeliminować wartości znacznie odbiegające od typowych (ang. *outliers*), chyba że określona wielkość fizyczna przyjmuje wartości spoza wartości rzeczywistych, co można wykryć stosunkowo łatwo przez ustalenie ograniczeń, których przekroczenie jest wskazaniem błędu. Mniej oczywistą sytuacją jest wystąpienie wartości mieszczących się w granicach dopuszczalnych, ale nieodzwierciedlających rzeczywistości. Ich wykrycie jest trudne i wymaga inspekcji manualnej. Dodatkowymi problemami pojawiającymi się w danych jest występowanie opóźnień oraz różnych częstotliwości próbkowania sygnałów. Poszczególne wielkości mogą mieć różną dynamikę, stąd potrzeba innej częstotliwości próbkowania.

Modele tworzone są najczęściej na podstawie samych danych historycznych. W niektórych przypadkach możliwe jest przeprowadzenie specjalnych pomiarów opartych na odpowiednio zaprojektowanych eksperymentach [75]. Dla dużych obiektów może on niestety być kosztowny. W zamian otrzymać można dane zarówno dla normalnych, jak też specyficznych stanów procesu. Wybrane zmienne mogą być mie-

rzony z większą częstotliwością oraz w miarę potrzeb z lepszą dokładnością, przez wykorzystanie dodatkowego sprzętu pomiarowego.

Podczas tworzenia modelu czujnika wyróżnić można kilka etapów. Kroki początkowe wymagają dużego nakładu pracy ze względu na jedynie ograniczoną możliwość ich automatyzacji. Pierwszym z nich jest przyjrzenie się dostępnym danym oraz ogólnej charakterystyce procesu. W jego trakcie należy wybrać docelowe wielkości modelowania oraz rodzaj modelu, jaki może być zastosowany. W kroku tym zdecydować można o potrzebie zastosowania modelu liniowego lub nieliniowego oraz wyborze zbioru jego wielkości wejściowych i wyjściowych.

Wstępne przetwarzanie danych ma na celu ich przystosowanie do modelowania właściwego. W kroku tym dane muszą zostać znormalizowane lub przeskalowane. Należy wykryć wartości odbiegające od prawidłowych, wykryć oraz skorygować dryft oraz ewentualne opóźnienia w poszczególnych zmiennych. Poradzić sobie należy również z ewentualnym brakiem określonych danych. Wybrać trzeba również reprezentatywne dane historyczne. Wymagana być może filtracja danych w celu usunięcia zakłóceń wysokoczęstotliwościowych oraz offsetu. Wybór istotnych wielkości wejściowych przeprowadza się najczęściej za pomocą analizy korelacyjnej. Korelację oblicza się między poszczególnymi wielkościami mierzonymi a wyjściową wielkością modelu. Wybór właściwych sygnałów wejściowych może również polegać na wiedzy eksperckiej.

Do redukcji wymiaru danych podczas budowania czujników softwareowych wykorzystuje się metodę (ang. *Principal Component Analysis* – PCA). Kolinearność danych jest wynikiem nadmiarowej liczby czujników oraz korelacji między poszczególnymi zmiennymi w procesie. Oznacza to mniejszą liczbę informacji, a często nawet jej niedostatek czy brak pomimo wielu dostępnych sygnałów. Naturalną metodą radzenia sobie z kolinearnością jest użycie PCA. Inną metodą jest wybranie najbardziej reprezentatywnych sygnałów i wykorzystanie ich ograniczonego zbioru do budowy modelu.

Najważniejszym krokiem jest identyfikacja modelu, która wymaga wyboru jego optymalnej struktury. Wykorzystywana jest jedna z wielu metod regresji do powiązania zależności wejściowych z wyjściowymi dla modelu. Wybór ten zależy od doświadczenia i preferencji eksperymentatora. W samym procesie wyboru zacząć można od prostych modeli. W razie ich niedostatecznych właściwości wybierać należy kolejno bardziej skomplikowane modele o lepszych właściwościach, przy czym dla optymalnego wyboru należy wykonać ocenę modelu. Ewaluację modelu przeprowadzić można przez ocenę rozbieżności jego działania dla danych testowych. W związku ze zmianami długoterminowymi określonego procesu takimi, jak dryft, model wymaga działań utrzymujących jego właściwości.

Pod wpływem zarówno dryftu czujników sprzętowych, jak też pod wpływem zmienności samego procesu lub warunków otoczenia wykorzystywane dane mogą również podlegać dryftowi. Jego cechą charakterystyczną jest wolna zmienność

w czasie. Kluczowe jest odróżnienie poszczególnych rodzajów dryftu. W przypadku czujników sprzętowych korekcja polega na przeprowadzeniu rekaliibracji czujnika lub korekcji samego sygnału z czujnika. W niektórych procesach wymagana jest adaptacja modelu do pewnych jego zmian czasowych wynikających z różnych przyczyn. Podstawą metod adaptacyjnych są rekursywne poprawki wprowadzane do znanych metod, takich jak: PCA czy PLS. Mechanizm poprawek wykorzystywać może [76, 77]: dane dostarczane on-line, wiedzę ekspercką lub informacje zwrotne dotyczące efektów, osiągnięć i wydajności procesu. Adaptacja polegać może na wykorzystaniu wstępnego przetwarzania danych. Adaptacja w tym kroku zapobiega dostosowywaniu się modelu do chwilowych zakłóceń lub nieprawidłowości w procesie. Podstawowym problemem metod adaptacyjnych jest wybór szybkości wprowadzanych zmian. Jeśli adaptacja będzie oparta na danych ze zbyt długiego okresu, poprawka może nie być w stanie przystosować się do występujących zmian. Adaptacja według danych o zbyt małym historycznym horyzoncie czasowym może nie wykorzystać ważnych informacji zawartych w nieco starszych danych. Jeśli dostępny jest reprezentatywny zbiór danych testujących, ocenić można zmiany adaptacyjne zachodzące w modelu. Jeśli model pierwotny oparty jest na znaczącej liczbie danych historycznych oraz reprezentatywnej wiedzy eksperckiej, ocenić można poziom adaptacji i konieczność dokonywanych zmian. Może się jednak zdarzyć zbyt duża stabilizacja modelu, co zaprzecza samej idei adaptacji.

3.2. FUZJA DANYCH CZUJNIKÓW

Zwiększenie liczby informacji otrzymywanych z wykorzystaniem pomiarów pojedynczym czujnikiem otrzymać można, wykorzystując różne metody. W przypadku braku wystarczająco selektywnych czujników stosowane są odpowiednie metody kompensacji wpływu czynników niepożądanych w postaci sprzętowej, programowej lub mieszanej. Metoda wieloczujnikowej fuzji danych jest przykładem metody mieszanej, w której niezbędne są dodatkowe czujniki, z których z kolei sygnał jest poddawany fuzji przeprowadzanej w postaci algorytmicznej. Fuzję danych czujników można przedstawić w postaci graficznej, co ułatwia jej interpretację [78, 79].

Fuzja czujników jest metodą na przezwycięzenie ograniczeń i niedoskonałości, jakie mają pojedyncze czujniki fizyczne. Do niedoskonałości tych zaliczyć można nieliniowość, czułości skrośne oraz niepożądane właściwości dynamiczne. Wymienione wady powodują konieczność korekcji wyniku pomiaru, co można uzyskać tylko za pomocą dodatkowych informacji. Najprostszym rodzajem fuzji jest wykorzystanie dodatkowych czujników do ich korekcji sygnału z czułością skrośną [80]. Przeważnie czujniki z małymi czułościami skrośnymi są dużo droższe, bardziej zawodne. Fuzja jest ekonomicznie opłacalna. Zmniejszenie wpływu czynników niepożądanych uzyskać można także, wykorzystując kilka czujników tego samego rodzaju o różniących się między sobą parametrach. Często z wykorzystaniem pojedynczego czujnika po-

miar określonej wielkości nie jest możliwy. Jednoczesne pozyskiwanie sygnałów z wielu czujników ogranicza wpływ niejednoznaczności, sprzeczności oraz błędów w pozyskanej z pomiaru wielkości. Dodatkowe sygnały mogą korygować oraz kompensować informacje z jednego czujnika. Dzięki czemu informacja staje się bardziej niezawodna, a przez to przydatna. Jest to efekt synergii poprawiający rozdzielczość pomiaru, jego dokładność czy też zmniejszający zakłócenia [81].

W niektórych przypadkach dzięki metodom fuzji możliwy jest pomiar wielkości, które są w inny sposób bardzo trudne lub nawet niemożliwe do obserwacji. Kilka źródeł danych daje pełniejszą i bardziej niezawodną informację, która jednocześnie ułatwia podjęcie właściwych decyzji. Zalety zastosowania wielu czujników to: zwiększenie odporności na awarię pojedynczego czujnika, zwiększenie obszaru pokrycia – informacja określa cały obiekt (proces), zwiększenie rozdzielczości analizy, zwiększenie dokładności informacji oraz zmniejszenie wpływu interferencji. Fuzja może sprawić, że informacja jest dostępna w krótszym czasie z mniejszym kosztem.

Fuzja czujników (ang. *multisensor data fusion*) może być zdefiniowana jako proces łączenia danych z kilku źródeł informacji (sygnałów z czujników), w celu poprawienia właściwości metrologicznych systemu oraz polepszenia jakości otrzymanych danych. Fuzja danych z czujników zajmuje się sposobem integracji wielu czujników, w celu osiągnięcia unikatowych właściwości systemu fuzji. Nie jest to dziedzina jednorodna, lecz korzystająca z wielu różnych gałęzi nauki, takich jak statystyka, przetwarzanie sygnałów oraz sztuczna inteligencja i logika rozmyta [74, 82]. Nie istnieje jedna, najlepsza metoda wykorzystania wielu źródeł informacji odpowiednia dla każdego przypadku. Pomysł integracji informacji z danych pochodzi bezpośrednio od sposobu ich przetwarzania przez człowieka, który do oceny zjawisk i rzeczy wykorzystuje jednocześnie kilka swoich zmysłów. Przykładem może być ocena jakości określonego produktu przez wzrok, dotyk, a w przypadku produktów żywnościowych dodatkowo smak i zapach.

Fuzja wydaje się interesującą technologią wszędzie tam, gdzie w bezpośredni sposób nie jest możliwe uzyskanie wystarczająco dokładnych i niezawodnych danych, szczególnie jeśli jest potrzeba uzyskiwania tych danych w czasie rzeczywistym. Sytuacja taka występuje w przypadku monitorowania różnego rodzaju procesów. Wśród nich wyróżnić z całą pewnością można procesy chemiczne, w których liczba kontrolowanych zmiennych może być znaczna [74]. Integracja danych z wielu czujników jest wykorzystywana w systemach kontroli procesów przemysłowych. Dzięki fuzji przekazywane do niego dane mają lepsze właściwości i nie wymagają znacznych nakładów obliczeniowych. Z perspektywy kontroli procesu informacja z systemu fuzji powoduje, że system sterowania otrzymuje informacje o lepszej wartości bez potrzeby jej przetwarzania. O wykorzystaniu tej informacji decyduje sam system sterowania procesem. Wymiana czujników lub nawet całych systemów pomiarowych nie musi oznaczać konieczności rekonfiguracji systemu sterowania.

Czujniki, które reagują na wybraną wielkość fizyczną nazywane są sondami wielkości procesu (ang. *proces quantity probes* [46]). Gdyby do każdej wielkości fizycznej

i procesowej dostępny był czujnik o idealnej charakterystyce, fuzja nie byłaby konieczna. Aby mówić o fuzji musimy mieć do czynienia z co najmniej dwoma czujnikami. W praktyce dodatkowy pojedynczy czujnik najczęściej służy do kompensacji lub korekcji sygnału z czujnika podstawowego. Jeśli zwiększona zostanie liczba czujników, mówimy o sieci czujników lub o ich macierzy. System czujników jest zbiorem kilku czujników wraz z infrastrukturą niezbędną do ich funkcjonowania. W procesie przetwarzania sygnału każdemu czujnikowi w macierzy przypisywana jest jednakowa waga. Gdy używane są dodatkowe czujniki służące do korekcji sygnału czujnika podstawowego, wpływ ich sygnału na wielkość wyjściową pomiaru nie jest jednakowy. Fuzja danych występuje, jeśli określona liczba czujników pojedynczych lub w postaci macierzy służy temu samemu celowi. Celem konstrukcji zbioru czujników tego samego rodzaju w postaci macierzy jest ograniczenie błędów przypadkowych i systematycznych.

W macierzach czujników czułość skrośna jest wykorzystywana do określania wielu wielkości, np. różnorodnych związków chemicznych podczas pomiarów gazów. Identyfikacja poszczególnych składników polega na wykorzystaniu określonego wzorca tworzonego przez wektor sygnałów pochodzących z macierzy czujników. Wykorzystane mogą zostać czujniki homogeniczne lub niehomogeniczne. Dla tych drugich sygnały z czujników nie mogą zostać bezpośrednio połączone, lecz wymagają wstępnego dopasowania. W dużej macierzy czujników zmienna czułość na poszczególne składniki umożliwia efektywniejsze wyznaczanie zawartości składników, jakimi jesteśmy zainteresowani. Większa liczba czujników w macierzy zwiększa odporność na interferencje oraz odporność na uszkodzenia pojedynczego czujnika. W pomiarach zawartości gazów o złożonym składzie oraz stosunkowo niewielkich ich zawartościach, jak na przykład w pomiarach zanieczyszczenia powietrza, wykorzystuje się nawet trzy różne rodzaje czujników w macierzy. Rozwiązanie takie zapewnia otrzymanie informacji niezależnych, które dodatkowo mogą być wzmocnione przez zmiany temperatury czujników. Macierze mogą być również czujnikami wieloparametrowymi potrafiącymi wykrywać różne wielkości (parametry) fizyczne oraz chemiczne [83]. Macierz taka może być wykonana w postaci jednego układu scalonego.

Cechą charakterystyczną systemu fuzji czujników jest przetwarzanie informacji na wyższy poziom reprezentacji [84]. Jeśli fuzji podlegają sygnały z pojedynczych czujników inteligentnych, to dodatkowo ich sygnał może być wstępnie przetworzony oraz zunifikowany, przez co integracja może być mniej skomplikowana i wymagać mniejszych zasobów obliczeniowych. W dużych systemach fuzja odbywa się w kilku krokach, z których każdy powoduje zmianę sposobu reprezentacji danych. Kilka kroków fuzji zmniejsza skomplikowanie całego systemu. Izolacja poszczególnych podsystemów fuzji sprawia, że informacja pochodząca z nich wymaga mniejszej obróbki na wyższych poziomach. Działanie takie może również zmniejszać czas obróbki danych. Wyróżnić można trzy podstawowe poziomy fuzji [85]:

- niskiego poziomu, wykorzystuje surowe dane z kilku czujników i przedstawia je w bardziej informacyjnej postaci,

- pośredniego poziomu, wykorzystuje przetworzone cechy sygnału,
- wysokiego poziomu, używa poziomu decyzji; dla każdego sygnału modele są budowane niezależnie, a łączone są wyniki z każdego modelu.

Podany podział poziomów jest podobny do zaprezentowanego w pracy [86] jako poziom przetwarzania sygnałów. Pierwszy poziom fuzji (fuzja surowych danych) wymaga centralnego przetwarzania danych. Dlatego też dane z wszystkich czujników muszą być dostępne jednocześnie. Pośredni poziom fuzji (fuzja cech) wymaga od poszczególnych czujników wstępnego przetwarzania sygnałów. Fuzja na poziomie decyzji wymaga od czujnika podejmowania decyzji na wyższym poziomie.

Najważniejszą częścią fuzji jest sposób integracji wielu sygnałów z czujników. Wyróżnić można trzy główne metody konfiguracji (integracji) fuzji czujników [81]:

- Komplementarna – sygnały z czujników nie zależą od siebie. Mogą reagować na inne wielkości, inny obszar procesu czy fragment obiektu. Stąd sygnały z kilku czujników dają pełniejszy obraz. Pojedynczy czujnik jest w stanie określić tylko kilka cech z całego zbioru, jakim określić można obiekt. Zebranie niezależnych cech z poszczególnych czujników przedstawia pełen obszar cech.
- Konkurencyjna (nadmiarowość) – czujniki mierzą tę samą wielkość fizyczną. Rozwiązanie to może służyć tolerowaniu awarii jednego lub kilku z czujników.
- Kooperacja – współdziałanie czujników służy otrzymaniu informacji niedostępnej przez pomiar pojedynczym czujnikiem. Kooperacja może dotyczyć również jednakowych czujników umieszczonych jednak w różnych miejscach, przez co dostarczana informacja przestrzenna jest niedostępna przez pojedynczy czujnik.

W zależności od rodzaju fuzji może być ona wykonywana na różnym poziomie. Informacje nadmiarowe zazwyczaj można łączyć na najniższym poziomie. Sygnały komplementarne zazwyczaj łączy się na najwyższym poziomie.

Fuzja bezpośrednia dotyczy sytuacji, gdy wykorzystywane są czujniki homogeniczne lub heterogeniczne w postaci sprzętowej lub softwareowej. Fuzja pośrednia występuje, gdy dodatkowo używana jest informacja *a priori* oraz dane przekazywane są przez operatora [85]. Informacja *a priori* często zarówno zmniejsza złożoność systemu, jak też umożliwia otrzymanie bardziej wartościowych informacji przy mniejszym stopniu przetwarzania sygnałów z czujników. Informacje *a-priori* mogą być pozyskane za pomocą matematycznego modelowania obiektu. Łączenie tych dwu rodzajów informacji uzyskać można za pomocą logiki rozmytej, jak to zostało pokazane w przypadku obliczania średniej temperatury przepływającego powietrza za pomocą ograniczonej liczby czujników [87].

Fuzja w przypadku czujników w procesach przemysłowych może być oparta na modelu samego procesu [86] zbudowanym już w fazie jego projektowania. Wykorzystać można również zależność między wielkościami, jakie są dostępne pomiarowo, a wielkościami będącymi w zakresie zainteresowań w postaci modelu empirycznego. Innym przykładem fuzji z wykorzystaniem modelu są pomiary tomograficzne, w których możliwy jest pomiar jedynie na brzegu badanego obszaru, którym jesteśmy za-

interesowani. Podobnym przykładem są zdalne pomiary spektralne w atmosferze, w których za pomocą widma docierającego do sondy określane są właściwości ścieżki obserwacji.

Informacje *a priori* mogą być także wprowadzane w postaci regularyzacji, szczególnie wtedy, gdy poszukiwane jest rozwiązanie zagadnienia odwrotnego źle postawionego w sensie Hadamarda. Wyróżnić można trzy podstawowe metody regularyzacji [88]:

- klasyczna opierająca się na członie kary, który jest pewnym wskaźnikiem energii,
- maksymalnej entropii, oparta na teorii informacji,
- podejścia Bayesa, oparta na wstępnym prawdopodobieństwie.

Za fuzję rozpatruje się również łączenie sygnałów z czujników poddanych dekompozycji falkowej. Jeśli model regresji będzie wykonywany z wykorzystaniem transformaty falkowej (regresja z wykorzystaniem falek), to współczynniki na poszczególnych skalach, jak też sam model, mogą być wybierane w różnych sposób. Wybór jest odpowiednikiem fuzji danych na różnych poziomach [89]:

- Fuzja niskiego poziomu odpowiada wyborowi współczynników ze wszystkich poziomów dekompozycji na podstawie określonych kryteriów, takich jak wariancja.
- Fuzja średniego poziomu odpowiada wyborowi współczynników osobno dla każdego poziomu dekompozycji, a następnie ich połączenie w modelu regresji.
- Fuzja wysokiego poziomu odpowiada zbudowaniu modelu dla każdego poziomu dekompozycji, a następnie fuzji wyników dla każdego z nich.

Wykorzystanie transformaty falkowej w wielowymiarowej regresji razem z fuzją danych skutkuje lepszą eliminacją szumu oraz zakłóceń, zmniejsza nadmiarowość i prowadzi do prostszych modeli regresji [90].

3.3. METODY ZESPOŁÓW MODELI

Uczenie zespołowe (ensemble) jest pewną metodą budowania modeli predykcyjnych (klasyfikacji, regresji) składających się ze zbioru kilku modeli. Za pomocą takiego łączenia modeli może być budowany skomplikowany model wykorzystujący określoną liczbę modeli prostych. Zaletą modeli ensemble jest zwiększenie dokładności, stabilności oraz odporności na czynniki zakłócające i warunki pomiarowe. Dzięki połączeniu poszczególnych modeli bazowych uzyskać można lepszą generalizację w porównaniu do każdego z poszczególnych modeli w zespole. Zespół modeli cechują lepsze właściwości niż dowolnego złożonego modelu pojedynczego, opartego na całym zbiorze kalibracyjnym. Jakość tego typu algorytmu zależy od różnorodności poszczególnych modeli członkowskich [91]. Różnorodność osiągnąć można przez zastosowanie kilku metod przedstawienia danych lub kilku metod uczenia (rodzaj modelu).

Poszczególne modele członkowskie powinny cechować brak korelacji błędów. Modele różnią się między sobą, jeśli popełniają inne błędy dla nowych danych nie użytych do ich budowania.

Odróżnić należy modele typu ensemble od klasycznego budowania modeli globalnych ze zbioru modeli lokalnych. W tym ostatnim przypadku modele członkowskie podlegają uczeniu w poszczególnych podprzestrzeniach. Jest to więc rozłożenie dużego problemu na odseparowane od siebie fragmenty. Budowanie modeli typu ensemble podzielić można na dwa podstawowe zagadnienia [92, 93]:

- W jaki sposób stworzyć kolekcję modeli członkowskich (ang. *ensemble generation*).
- W jaki sposób połączyć modele członkowskie w całość (ang. *ensemble fusion*).

Wytworzony zbiór modeli może zawierać modele o bardzo podobnych właściwościach, które nie wnoszą żadnej dodatkowej różnorodności, stanowiąc jedynie nadmiarowość ilościową. W takiej sytuacji niektóre modele członkowskie powinny zostać usunięte ze zbioru. Metody zmniejszania liczby podzbiorów członkowskich są różne, a ich właściwości są oceniane dla bardzo dużej liczby podzbiorów [94], co czyni je mało praktycznymi w typowych zastosowaniach czujnikowych.

Podział metod ensemble [95]:

- Metody generatywne – tworzą zbiór modeli bazowych przez różne metody uczenia lub generują podzbiory do uczenia poszczególnych modeli. Metody z tej grupy starają się osiągnąć zwiększenie różnorodności poszczególnych modeli w grupie. Do tej metody zaliczamy metody resamplingu oraz wyboru podzbiorów.
- Metody niegeneratywne – metody zajmują się głównie łączeniem poszczególnych modeli lub ich wyborem, natomiast w mniejszym stopniu metodami tworzenia modeli bazowych. Wykorzystywane są predefiniowane zbiory modeli. Do metod tych zaliczamy metody fuzji (np. głosowanie większościowe) oraz selekcji modeli (wybór najlepszego).

Do najbardziej popularnych metod wyboru poszczególnych zbiorów uczących należą bagging (ang. *bootstrap aggregating*) [96], boosting oraz wybór losowy [97, 98].

W metodzie bagging zbiór uczący jest losowo próbkowany wielokrotnie, tworząc podzbiory uczące o tym samym rozmiarze. Bagging używa jednakowych wag dla wszystkich modeli. Bagging zmniejsza wariancję błędów pod warunkiem, że użyte modele regresji mają dużą wrażliwość parametrów modelu na zmiany w zbiorze uczącym. Uśrednianie tą metodą zmniejsza błędy spowodowane próbkami oddalonymi w zbiorze kalibracyjnym.

Boosting polega na dopasowaniu kolejnego zbioru na podstawie właściwości poprzedniego. Boosting służy do zmniejszenia odchylenia predykcji. W tej metodzie stosuje się wagi dla poszczególnych modeli w zależności od ich właściwości, jakości dopasowania do danych uczących. Boosting jest algorytmem sekwencyjnym, powtarzanym kilkakrotnie. Z każdym nowym krokiem algorytm zwraca większą uwagę na

punkty, które w poprzednich modelach były znacznie oddalone od celu. Boosting jest ogólną metodą zwiększania dokładności algorytmów uczących, do takich algorytmów należy AdaBoost [99, 100]. AdaBoost tworzy model o dobrych właściwościach na postawie ważonej sumy modeli o słabych właściwościach.

Bagging ma lepsze właściwości niż boosting dla zaszumionych danych. Dla danych mało zaszumionych lepszy jest boosting [101]. Dla 33 przeprowadzonych eksperymentów klasyfikacji w większości przypadków lepszy okazał się boosting [102].

Nieco inne podejście do metod uczenia zespołowego polega na rywalizacji poszczególnych modeli w zespole i wyborze (selekcji) takiego, który w danym obszarze jest najlepszy [93]. Stworzony w ten sposób model globalny, w którym poszczególne modele członkowskie są wybierane przez przełącznik z wykorzystaniem odpowiednich reguł, przy czym dla algorytmów selekcji wyjście modelu jest wyznaczone w różny sposób w zależności od wyjść modeli członkowskich, a w metodzie integracyjnej istnieje na to określony i zawsze taki sam algorytm.

Metody uczenia zespołowego są używane w analizie składu substancji. Dla systemów elektronicznego nosa użyte zostały 4 różne metody regresji, w tym maszyny wektorów wspierających, sieci neuronowych oraz dwóch rodzajów regresji wielorakiej do wykrywania typu odoru oraz wyznaczania jego koncentracji [103].

Spektroskopia nie jest typowym przypadkiem zastosowania metod ensemble. W danych spektroskopowych przeważnie jest bardzo wiele zmiennych objaśniających, przy stosunkowo niewielkich zbiorach kalibracyjnych. Metody ensemble zastosowane dla spektroskopii muszą powodować wirtualne zwiększenie liczby danych kalibracyjnych.

Podział zbioru kalibracyjnego z zespołem sieci neuronowych zastosowano z powodzeniem w kalibracji spektrometrii bliskiej podczerwieni [104]. Metoda ta sprawdza się w stosunku do sieci neuronowych, które mają tendencję do wyszukiwania lokalnych minimów.

Consensus modeling jest to metoda modelowania polegająca na budowaniu wielu modeli członkowskich, które następnie służą do określania pewnej wielkości modelowanej. Podobną metodą jest consensus PLS, która wyszukuje poszczególne korelacje między zmiennymi, jakie nie mogłyby być wyznaczone z pomocą pojedynczego modelu [105]. Wykorzystanie kilku modeli PLS zapewnia polepszenie jakości pełnego modelu.

W przypadku modeli budowanych z widm metoda ich zbioru może wykorzystywać strategię wyboru zmiennych do każdego z modeli członkowskich. Metody wyboru zmiennych oparte na wyborze przypadkowym mogą generować część modeli członkowskich o bardzo słabych właściwościach. Stąd też opracowywane są strategie wyboru podprzestrzeni, oparte np. na grupowaniu danych (ang. *cluster analysis*) [106].

Model zespołu modeli regresji wielorakiej z losowym wyborem podprzestrzeni (ang. *Random Subspace, Multiple Linear Regression – RS-MLP*) ma dobre właściwo-

ści do celów kalibracji w bliskiej podczerwieni [107]. Za główną zaletę losowego wyboru podprzestrzeni należy uznać uproszczenie algorytmu w stosunku do algorytmu bagging czy boosting. Wyborowi losowemu do budowy każdego z modeli członkowskich podlegają długości fali.

Do tworzenia różnorodnych danych uczących można użyć również technikę *noise augmentatnion*, czyli dodawania różnego rodzaju szumu do tych samych danych [108]. Dodanie szumu zmniejsza prawdopodobieństwo przetrenowania modelu. Metoda ta zmniejsza również czułości na interferencje i czynniki wpływające. Niewątpliwą zaletą jest również zmniejszenie liczby wykonanych punktów kalibracji, co skraca czas oraz zmniejsza koszty jej wykonania [109]. Metoda sprawdza się bardziej dla modeli PLS niż bagging. Porównanie przeprowadzono dla kombinacji następujących rodzajów szumu [110]:

- lokalne przesunięcie pików widma przez dodanie pochodnej widma,
- szum multiplikatywny,
- szum addytywny,
- szum zależny od intensywności,
- szum mieszany powstający z dodania powyższych rodzajów szumu.

4. ALGORYTMY KALIBRACJI W SPEKTROSKOPII

W przypadku pomiarów wielkości chemicznych duże znaczenie dla dokładności pomiaru ma wstępne przygotowanie próbki oraz jej pobranie. Dla pomiarów duże znaczenie mają również czynniki dodatkowe. Kalibracja do pomiarów chemicznych wymaga wielu kroków i jest bardziej skomplikowana niż w przypadku pomiarów fizycznych. Dodatkowo podczas kalibracji potrzebny jest zbiór standardów dla pomiarów chemicznych, podczas gdy dla pomiarów fizycznych często wystarczy jeden. Kalibracja razem z wstępnym przygotowaniem próbek i innymi niezbędnymi czynnościami w pomiarach analitycznych nazywana jest kalibracją procesu (ang. *process calibration*) [111].

Kalibrację określić można jako operację uzyskania zależności między wartościami wskazywanymi przez instrument pomiarowy lub systemem pomiarowym a wartościami reprezentowanymi przez mierzony standard (próbkę kalibracyjną) [112], przy czym operacja ta jest przeprowadzana w określonych warunkach.

Wyróżnić można kalibrację bezpośrednią i pośrednią [113]. Kalibracja bezpośrednia polega na porównaniu mierzonej wielkości z wielkością standardową, jest ona wykonywana tylko w jednym punkcie pomiarowym. Cechą charakterystyczną porównania jest to, że zarówno próbka kalibracyjna (standard), jak też mierzona są tej samej wielkości. Po wykonaniu pomiaru bezpośredniego otrzymuje się wartość korekcji, którą dodaje się, aby wykonać kompensację błędu systematycznego (co oznacza się określeniem bias).

W kalibracji pośredniej wskazania instrumentu pomiarowego różnią się wielkością fizyczną od mierzonej. Kalibracja pośrednia jest to więc przekształcenie wskazań (sygnału) instrumentu pomiarowego w wielkość mierzoną. Aby możliwe było utworzenie empirycznej zależności funkcyjnej między wskazaniami instrumentu pomiarowego a wielkością mierzoną, potrzebny jest zbiór standardów. Wyróżnić można kalibrację instrumentu analitycznego oraz kalibrację procesu [113]. Pierwszy rodzaj jest równoznaczny z kalibracją fizyczną, w której próbki są mierzone bez poddawania ich wcześniejszej modyfikacji. W uproszczeniu kalibracja instrumentu polega zarówno na wybraniu odpowiedniego modelu (identyfikacja, struktura matematyczna modelu), estymacji jego parametrów, jak też sprawdzeniu poprawności modelu dla danych testowych [114]. To właśnie właściwa weryfikacja stanowi o poprawności przyjętego modelu oraz wykonanej kalibracji. W przypadku pomiarów chemicznych (procedury analityczne) przeprowadzane są dodatkowe operacje, takie jak przygotowanie próbek do pomiarów.

W praktyce kalibracja jest to określenie empirycznej zależności funkcyjnej $y = f(x)$ zmiennej y (zmiennej zależnej) w funkcji zmiennej x (niezależnej). Przy czym pomiar wielkości x jest łatwiej dostępny, szybszy lub znacznie tańszy od bezpośredniego pomiaru y . Dane używane w procedurze znajdowania parametrów kalibracyjnych (postaci funkcji lub jej parametrów) nazywane są zbiorem kalibracyjnym. Następnie można przeprowadzać wyznaczanie wartości zmiennych niezależnych na podstawie pomiarów zmiennych zależnych. Mówimy wtedy o predykcji lub też o testowaniu. Według poziomu znajomości zależności zmiennej x oraz y możemy wyróżnić następujące sytuacje:

- Zależność jest ściśle określona przez prawo w postaci równania matematycznego o znanych wszystkich parametrach.
- Zależność jest ściśle określona przez prawo w postaci równania matematycznego, ale nie znamy wartości wszystkich jego parametrów.
- Zależność jest teoretycznie określona, lecz nie można jej bezpośrednio wykorzystać lub też teoretyczna zależność matematyczna nie istnieje.

W ostatnim przypadku modele kalibracji są czysto empiryczne. Modele kalibracji mogą być modelami liniowymi lub nieliniowymi.

Obliczeniową (matematyczną) część kalibracji wykonują w większości specjalizowane programy komputerowe. Dane wejściowe i wyjściowe są wtedy traktowane jako wejście do programu oraz wynik jego obliczeń. W tym sensie kalibracja jest rozumiana nieco inaczej. Kalibracja jest to dopasowanie zbioru parametrów wejściowych kodu związanego z określonymi obliczeniami, tak by wynik obliczeń oraz zbioru danych eksperymentalnych zgadzał się w jak największym stopniu [115]. Wymaga to ilościowego określenia zgodności. Po wykonaniu procesu kalibracji wymagane jest wykonanie procesu walidacji (kontroli prawidłowości, zgodności). Stwierdzenie ważności jest to określenie pewności w stosunku do możliwości predykcji danego programu (kodu) dla danego zastosowania w stosunku do danych eksperymentalnych. Zarówno kalibracja, jak i walidacja wymaga rygorystycznego porównania z wzorcem, odpowiednim dla danego zastosowania. W przypadku kodu komputerowego nastąpić

powinna również weryfikacja, czyli potwierdzenie, że dany kod poprawnie implementuje dany algorytm, tzn. reprezentuje model kalibracji.

W związku z kalibracją można również mówić o predykcji, czyli obliczeniach, które przewidują wartość lub zbiór wartości poprzedzający fizyczne pomiary. Predykcja jest to więc przewidywanie pomiarów, określanie właściwości zjawiska fizycznego bez jego bezpośrednich eksperymentalnych obserwacji.

Wzorzec są to informacje, co do których mamy pewność, że są właściwe i prawdziwe do celów zarówno kalibracji, jak i walidacji. Informacje te są porównywane z wynikami obliczeń (predykcji) [115].

Do estymacji parametrów modelu używa się najczęściej liniowych metod regresji najmniejszych kwadratów. Regresję można rozpatrywać jako estymację (kalibracja kontrolowana) oraz predykcję (kalibracja naturalna) [116]. Kalibracja naturalna polega na pomiarach par wartości (x, y) pewnej populacji. Jeśli następnie poszukujemy pewnego komponentu x , możemy go przewidzieć na podstawie regresji x na y . Dopasowanie regresji do próbek kalibracyjnych pomaga estymować potrzebny parametr. W innym podejściu uważamy nieznaną wartość x za pewną stałą do estymacji. Wtedy stosujemy regresję y na x . O wartości x wnioskujemy przez odwrócenie zależności regresji. Ostatni sposób jest również nazywany regresją pośrednią, ponieważ sama regresja nie daje tu wyniku [116].

Problem regresji to znalezienie modelu jednej lub kilku zmiennych zależnych, odpowiedzi Y , w zależności od zbioru zmiennych niezależnych X . Wartość wyznaczana y może być funkcją jednej lub wielu zmiennych.

$$y = f(x_1, x_2, x_3, \dots, x_n) \quad (26)$$

Sam model f może być przedstawiony w postaci funkcji matematycznej lub w postaci mniej bezpośredniej jako sztuczna sieć neuronowa, funkcje bazowe, funkcje sklepane, stabilaryzowane parametry. Podstawowym modelem używanym w kalibracji jest model liniowy następującej postaci:

$$y = b_0 + b_1x_1 + \dots + b_2x_2 + b_3x_3 + \dots + b_nx_n + e \quad (27)$$

gdzie: b_0 – punkt przecięcia z osią współrzędnych, b_1, \dots, b_n – współczynniki regresji, n – liczba zmiennych, e – pozostałość (błąd).

Często do kalibracji używa się danych, z których została usunięta wartość średnia, co oznacza, że wartość b_0 wynosi zero. Parametry modelu $b_0 - b_n$ estymowane są z wykorzystaniem zbioru kalibracyjnego, zwanego także zbiorem trenującym, który zawiera wartości x oraz y dla m próbek. Ocena parametrów modelu następuje z wykorzystaniem zbioru testowego. Takie podejście do danych nazywane jest uczeniem nadzorowanym.

Wszystkie metody regresji polegają na zmniejszaniu wartości pozostałości. Najbardziej znaną metodą jest minimalizacja sumy kwadratów pozostałości. Ocena modelu jest zadaniem dość skomplikowanym i opiera się nie tylko na samej analizie błęd-

du dopasowania do danych trenujących. Często używana miara jakości modelu to standardowy błąd predykcji, a więc miara określająca dopasowanie modelu do danych testowych, a nie danych kalibracyjnych. Możliwości wiarygodnej kalibracji oraz walidacji modelu zależą od wielkości posiadanego przez nas zbioru danych. Zazwyczaj stosuje się podział zbioru danych kalibracyjnych na zbiór treningowy oraz zbiór testowy [117]. Jeśli liczba danych jest duża, można je podzielić na zbiór treningowy służący do tworzenia modelu (np. 50%), zbiór kontrolny (np. 25%) mogący służyć do optymalizacji modelu oraz zbiór testowy (np. 25%), za pomocą którego określamy jakość modelu. Istnieją specjalne metody podziału danych na podane zbiory, ale używa się również metody wyboru przypadkowego.

Użycie bardziej skomplikowanego modelu zawierającego więcej parametrów spowoduje zmniejszenie błędu dopasowania do zbioru kalibracyjnego. Bardzo skomplikowany model może zmniejszyć błąd pozostałości nawet do zera, między wartością eksperymentalną a modelowaną y . Modele takie mogą nie mieć właściwości odpowiedniej generalizacji i dopasowania do danych testowych. Jest więc określona optymalna liczba parametrów, dzięki którym model jest poprawny w określonym sensie.

Problemy pojawiające się podczas kalibracji to brak selektywności, kolinearność danych, nieliniowość, błędy grube w danych kalibracyjnych. Problem pierwszy w spektroskopii oznacza, że nie ma jednej długości fali, na podstawie której wyznaczyć można zawartość danego składnika. Na wszystkich mierzonych długościach fali transmisyjność pochodzi od dwóch lub większej liczby składników. Problem selektywności może być rozwiązany przez użycie kilku długości fali w procesie kalibracji, a więc zastosowanie kalibracji wielowymiarowej. Kolinearność zmiennych oznacza liniową lub prawie liniową zależność między nimi, co powoduje kłopoty w klasycznych metodach kalibracji wykorzystujących metodę najmniejszych kwadratów. W przypadku danych nieliniowych stosowane są metody przekształcania ich do zależności liniowej lub ostatecznie stosowane są modele nieliniowe. Ostatni problem wynika z jakości danych kalibracyjnych i pewnego rodzaju odstępstwa części danych od reszty „typowych” danych. Powodem takiej sytuacji może być np. błąd instrumentalny. Odstępstwa takie mogą, jeśli nie są zauważone, powodować znaczne błędy kalibracji.

Podstawowym wskaźnikiem oceny jakości wszystkich modeli kalibracyjnych jest błąd predykcji

$$e_i = y_i - \hat{y}_i \quad (28)$$

Uzyskuje się go na podstawie niezależnego zbioru testującego. Standardowy błąd predykcji określany jest następującym wzorem

$$SBP = \sqrt{\frac{1}{k-1} \sum_{i=1}^k (y_i - \hat{y}_i - bias)^2} \quad (29)$$

$$bias = \frac{1}{k} \sum_{i=1}^k (y_i - \hat{y}_i) \quad (30)$$

gdzie: y_i – wartości eksperymentalne, \hat{y}_i – wartości uzyskane z modelu, k – liczba danych.

Kwadrat średniego błędu

$$KSB = \frac{1}{k} \sum_{i=1}^k (y_i - \hat{y}_i)^2 \quad (31)$$

Błąd średni kwadratowy wyraża się zależnością

$$BSK = \sqrt{\frac{1}{k} \sum_{i=1}^k (y_i - \hat{y}_i)^2} \quad (32)$$

W najprostszej wersji mamy do czynienia z jedną zmienną zależną y oraz jedną zmienną niezależną x , a zależność między nimi jest liniowa. Liniowy związek między zmiennymi określony jest następująco:

$$y = b_0 + bx + e \quad (33)$$

Zadaniem jest znalezienie najlepszej estymaty parametrów b_0 , b_1 , co w praktyce sprowadza się do minimalizacji błędu e . Wykorzystując metodę najmniejszych kwadratów, wartość parametrów przedstawić można następująco:

$$b = \frac{\sum_{i=1}^m (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^m (x_i - \bar{x})^2} \quad (34)$$

$$b_0 = \bar{y} - b\bar{x} \quad (35)$$

gdzie \bar{x} jest średnią arytmetyczną wszystkich x_i a \bar{y} średnią arytmetyczną y_i . Wartość licznika jest proporcjonalna do kowariancji między x oraz y . Mianownik jest proporcjonalny do wariancji x . Jeśli dane zostaną przekształcone, tak by ich wartość średnia wynosiła zero, tzn. $\bar{x} = 0$ oraz $\bar{y} = 0$, otrzymamy $b_0 = 0$ oraz

$$b = \frac{\sum_{i=1}^m x_i y_i}{\sum_{i=1}^m x_i^2} = \frac{\mathbf{x}^T \mathbf{y}}{\mathbf{x}^T \mathbf{x}} \quad (36)$$

Klasyczne podejście najmniejszych kwadratów nie sprawdza się, jeśli w danych trenujących zawarte są punkty o dużym oddaleniu od przewidywanej krzywej regresji. Jakość estymacji ocenić można przez ocenę wariancji błędu.

Opisana metoda często nazywana jest metodą odwrotnej kalibracji (ang. *inverse calibration*). Kalibracją klasyczną nazywana jest metoda opisana następującym równaniem

$$x = a_0 + a_1 y + e \quad (37)$$

Do określania przyszłych wartości używa się następującego równania:

$$\hat{y} = -a_0 / a_1 + (1 / a_1)x \quad (38)$$

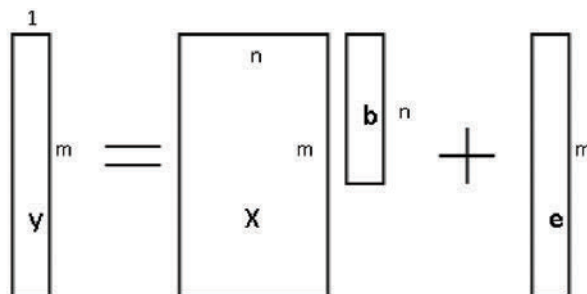
Kalibracja klasyczna była używana głównie przez chemików, natomiast odwrotna przez statystyków. Stąd takie ich określanie i nazwy w chemometrii. Model oparty na jednej zmiennej daje poprawne wyniki jedynie dla pomiarów, w których nie ma zakłócenia. W mieszaninie występować może jedynie interesujący nas składnik. Jeśli próbka zawiera inny składnik, to wynik będzie zafałszowany (występuje tzw. przesunięcie). Nie ma metody wykrywania zakłóceń, a więc przyjmujemy, że wynik jest poprawny bez sprawdzania. Użycie dużej liczby zmiennych mierzonych umożliwia przede wszystkim zredukowanie szumu. Drugą zaletą jest możliwość uzyskania selektywności, dzięki operacjom matematycznym. Nie jest wymagana pełna selektywność instrumentu pomiarowego. Wybór czujnika nie musi opierać się na selektywności spektralnej, lecz również na innych parametrach, takich jak solidność wykonania, mocna obudowa [118], szczególnie ważnych w warunkach przemysłowych. Metody wielowymiarowe umożliwiają wykrywanie danych leżących poza głównym trendem (ang. *outliers*). Dane znacząco różniące się od pozostałych mogą wskazywać na popełnienie pewnego błędu, na który nie zwróciliśmy uwagi. A może nas to kosztować znaczne pogorszenie modelu predykcji. Są to dane, które znacznie odbiegają od pozostałych i nie pasują do głównego kształtu. W zależności od stopnia skomplikowania sygnału do kalibracji potrzebujemy typowo 10–50 próbek [118]. W wyniku operacji wielowymiarowych otrzymujemy również pozostałość, która nie należy do żadnej z poszukiwanych przez nas wartości zmiennych. Może to służyć do określania poziomu ważności modelu, poprawiania go oraz określenia, że w próbkach znajduje się również jakiś inny niespodziewany przez nas składnik.

W wielu przypadkach użycie razem kilku zmiennych przewidujących zwiększa dokładność wyniku i rozwiązuje problem selektywności. Model zawierać może wiele zmiennych x_1, x_2, \dots, x_n , które można mierzyć. Jeśli mierzonych jest m próbek, otrzymamy następujący zbiór równań:

$$\begin{aligned}
 y_1 &= b_0 + b_1x_{11} + b_2x_{12} + \dots + b_nx_{1n} + e_1 \\
 y_2 &= b_0 + b_1x_{21} + b_2x_{22} + \dots + b_nx_{2n} + e_2 \\
 &\vdots \\
 y_m &= b_0 + b_1x_{m1} + b_2x_{m2} + \dots + b_nx_{mn} + e_m
 \end{aligned}
 \tag{39}$$

Wartość y jest powiązana liniowymi zależnościami ze zmiennymi x z uwzględnieniem wartości błędu e , który jest addytywny. W tym przypadku poszukiwane są współczynniki $b_0, b_1, b_2, \dots, b_n$. W zapisie macierzowym otrzymujemy

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e}
 \tag{40}$$



Rys. 1. Regresja wielowymiarowa
Fig. 1. Multivariate regression

Wykorzystując klasyczną metodę najmniejszych kwadratów, wartość wektora \mathbf{b} obliczymy następująco:

$$\mathbf{b} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}
 \tag{41}$$

Wektor błędu wynosić będzie

$$\mathbf{e} = \mathbf{y} - \hat{\mathbf{y}}
 \tag{42}$$

a minimalizacji podlegała będzie suma kwadratów błędów $\mathbf{e}^T\mathbf{e}$.

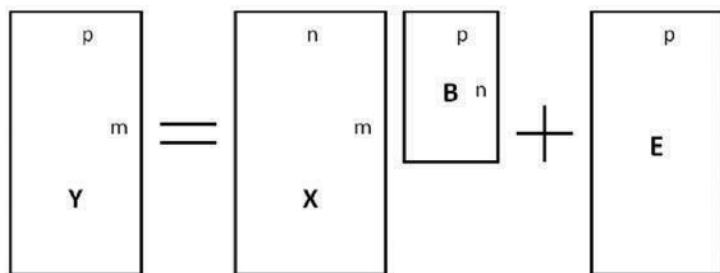
Problem, jaki pojawia się podczas obliczania wektora współczynników \mathbf{b} , to potrzeba odwracania macierzy $\mathbf{X}^T\mathbf{X}$ dla skorelowanych zmiennych x . Oznacza to bowiem niestabilność, a czasami nawet brak możliwości obliczenia macierzy odwrotnej. Innym problemem może być mniejsza liczba obserwacji w stosunku do liczby estymowanych parametrów. W takim przypadku stosuje się metody zmniejszenia liczby zmiennych oraz dodatkowe metody stabilizujące rozwiązanie.

Jeśli na podstawie pomiaru wektora \mathbf{x} (n wymiarów) poszukujemy wektor \mathbf{y} (m wymiarów), mamy do czynienia z regresją wielu zmiennych. Model regresji zapisać można następująco:

$$\mathbf{Y} = \mathbf{X}\mathbf{B} + \mathbf{E} \quad (43)$$

Dla p wartości obserwacji poszczególne macierze mają następujące wymiary \mathbf{Y} $m \times p$, \mathbf{X} $m \times (n+1)$, \mathbf{B} $(n+1) \times p$, \mathbf{E} $m \times p$. Macierz współczynników obliczamy następująco:

$$\mathbf{B} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y} \quad (44)$$



Rys. 2. Regresja wielowymiarowa, kilka zmiennych wyjściowych
Fig. 2. Multivariate regression, multiple variable

Macierz \mathbf{B} może być obliczona dla niekolinearnych wartości x , oraz jeśli $n < m$. Kolejnym praktycznym problemem podczas kalibracji wieloskładnikowej mieszaniny jest konieczność dokładnej znajomości koncentracji wszystkich składników wnoszących jakąkolwiek zmianę absorpcyjności. Jest to bardzo niepraktyczne, a czasami niemożliwe [119].

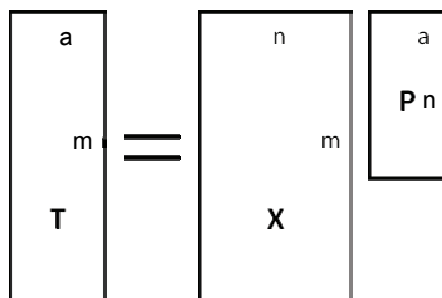
Jedną z metod rozwiązujących problem kolinearności danych jest regresja składowych głównych. Zmniejsza ona liczbę zmiennych regresji, które są liniową kombinacją zmiennych wejściowych. Regresja składowych głównych jest więc kombinacją metody składowych głównych oraz regresji wielu zmiennych. W ten sposób nieważne i powodujące niestabilność informacje są odrzucane, a tylko mająca wpływ część \mathbf{X} jest używana w procedurze regresji. Dodatkowo zmniejszenie wymiarowości danych ułatwia ich wizualizację oraz interpretację. Metoda składowych głównych polega na dekompozycji macierzy danych \mathbf{X} na wektor pasm (*scores*) \mathbf{T} oraz wektor wartości ładujących (*loadings*) \mathbf{P} . Dla pewnej liczby składowych głównych, których jest zwykle mniej niż wynosi wartość rzędu macierzy \mathbf{X} , macierz \mathbf{T} zawiera maksymalną wartość informacji zawartej w \mathbf{X} . Dekompozycje macierzy \mathbf{X} można przedstawić następująco:

$$\mathbf{X} = \mathbf{TP}^T + \mathbf{E} \quad (45)$$

Co oznacza, że macierz \mathbf{X} może być aproksymowana przez określoną liczbę składowych głównych (liczba kolumn \mathbf{T}) oraz odpowiadającym im wartościami ładującym. Liczba składowych głównych zależy od dokładności odwzorowania \mathbf{X} , czyli

minimalizacji sumy kwadratów elementów \mathbf{E} . Składowe główne można obliczyć, wykorzystując dekompozycję singularną $\mathbf{X}^T \mathbf{X}$. Kolumny macierzy \mathbf{P} są wektorami własnymi o długości jednostkowej tej macierzy. Macierz \mathbf{P} może być wtedy określona z wykorzystaniem macierzy kalibracyjnej \mathbf{X}_k równania

$$\mathbf{T} = \mathbf{X}_k \mathbf{P} \quad (46)$$



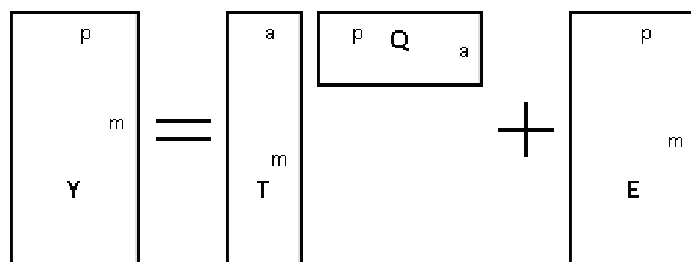
Rys. 3. Dekompozycja macierzy \mathbf{X} na składowe główne
Fig. 3. Principal component \mathbf{X} matrix decomposition

Następnie w równaniu regresji zastępujemy macierz \mathbf{X} macierzą \mathbf{T}

$$\mathbf{Y} = \mathbf{TQ} + \mathbf{E}_s \quad (47)$$

obliczając wartość \mathbf{Q} normalną metodą najmniejszych kwadratów

$$\mathbf{Q} = \mathbf{T}^T \mathbf{Y}_k \quad (48)$$



Rys. 4. Regresja składowych głównych
Fig. 4. Principal component regression

W tym przypadku nie ma konieczności odwracania macierzy, a ostateczny wynik obliczamy następująco:

$$\mathbf{Y} = \mathbf{XPQ} \quad (49)$$

Zachodzi więc tutaj projekcja danych macierzy \mathbf{X} na przestrzeń rozpinaną przez \mathbf{P} . Kompresujemy zawartość danych do znacząco mniejszej macierzy.

Macierz pasm ma następujące właściwości [120]: liczba wierszy jest równa liczbie wierszy macierzy danych \mathbf{X} , liczba kolumn jest równa liczbie ważnych czynników znajdujących się w danych, suma kwadratów elementów każdej kolumny macierzy pasm odnosi się do wartości własnej (*eigenvalue*), która im większa, tym ważniejszy jest dany składnik.

Macierz wartości ładujących ma taką liczbę kolumn, jak liczba kolumn macierzy \mathbf{X} . Liczba wierszy jest równa liczbie znaczących czynników w danych. Każdy wiersz odpowiada jednemu składnikowi (czynnikowi) głównemu. Suma kwadratów wszystkich elementów każdej kolumny równa się jedności.

Interpretacja geometryczna składników głównych wymaga odpowiedniej przestrzeni wielowymiarowej. Każdy główny składnik jest linią w tej przestrzeni. Wektor pasm określa długość wzdłuż tej linii, natomiast wektor ładujący kierunek kątowej linii. Niestety metoda PCR ma wadę modelowania zmienności w macierzy \mathbf{X} , która to może nie mieć nic wspólnego ze zmiennością \mathbf{Y} [121].

Kolejnym krokiem jest zastosowanie, oprócz macierzy \mathbf{X} , również dekompozycji macierzy \mathbf{Y} . Metoda regresji cząstkowej polega na uzyskaniu każdego komponentu przez maksymalizację kowariancji pomiędzy \mathbf{Y} oraz wszystkimi możliwymi liniowymi funkcjami \mathbf{X} . Poszczególne składowe dekompozycji są zatem powiązane ze zmiennością \mathbf{Y} w większym stopniu niż składniki główne.

$$\begin{array}{|c|} \hline p \\ \hline \mathbf{Y} \\ \hline n \\ \hline \end{array} = \begin{array}{|c|} \hline a \\ \hline \mathbf{U} \\ \hline n \\ \hline \end{array} \begin{array}{|c|c|} \hline p & a \\ \hline \mathbf{Q}' & \\ \hline \end{array} + \begin{array}{|c|} \hline p \\ \hline \mathbf{E}_Y \\ \hline n \\ \hline \end{array}$$

Rys. 5. Rozkład macierzy \mathbf{Y} dla PLS
Fig. 5. PLS decomposition of \mathbf{Y} matrix

$$\begin{array}{|c|} \hline n \\ \hline \mathbf{X} \\ \hline m \\ \hline \end{array} = \begin{array}{|c|} \hline a \\ \hline \mathbf{T} \\ \hline m \\ \hline \end{array} \begin{array}{|c|c|} \hline m & a \\ \hline \mathbf{P}' & \\ \hline \end{array} + \begin{array}{|c|} \hline n \\ \hline \mathbf{E}_X \\ \hline m \\ \hline \end{array}$$

Rys. 6. Rozkład macierzy \mathbf{X} dla PLS
Fig. 6. PLS decomposition of \mathbf{X} matrix

Podstawowe zalety metody PLS to możliwość wykorzystania całego zakresu spektrum, co jest szczególnie istotne w wieloskładnikowych miksturach i umożliwia zastosowanie mniejszej rozdzielczości pomiaru. Procedura obliczania wyniku jest szybka i nie wymaga oczyszczania badanych próbek. Zawartość pozostałych składników nie wpływa na wynik, gdyż są one ignorowane przez procedurę, która jest bardzo selektywna. Jej zalety ujawniają się w pełni, gdy, oprócz interesującego nas składnika, badana próbka cechuje się dużą absorpcyjnością pozostałych składników o szeroko zachodzących na siebie widmach.

Kierunek pierwszego składnika PLS jest otrzymywany przez kryterium maksymalizacji kowariancji. Jest to jednostkowy wektor \mathbf{w}_1 nazywany pierwszym wektorem ładującym. Pasma wzdłuż tego kierunku jest obliczane następująco:

$$\mathbf{t}_1 = \mathbf{X}\mathbf{w}_1 \quad (50)$$

Co oznacza, że wszystkie zmienne z \mathbf{X} są sprowadzane do \mathbf{t}_1 w celu otrzymania wektora ładującego \mathbf{p}_1 . Podobnie jest otrzymywany współczynnik \mathbf{q}_1 , przez sprowadzenie \mathbf{Y} do \mathbf{t}_1 . Następnie wynik \mathbf{t}_1 i \mathbf{p}_1 jest odejmowany od \mathbf{X} , natomiast $\mathbf{t}_1\mathbf{q}_1$ jest odejmowany od \mathbf{Y} . Następnie poszukiwany jest drugi kierunek dla pomniejszych macierzy \mathbf{X} i \mathbf{Y} . Proces jest kontynuowany do momentu znalezienia odpowiedniej liczby komponentów.

Metoda PLS wymaga mniej komponentów (składowych) w stosunku do PCR, co ułatwia również interpretację.

Historycznie pierwszą, prostą, ale efektywną metodą wyznaczania parametrów PLS jest nieliniowa metoda iteracyjna (ang. *Non-linear Iterative Partial Least Squares* – NIPALS). To właśnie od niej wzięła się nazwa metody PLS. Przed zastosowaniem algorytmu dane należy przeskalować i wycentrować. Dla jednej zmiennej y algorytm ogranicza się do jednej iteracji. Algorytm NIPALS przedstawić można następująco [122]:

- Wybieramy startowy wektor \mathbf{u} , najczęściej jest to jedna z kolumn \mathbf{Y} .
- Obliczamy wagi wektora \mathbf{X} :

$$\mathbf{w} = \mathbf{X}'\mathbf{u} / \mathbf{u}'\mathbf{u} \quad (51)$$

- Normujemy wartość wektora, tak by $\|\mathbf{w}\| = 1,0$
- Obliczamy wektor ładujący

$$\mathbf{t} = \mathbf{X}\mathbf{w} \quad (52)$$

- Obliczamy wagi \mathbf{Y}

$$\mathbf{c} = \mathbf{Y}'\mathbf{t} / \mathbf{t}'\mathbf{t} \quad (53)$$

- Modyfikujemy wektor ładujący \mathbf{Y}

$$\mathbf{u} = \mathbf{Y}\mathbf{c}/\mathbf{c}'\mathbf{c} \quad (54)$$

Jeśli nie osiągnięto zbieżności, następuje powrót do punktu drugiego. Zbieżność oblicza się najczęściej następująco

$$\|\mathbf{t}_{n-1} - \mathbf{t}_n\|/\|\mathbf{t}_n\| < \varepsilon, \quad \varepsilon \in (10^{-5} - 10^{-8}) \quad (55)$$

- Następnie należy usunąć z macierzy \mathbf{Y} oraz \mathbf{X} obliczone wcześniej składniki w następujący sposób:

$$\mathbf{p} = \mathbf{X}'\mathbf{t}/(\mathbf{t}'\mathbf{t}) \quad (56)$$

$$\mathbf{X} = \mathbf{X} - \mathbf{t}\mathbf{p}' \quad (57)$$

$$\mathbf{Y} = \mathbf{Y} - \mathbf{t}\mathbf{c}' \quad (58)$$

- Kontynuujemy procedurę dla następnego składnika (komponentu) od punktu pierwszego aż do momentu, gdy otrzymamy informację, że w macierzy \mathbf{X} nie ma już danych na temat macierzy \mathbf{Y} . Informację tę można otrzymać, wykorzystując procedurę walidacji skróśnej.

Modelowanie zależności między zmiennymi w PLS opiera się na założeniu, że proces (dane, model) jest zależny od kilku ukrytych w danych zmiennych. Liczba ukrytych zmiennych (ang. *latent variables* – LV) przeważnie nie jest znana, a ich liczba jest estymowana podczas procedury. Liczba ukrytych zmiennych zarówno w modelu PLS, jak i PCA jest nazywana chemicznym rzędem systemu [123]. Ponieważ zarówno zmienne \mathbf{X} oraz \mathbf{Y} są pewnymi realizacjami zmiennych ukrytych, to oznaczają, że nie są one zmiennymi niezależnymi od siebie. Przykładem, który dobrze ilustruje sens zmiennych utajonych, są pomiary spektroskopowe. Jeśli mierzona jest mieszanina składająca się z kilku składników, których widma zachodzą na siebie wzajemnie, to absorbancja na poszczególnych liczbach falowych jest ważoną sumą absorbancji poszczególnych składników. Często mówi się wtedy o liczbie czynników, jakie wpływają na całość absorbancji, a więc macierz danych \mathbf{X} . Jeśli liczba LV jest równa liczbie zmiennych \mathbf{X} , oznacza to równoważność regresji PLS oraz klastycznej regresji wielu zmiennych.

4.3. METODY NIELINIOWE

W spektroskopii nieliniowość w danych objawiać się może jako nieliniowość intensywności piku w zależności od koncentracji, jak też w wyniku przesunięcia pozycji piku linii spektralnej w zależności od koncentracji. Nieliniowa regresja najmniejszych kwadratów rozszerza metodę liniową na większy zbiór funkcji. W regresji nieliniowej między odpowiedzią y jak też predykatorem x występuje model nieliniowy. Regresja nieliniowa wymaga metod optymalizacji do określenia współczynników. Często sto-

sowane są metody nieparametryczne, które lepiej dopasowują się do danych w poszczególnych miejscach (lepsze dopasowanie lokalne). Są to metody umiejscowione lokalnie, niemające właściwości regresji globalnej.

Nieliniowość między wielkościami wejściowymi i wyjściowymi jest częstym zjawiskiem w różnych czujnikach gazu. Przykładowo dla chemorezystancyjnych czujników gazu funkcja przetwarzania może być zapisana jako [124]

$$R_g = sc^m + r \quad (59)$$

gdzie: s , m , r są parametrami modelu wyznaczanymi empirycznie, a c jest koncentracją składnika.

Dla czujników opartych na SnO_2 model zmian konduktywności w zależności od koncentracji wyraża się następującym wzorem [125]:

$$S(c) = a - b \cdot \ln(c + 0.5) \quad (60)$$

gdzie a , b są parametrami modelu wyznaczanymi empirycznie. Modele logarytmiczne z powodzeniem stosowano również do wyznaczania zawartości mieszanin dwu- i trójskładnikowych z wykorzystaniem półprzewodnikowych czujników z tlenków cyny [126].

Dla obydwu podanych przykładów modele są ważne, jeśli nie zachodzą zmiany warunków pracy czujników. Dla macierzy czujników zależność empiryczna sygnału pojedynczego czujnika może być zapisana następująco [127]

$$y_i = a_{i0} + a_{i1}x_1^{k_i} + a_{i2}x_2^{k_i} + a_{i3}x_3^{k_i} + \dots \quad (61)$$

gdzie współczynniki a_{ij} są czułościami danego czujnika na poszczególne składniki x_j , natomiast k_i są wykładnikami przyjmującymi wartości 0,5–0,9 [127]. W zapisie macierzowym otrzymamy

$$y = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & \ddots & & \\ \vdots & & & \\ a_{n1} & \dots & & a_{nm} \end{bmatrix} \begin{bmatrix} x_1^{k_i} \\ x_2^{k_i} \\ \vdots \\ x_m^{k_i} \end{bmatrix} + \begin{bmatrix} a_{10} \\ a_{20} \\ \vdots \\ a_{n0} \end{bmatrix} \quad (62)$$

Jedną z nieliniowych metod wykorzystujących regresję składowych głównych PCR jest wykonanie transformacji nieliniowej wektora pasm \mathbf{t} , metoda ta to wielomianowa regresja PCR [128]:

$$\hat{\mathbf{y}} = b_1 t_1 + b_2 t_2 + \dots + b_{11} t_1^2 + b_{22} t_2^2 + \dots + b_{12} t_1 t_2 + \dots \quad (63)$$

Inną możliwością jest dodanie do oryginalnej macierzy \mathbf{X} elementów tej macierzy odpowiednio nieliniowo przetransformowanych. Jedną z propozycji jest utworzenie macierzy z elementów podniesionych do kwadratu [128]

$$x_{i,j}'_{\text{extra}} = x_{i,j}^2 \quad (64)$$

$$\mathbf{X}_{\text{new}} = [\mathbf{X}_{\text{old}}, \mathbf{X}_{\text{extra}}] \quad (65)$$

Podobnie do metody PCR, przystosować można metodę PLS. Najprościej wykonać to przez przekształcenia zmiennych wejściowych lub/i wyjściowych do budowania modelu. Przekształcenia takie ułatwiają użycie liniowego modelu kalibracji, dzięki któremu rozwiązanie takie jest wydajne obliczeniowo, przy czym rodzaj manipulacji zależy od rodzaju nieliniowości. W wielu przypadkach dane mają nieliniowość, którą można opisać jako miękkie nasycenie (ang. *soft saturation*) [129]. Nieliniowość w takim przypadku pojawia się po przekroczeniu pewnej wartości sygnału. Dla tego typu nieliniowości proponuje się podzielenie intensywności macierzy \mathbf{X} na dwa obszary. Pierwszy z liniowym zakresem oraz drugi, po przekroczeniu zakresu liniowego (wartość progowa k), gdzie z wartości elementów macierzy \mathbf{X} zostaje usunięta wartość progowa. Powstaje w ten sposób macierz, której mają wartości jedynie, gdy przekroczona zostaje wartość progowa $x_{ij} > k$ [129]

$$z_{ij} = (x_{ij} - k) \quad (66)$$

Następnie elementy tej macierzy zostają podniesione do kwadratu

$$\mathbf{Z}^{(2)} = (z_{ij}^2) \quad (67)$$

Rozszerzona macierz do kalibracji przyjmuje postać

$$\mathbf{X}_{\text{extend}} = [\mathbf{X} \quad \mathbf{Z}^{(2)}] \quad (68)$$

Zaletą tego rozwiązania jest mniejsza liczba rozszerzonych zmiennych w stosunku do rozszerzenia, w którym podnoszona do kwadratu jest cała macierz \mathbf{X} . Również model PLS opiera się wtedy na mniejszej liczbie zmiennych utajonych.

Występują również nieliniowe odpowiedniki modelu liniowego PLS, takie jak wielomianowe lub oparte na funkcjach sklepanych modeli PLS [130]. Wykazano jednak w wielu publikacjach, że sieci neuronowe mają lepsze właściwości od nieliniowych metod PLS [131, 132].

Do kalibracji wielowymiarowej w przypadkach nieliniowych stosować można również sieci neuronowe. Podczas uczenia sieci neuronowe mają tendencję do zatrzymywania procesu w najbliższym minimum lokalnym. Dla zbyt małego zbioru uczącego sieć dobrze dopasowuje się do danych uczących, ale dla danych spoza tego zbioru generuje błędy grube [133]. Stąd pomysły takie jak jednoczesne stosowanie wielu modeli predykcji budowanych na podstawie różnych zbiorów uczących (ang. *ensemble model*) [104].

5. KALIBRACJA W WARUNKACH WRAŻLIWOŚCI SYGNAŁÓW POMIAROWYCH NA ZMIANY WIELKOŚCI WPLYWAJĄCYCH

5.1. MODELE KALIBRACJI ODPORNE NA CZYNNIKI WPLYWAJĄCE

Zbudowanie niezawodnego modelu kalibracji w spektroskopii wymaga długiego czasu wykonywania pomiarów, dobrej jakości próbek kalibracyjnych oraz wytworzenia warunków odpowiadającym możliwym czynnikom wpływającym. Jest to procedura związana ze znaczącymi nakładami finansowymi. Stąd też ważne jest, aby nie trzeba jej było wykonywać często. Jeśli już wystąpi zmiana odpowiedzi instrumentalnej – wynikająca z naprawy instrumentu, jego starzenia się, zmiany warunków pomiarów wpływających na przyrząd pomiarowy lub badaną próbkę – ważne, by całej procedury nie trzeba było wykonywać od nowa. Jeśli niezbędne są nowe pomiary korygujące wskazania, to istotne jest to, aby ich liczba była nieznaczna, a przynajmniej mniejsza od liczby pomiarów podstawowych. Jeśli zaszła zmiana parametrów, dodanie kilku nowych punktów kalibracji do punktów wykonanych w poprzednich warunkach może się powieść tylko przy niewielkich zmianach. Większa liczba próbek daje lepsze rezultaty, nie są one jednak w stanie w pełni odwzorować nowych warunków. Nowe punkty kalibracyjne mogą być dodawane do modelu z odpowiednio większymi wagami. Dodatkowo najstarsze punkty kalibracji można stopniowo usuwać, zastępując je nowymi.

W spektroskopowych metodach pomiaru otrzymujemy widmo zawierające cechy spektralne mierzonych składników (analitów), ale również dodatkowe elementy niepożądane, takie jak:

- widmo od znanych składników powodujących interferencje (interferentów),
- interferencje od nieznanymi składników,
- przesunięcie linii podstawowej widma,
- dryft spowodowany wielkościami wpływającymi,
- zniekształcenia spowodowane interakcjami chemicznymi i fizycznymi,
- efekt powodowany przez rozproszenie światła.

W związku z obecnością w widmie niepożądanych elementów stosowane są różnorodne metody im przeciwdziałające, takie jak modele kalibracji niewrażliwe na dryft, metody korekcji lub wstępnego przetwarzania danych niwelujące dryft. Niestety wszystkie te metody wymagają przed ich zastosowaniem znajomości zachodzących zmian. Czynniki nieuwzględnione w kalibracji nie mogą podlegać korekcji. W niektórych przypadkach przy znanych czynnikach wpływających oraz interferencjach zastosowane mogą być metody planowania eksperymentu kalibracji uwzględniające je wszystkie. Jeśli zastosujemy model adaptacyjny i możliwe jest wykonanie referencyjnych pomiarów w takcie pomiarów on-line, to model kalibracji może być rekursywnie adoptowany do danych. Tego typu metody nazywają się niejawną korekcją bezwarun-

kową (ang. *Implicit Correction Method* – ICM) [134, 135]. Metody korekcji ortogonalnej, w których wyszukuje się model ortogonalny w stosunku do zmian wielkości powodujących dryft, nazywane są metodami korekcji jawnej (ang. *Explicit Correction Method*).

W spektroskopii bardzo często zastosowanie mają metody pochodnych [136]. Pochodna pierwszego rzędu umożliwia usunięcie stałej wartości poziomu przesunięcia w widmie, przesunięcie to jest niezależne od długości fali. Pochodna drugiego rzędu ułatwia usunięcie przesunięcia liniowo zmieniającego się z długością fali. Zastosowanie pochodnych zawsze zwiększa rozdzielczość analizy kosztem zmniejszenia stosunku sygnału do szumu [137].

Główny podział metod ograniczania wpływu zmiany warunków środowiskowych lub parametrów instrumentalnych w spektroskopii można przedstawić następująco [138]:

- Metody bez przetwarzania danych lub transferu modeli: dopasowanie instrumentów pomiarowych, modele globalne (uniwersalne), dopasowanie modelu, wybór czujników (zakresu widma).
- Metody standaryzacji: standaryzacja (transfer) współczynników modelu, odpowiedzi spektralnej instrumentu pomiarowego jednowymiarowa lub wielowymiarowa (standaryzacja bezpośrednia, standaryzacja kawałkami gładka), wielkości wyjściowej modelu (nachylenia i składowej stałej odpowiedzi), przekształcenie widma przed standaryzacją (np. transformata falkowa) oraz pozostałe metody (np. sieci neuronowe).

Modele stosowane w przypadku wpływu czynników zewnętrznych muszą być odporne na czynniki wpływające (*robustness, ruggedness*). Odporność taką mierzy się jako zdolność do pozostawania poprawnym i wskazywania prawidłowych wyników pomimo oddziaływań zakłócających. Stworzenie lepszego modelu polega na zmniejszeniu jego czułości na parametry wpływające. Brak odporności modelu manifestuje się w postaci dużych błędów powstających w przypadku małych zmian w funkcji przetwarzania instrumentu pomiarowego oraz/lub warunków pomiaru [139]. Stworzenie odpornego modelu kalibracji można osiągnąć za pomocą następujących metod [140]:

- Tworzenie uniwersalnego modelu przez optymalizację zbioru kalibracyjnego. Musi on zawierać wszystkie możliwe warunki eksperymentalne oraz uwzględnić wszystkie ewentualne czynniki zakłócające. Jeśli zbiór kalibracyjny będzie zawierał te czynniki, to zostaną one niejako wbudowane w model. Tak rozbudowany model kalibracji może niestety przejawiać nieco mniejsze gorsze właściwości. Jest to koszt, jaki ponosi się za uniwersalność.
- Tworzenie wielu modeli lokalnych. Współczynniki poszczególnych modeli dobiera się w zależności od wartości wielkości wpływających.
- Optymalizując wybór cech, najczęściej w postaci doboru czujnika lub wyboru długości fal, które, przenosząc potrzebną informację, są mało wrażliwe na czynniki zakłócające.

- Wstępne przetwarzanie sygnałów, którego zadaniem jest dopasowanie widma do warunków ogólnych. Usunięcie części sygnału odpowiadającego czynnikom zakłócającym.

Model uniwersalny zawierać może jako zmienną wejściową sygnał zakłócający lub też nie. Aby otrzymać model uniwersalny, nazywany również modelem globalnym, należy włączyć podczas kalibracji wszystkie możliwe zmienności parametrów instrumentalnych oraz warunków środowiskowych. Wykonanie takiego modelu może wymagać kontrolowania wielu czynników podczas kalibracji, ale przede wszystkim przewidzenia ich zmienności podczas normalnych pomiarów. Jakość modeli globalnych zależy od możliwości eksperymentatora oraz przewidzenia możliwych zmienności wszystkich parametrów procesu [141]. Ze względu na szeroką funkcjonalność modele uniwersalne mogą być bardzo skomplikowane. Uniwersalność może też nieść ze sobą wadę gorszej dokładności. Za przykład można przytoczyć pomiary zawartości wilgotności w ziemi za pomocą spektroskopii przy zmianach temperatury, użyto w nich pomiar temperatury jako dodatkową zmienną niezależną [142]. Model globalny, w którym nie wykonuje się dodatkowego pomiaru temperatury, zaproponowano do zmierzenia zawartości tłuszczów i protein w pożywieniu [143] oraz do zmierzenia zawartości składników w jabłkach [144].

Metody wyboru zmiennych do zbudowania modelu próbują wyłączyć zakresy, w których występuje zmienność niezwiązana z wyznaczaną wielkością. Następnie model jest budowany na podstawie zmiennych, na które dodatkowe czynniki nie wpływają. Wybór odpowiednich zmiennych jest zadaniem złożonym obliczeniowo [145]. Liczba zmiennych, jakie pozostaną po wyborze, będzie znacznie zredukowana w stosunku do ich wyjściowej. Metody wyboru zmiennych stosowane są raczej w przypadku, gdy nie jest możliwe określenie modelu wpływu czynników dodatkowych na model kalibracji.

Metody wstępnego przetwarzania sygnałów oparte są na dwóch głównych technikach. Pierwsza to znajdowanie określonych czynników w danych, które są ortogonalne w stosunku do mierzonych wielkości. Alternatywnie poszukiwana jest określona przestrzeń danych, w których mieści się oddziaływanie czynników zewnętrznych. Dla wielowymiarowych metod liniowych metody te wypracowują macierze różnic pomiędzy takim samymi składnikami w różnych warunkach. Są to macierze, które reprezentują czynniki zewnętrzne [140]. Najprościej byłoby je obliczyć, wykonując pomiar bez interferencji oraz z interferencjami. Metodę taką zaproponowano do wytwarzania syntetycznych widm, które dodawano następnie do kalibracyjnych widm pomiarowych [146]. Metoda ta zmniejsza liczbę pomiarów kalibracyjnych oraz ułatwia dołączenie do modelu różnorodnych interferencji trudnych do uzyskania jako rozszerzony zbiór kalibracyjny w warunkach pomiarowych.

Ponowne wykonanie kalibracji dla wszystkich punktów kalibracyjnych wiąże się z dużym kosztem. Wykonanie kalibracji niezależnej od możliwych zmian środowiskowych oraz instrumentalnych jest w praktyce bardzo trudne. W pomiarach spektrometrycznych korekcia dryftu może być wykonana metodami transferu kalibracji. Me-

tody te opierają się głównie na użyciu zbioru próbek i dopasowaniu widm dla poprzedniego i nowego instrumentu lub też dla poprzednich i nowych warunków środowiskowych. Inną metodą jest dopasowanie nachylenia oraz offsetu charakterystyki przetwarzania. Jeśli zmianie podlega instrument pomiarowy, zwykle się ten przypadek nazywać standaryzacją instrumentu pomiarowego. Budowanie metod transferu kalibracji powinno w efekcie doprowadzić bądź do izolacji czynnika zakłócającego, bądź do korekcji jego wpływu. Każda zmienność oraz czynnik nieuwzględniony w kalibracji może powodować, że wynik odbiega od prawdziwego.

Metody standaryzacji opierają się na pomiarze pewnego zbioru próbek na dwóch instrumentach pomiarowych, a więc musi zachodzić możliwość takiego pomiaru. W niektórych przypadkach nie jest możliwe zapewnienie tych samych próbek kalibracyjnych w dwóch różnych lokalizacjach. W tego typu przypadkach sprawdzają się metody wstępnego przetwarzania widma.

Transfer kalibracji dla modeli wielowymiarowych jest możliwy dzięki podobieństwu strukturalnym występującym w danych w różnych warunkach [147]. Aby wykonać transfer kalibracji, zazwyczaj niezbędne są dwa kroki. W pierwszym należy wykonać pomiary tych samych próbek na dwóch instrumentach lub dwóch warunkach pomiarowych. W drugim wyznaczyć należy parametry modelu użytego do transferu kalibracji [148]. Podstawowe metody transferu kalibracji to [149]:

- korekcja nachylenia i składowej stałej funkcji kalibracyjnej,
- kawałkami gładka standaryzacja bezpośrednia (ang. *Picewise Direct Standardization* – PDS),
- ortogonalna korekcja sygnałów (ang. *Orthogonal Signal Correction* – OSC),
- globalne metody kalibracji,
- aktualizacja modelu kalibracji,
- zastosowanie filtrów cyfrowych typu FIR.

Transfer kalibracji rozumiany jest najczęściej na dwa sposoby [150]:

- transfer widm, korekcji podlega widmo tak, że widmo drugiego instrumentu jest dopasowane do widma pierwszego,
- transfer modelu, zbudowany na pierwszym instrumencie jest dopasowywany do drugiego instrumentu tak, by wyniki były jednakowe.

Wśród metod transferu widma wymienić można [150]:

- standaryzacja bezpośrednia (ang. *direct standardization*),
- kawałkami gładka standaryzacja bezpośrednia (ang. *piecewise direct standardization*),
- ortogonalna korekcja sygnału (ang. *Orthogonal Signal Correction* – OSC).

Wśród metod transferu modelu wymienić można [150]:

- korekcja nachylenia i poziomu zerowego (ang. *slope and bias correction*),
- dopasowanie modelu (ang. *model updating*),
- transfer równań regresji (ang. *regression equation transfer*).

Najprostsza metoda korekcji zajmuje się wynikiem uzyskanym z modelu dla stałych warunków (ang. *Slope and Bias Correction* – SBC) [151]. Korekcji podlegają nachylenie oraz offset. Zaletą tej metody jest to, że jest ona metodą jednowymiarową. Jest ona także liniowa i sprawdza się najlepiej dla systematycznego wpływu warunków w każdej próbie [152]. Metody kalibracji jednowymiarowej nie są w stanie korygować efektów, takich jak rozszerzanie się poszczególnych linii spektralnych, szczególnie dla pomiarów o wysokiej rozdzielczości. W takich przypadkach lepsze rezultaty uzyskuje się, wykorzystując metody wielowymiarowe.

Najprostszą metodą wielowymiarową jest standaryzacja bezpośrednia (ang. *Direct Standardization* – DS), w której widmo z jednego instrumentu jest przekształcane na widmo innego instrumentu w następujący sposób:

$$\mathbf{X}_1 = \mathbf{X}_2 \mathbf{P} \quad (69)$$

gdzie \mathbf{P} jest macierzą transformacji, natomiast \mathbf{X}_1 jest odpowiednio widmem tego samego zbioru kalibracyjnego próbek w pierwszym oraz \mathbf{X}_2 drugim instrumencie pomiarowym. Przekształcenie to jest oczywiście liniowe, a macierz transformacji może być modelem regresji typu PLS. Model transferu może również zostać wykonany jako standaryzacja pośrednia w pewnej podprzestrzeni, czyli po wstępnym przetworzeniu macierzy \mathbf{X} obydwu instrumentów [153].

Metoda kawałkami gładkiej standaryzacji bezpośredniej (ang. *Picewise Direct Standardization* – PDS) oparta jest na stwierdzeniu, że zmiana w widmie lub mierzonych kanałach odnosi się jedynie do kilku sąsiednich długości fal lub kanałów. Dlatego też macierz przekształcenia \mathbf{P} można zredukować do oddziaływań jedynie na sąsiednie kanały spektralne [154–157].

Metoda kawałkami gładkiej ciągłej standaryzacji bezpośredniej (ang. *Continuous Picewise Direct Standardization* – CPDS) opiera się na modelu PLS budowanym między widmem referencyjnym w określonej (jednej) temperaturze oraz zawartością składników. W kolejnym kroku obliczana jest macierz transferu widma między mierzoną a referencyjną temperaturą. Macierz ta służy do obliczania widma odpowiadającego widmu w temperaturze referencyjnej dla mierzonego w innej temperaturze [158]:

$$\mathbf{X}(T_{\text{ref}}) = \mathbf{X}(T_k) \mathbf{P}(T_k) \quad (70)$$

Elementy macierzy transferu widma obliczane są w następujący sposób [159]:

$$p_{i,j}(T_k) = a_{i,j} + b_{i,j} \Delta T_k + c_{i,j} \Delta T_k^2 \quad (71)$$

Metoda indywidualnej standaryzacji wpływu (ang. *Individual Contribution Standardization* – ICS) została pokazana w celu kompensacji wpływu temperatury na widmo. Jest ona rozszerzeniem metody CPDS, w której uwzględnia się różny wpływ temperatury na poszczególne składniki mieszaniny na różnych długościach fali [160].

Wielomian drugiego stopnia może być również zastosowany w stosunku do standaryzacji wektorów ładujących [161].

Nieco inne podejście polega na włączeniu temperatury w wielowymiarowy model kalibracji [162]. W procesach o skończonym czasie trwania, o charakterystycznym przebiegu zmienności temperatury może ona być wykorzystana jako dodatkowa informacja. Podczas systematycznej zmiany temperatury zmieniają się parametry procesu, a ich wartość w każdym przedziale jej zmian może nieść ze sobą znaczące informacje.

Podział metod przeciwdziałających temperaturowym zmianom widma w pomiarach spektroskopowych może być następujący [163]:

- wyszukiwanie niewrażliwych na temperaturę części widma,
- dopasowanie widma odbywające się najczęściej za pomocą transferowania widma do temperatury podstawowej lub też usuwanie komponentów związanych z jej zmianami,
- modele globalne (nazywane tutaj czystymi metodami regresji).

Metody globalne przeciwdziałające temperaturowym zmianom widma można z kolei podzielić następująco [35]:

- niewymagające pomiaru aktualnej temperatury próbki oparte na odpowiednich metodach planowania eksperymentu,
- z pomiarem aktualnej temperatury próbki, w których temperatura może być zmienną dodatkową dla modelu PLS lub jako zmienna dodatkowa dla modelu regresji liniowej korzystającego z modelu PLS wykorzystującego jedynie widmo.

5.2. ORTOGONALNA KOREKCJA SYGNAŁÓW

Jeśli w wielowymiarowych danych pomiarowych występują wzajemne zależności, to dane zawierają mniej informacji ze względu na ich powtarzanie się. Występowanie zależności oznacza, że dane przedstawić można w nieco innej postaci z mniejszym wymiarem. Zmienne w danych są wzajemnie skorelowane, przez co pojawia się redundancja informacji. Aby liczba nowych zmiennych, do których przekształcone zostaną dane pierwotne, była jak najmniejsza, muszą być one wzajemnie ortogonalne. Wymaganie to spełnia technika analizy składowych głównych (ang. *Principal Component Analysis*). Składowe główne są to nowe zmienne zawierające zmienność występującą w danych oryginalnych. Powstają one jako liniowa kombinacja zmiennych niezależnych. Pierwsza składowa główna (wektor ładujący) zawiera najwięcej zmienności zawartych w danych. Kolejne zawierają coraz mniej informacji. W danych występować mogą zmienności niosące informację oraz zmienności niepożądane mające najczęściej charakter losowy. Dalsze składowe główne najczęściej nie przekazują informacji, lecz zawierają zmienności niepożądane. Ważnym zadaniem jest określenie liczby istotnych składowych głównych i odrzucenie pozostałych. Najczęściej liczba użytecznych składowych głównych jest określana na podstawie analizy wykresów lub wielkości składowych głównych i zawartości występujących w nich zmienności. Anali-

za składowych głównych w postaci wykresów umożliwia przedstawienie pewnych właściwości danych, takich jak charakter zmienności danych oraz ich wizualizację.

Przed zastosowaniem metody PCA dane muszą być poddane operacji centrowania. Podstawowej analizie danych i informacji dostarcza macierz wariancji/korelacji obliczana po autoskalowaniu

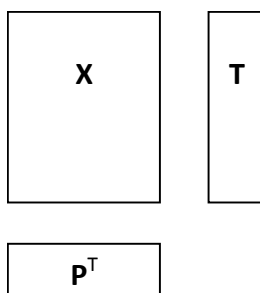
$$\mathbf{Z} = \mathbf{X}^T \mathbf{X} \quad (72)$$

W macierzy tej zawartość głównej przekątnej stanowią wariancje poszczególnych zmiennych. Pozostałe elementy macierzy to korelacja między poszczególnymi zmiennymi. Jeśli w macierzy występują znaczne wartości, oznacza to, że między zmiennymi występują silne korelacje, a, co z tym związane, poszczególne zmienne są wzajemnie skorelowane.

Macierz danych $\mathbf{X}(m \times n)$ zostaje przedstawiona jako iloczyn dwóch macierzy $\mathbf{T}(m \times f)$ – macierz pasm oraz $\mathbf{P}(n \times f)$ – macierz wektorów ładujących (wektorów własnych)

$$\mathbf{X} = \mathbf{TP}^T + \mathbf{E} \quad (73)$$

gdzie: \mathbf{E} – macierz pozostałości, m – liczba obiektów (pomiarów), n – liczba zmiennych niezależnych, f – liczba składowych głównych odpowiadająca liczbie kolumn macierzy \mathbf{T} oraz \mathbf{P} .



Rys. 7. Rozkład macierzy na składowe główne
Fig. 7. PC matrix decomposition

Co można zapisać następująco [164]:

$$\mathbf{X} = \mathbf{t}_1 \mathbf{p}_1^T + \mathbf{t}_2 \mathbf{p}_2^T + \dots + \mathbf{t}_A \mathbf{p}_A^T + \mathbf{E} \quad (74)$$

Liczba składowych głównych określa złożoność modelu PCA, a więc liczbę istotnych składowych zawierających zmienność danych. Jest ona znacząco mniejsza od liczby zmiennych niezależnych. Kolumny macierzy \mathbf{T} oraz \mathbf{P} są ortogonalne. Dodatkowo kolumny \mathbf{P} mają długość jednostkową, co oznacza, że są ortonormalne.

Konstrukcja modeli PCA może być wykonana za pomocą wielu różnych algorytmów. Często są to algorytmy iteracyjne, w których maksymalizacji podlega zmienność danych reprezentowana przez poszczególne składowe główne.

Najczęściej używanym algorytmem obliczania PCA jest metoda rozkładu singularnego (ang. *Singular-Value Decomposition* – SVD), którą można zapisać następująco:

$$\mathbf{X} = \mathbf{USV}^T + \mathbf{E} \quad (75)$$

Poszczególne elementy SVD mają następujące własności:

$$\mathbf{UU}^T = \mathbf{I} \quad (76)$$

$$\mathbf{V}^T \mathbf{V} = \mathbf{I} \quad (77)$$

gdzie \mathbf{S} jest macierzą diagonalną.

W odniesieniu do PCA poszczególne wektory można zapisać następująco:

$$\mathbf{P} = \mathbf{V} \quad (78)$$

$$\mathbf{T} = \mathbf{US} \quad (79)$$

Poszczególne wektory \mathbf{v}_i macierzy \mathbf{V} są wektorami własnymi macierzy $\mathbf{X}^T \mathbf{X}$. Jeśli λ_i jest odpowiednią wartością własną, to rozkład na wektory i wartości własne

$$\mathbf{X}^T \mathbf{X} \mathbf{v} = \lambda \mathbf{v} \quad (80)$$

Wektory własne są wzajemnie ortogonalne. Pierwszy z nich określa kierunek największej zmienności danych, drugi i kolejne odpowiednio następne co do wielkości kierunku zmienności. Ponieważ wektory własne są ortogonalne, to zapisując diagonalną macierz, która będzie zawierała poszczególne wartości własne, otrzymamy

$$\mathbf{X}^T \mathbf{X} \mathbf{V} = \mathbf{\Lambda} \mathbf{V} \quad (81)$$

Co dalej prowadzi do uzyskania następującej zależności

$$\mathbf{S}^2 = \mathbf{\Lambda} \quad (82)$$

Bardzo popularną metodą obliczania składowych głównych jest iteracyjny algorytm (ang. *Nonlinear Iterative Partial Least Squares* – NIPAS):

- wybór początkowy wartości wektora \mathbf{t}_1 , najczęściej w postaci jednej z kolumn macierzy \mathbf{X} ,
- obliczenie $\mathbf{p}_1^T = \mathbf{t}_1^T \mathbf{X}$,
- wykonanie normalizacji wektora \mathbf{p}_1^T ,
- obliczenie $\mathbf{t}_1 = \mathbf{X} \mathbf{p}_1$,

- sprawdzenie różnicy między obliczonym wektorem \mathbf{t}_1^T a jego poprzednią wersją. Jeśli zmiana jest wystarczająco mała, to przyjmujemy obliczony wektor jako wynik. W innym przypadku wracamy do punktu drugiego.

Kolejne składniki główne obliczamy dla macierzy \mathbf{E} powstałej przez usunięcie poprzedniego składnika głównego, co dla pierwszej macierzy zapiszemy jako

$$\mathbf{E}_1 = \mathbf{X} - \mathbf{t}_1 \mathbf{p}_1^T \quad (83)$$

Podczas różnorodnych pomiarów gromadzone są dane, które mogą zawierać wariancje od różnorodnych źródeł. W niektórych przypadkach zmienność od wielkości niezwiązanych może być tak znacząca, że należy ją usuwać przed wykonaniem właściwej interpretacji danych. Te dodatkowe zmienności mogą pochodzić od zmian występujących w instrumencie pomiarowym lub w warunkach pomiarowych. Typowym przykładem są fluktuacje o niskiej częstotliwości spotykane w pomiarach widm różnorodnych procesów. Metoda ortogonalnej korekcji sygnału (ang. *Orthogonal Signal Correction* – OSC) służy do usunięcia wariancji w \mathbf{X} , która jest ortogonalna do wariancji w \mathbf{Y} [165]. Ponadto zależy nam na pozostawieniu całej wariancji niosącej informacje. Metoda OSC jest pewnego rodzaju filtracją, która skupia się na usuwaniu niepożądanych sygnałów w danych, które nie zawierają informacji. Klasycznie rozumiana filtracja usuwa pewne składniki częstotliwościowe sygnału. Tego typu postępowanie nie uwzględnia zależności między zmiennymi wejściowymi oraz wyjściowymi w kalibracji. OSC jest metodą, w której usuwane są składniki niezwiązane ze zmienną wyjściową, a więc ortogonalne do niej. Metoda OSC może również służyć do analizy usuwanych komponentów, które mogą być interpretowane fizycznie jako interferencje. Sprawdziła się także w przypadku, gdy badane składniki mają bardzo podobne widma [137]. Metoda OSC daje dobre rezultaty także w przypadku transferu kalibracji [140]. Algorytmy filtrujące używane podczas wstępnego przetwarzania danych określane są mianem algorytmów izolujących sygnał analityczny [166]. Metoda jest przydatna w monitorowaniu procesów. Jedną z technik monitorowania jest spektroskopia odbiciowa FTIR. Przykładem jest proces krystalizacji, w którym problemem jest korekcja temperatury widma [167]. Powstało wiele odmian OSC, które różnią się głównie metodą obliczania ortogonalnych wektorów, jak też zapewnianiem samej ortogonalności między nimi. Są również metody wykorzystujące OSC wraz z innymi technikami wstępnego przetwarzania sygnałów, takimi jak na przykład transformata falkowa [168]. Metodę OSC najczęściej wykonuje się razem z metodą regresji PLS [169–173]. Elementy metody OSC zostały także użyte do poprawienia możliwości predykcyjnych modeli dla spektroskopii otwartej ścieżki OP-FTIR [174].

Ideę ortogonalizacji można przedstawić, wykorzystując model kalibracji następująco [175]:

$$\hat{\mathbf{y}} = \mathbf{x}^T \mathbf{b} + b_0 \quad (84)$$

gdzie: \hat{y} – wartość predykcji dla modelu odpowiadająca wartości prawdziwej y ,
 \mathbf{x} – mierzone widmo.

Zmiany mierzonego widma $\delta\mathbf{x}$, spowodowane wpływem czynników niepożądanych, można przedstawić następująco:

$$\delta\hat{y} = \delta\mathbf{x}^T \mathbf{b} \quad (85)$$

Błąd bezwzględny wynosi

$$|\delta\hat{y}| = \|\delta\mathbf{x}\| \times \|\mathbf{b}\| \times |\cos(\delta\mathbf{x}, \mathbf{b})| \quad (86)$$

Ortogonalizacja polega na maksymalnym zmniejszeniu członu $\cos(\delta\mathbf{x}, \mathbf{b})$ dla zmian $\delta\mathbf{x}$ wywołanych przez czynniki zakłócające.

Filtr OSC składa się z trzech podstawowych elementów [176], które można zapisać następująco:

$$\{\mathbf{w}_j, \mathbf{t}_j, \mathbf{p}_j\}_{j=1}^A \quad (87)$$

Dla pojedynczego komponentu ortogonalnego macierz \mathbf{X} może być przedstawiona następująco:

$$\mathbf{X} = \mathbf{t}_{\text{OSC}} \mathbf{p}_{\text{OSC}}^T + \mathbf{E} \quad (88)$$

$$\mathbf{t}_{\text{OSC}} = \mathbf{X} \mathbf{w}_{\text{OSC}} \quad (89)$$

$$\|\mathbf{w}_{\text{OSC}}\| = 1 \quad (90)$$

W każdym algorytmie OSC najpierw obliczana jest pierwsza główna składowa oraz wektor ładunków, który następnie podlega ortogonalizacji względem \mathbf{Y} [165]

$$\mathbf{t}^* = (1 - \mathbf{Y}(\mathbf{Y}'\mathbf{Y})^{-1}\mathbf{Y}')\mathbf{t} \quad (91)$$

następnie obliczana jest waga \mathbf{w} zgodnie z algorytmem regresji cząstkowej PLS

$$\mathbf{X}\mathbf{w} = \mathbf{t}^* \quad (92)$$

po czym należy obliczyć nowy wektor ładunków

$$\mathbf{t}_{\text{OSC}} = \mathbf{X}\mathbf{w} \quad (93)$$

oraz sprawdzić czy wektor ładunków osiągnął stabilną wartość. Jeśli nie, to należy ponownie obliczyć wektor \mathbf{t}^* . Jeżeli tak, obliczany jest wektor ładujący

$$\mathbf{p}_{\text{OSC}}^T = \mathbf{t}^T \mathbf{X} / (\mathbf{t}^T \mathbf{t}^*) \quad (94)$$

W tym momencie można już wykonać korekcję z wykorzystaniem pierwszego składnika

$$\mathbf{E} = \mathbf{X} - \mathbf{t}_{\text{OSC}} \mathbf{P}_{\text{OSC}}^T \quad (95)$$

Jeśli potrzebne są następne składniki, procedurę należy powtórzyć określoną liczbę razy dla macierzy \mathbf{E} .

Po wykonaniu pełnej procedury dostępne są macierze \mathbf{W} oraz \mathbf{P} , które stosuje się do korekcji nowych danych. Kolejno obliczane są poszczególne wektory score, następnie wykonywana jest korekcja

$$\mathbf{t}_1 = \mathbf{x}_{\text{data}}^T \mathbf{w}_1 \quad (96)$$

$$\mathbf{e}_1 = \mathbf{x}_{\text{data}}^T - \mathbf{t}_1 \mathbf{p}_1^T \quad (97)$$

Inna metoda obliczenia \mathbf{t} , który to wektor przedstawia zawartość \mathbf{X} ortogonalną do \mathbf{Y} , wykorzystuje bezpośrednio metodę PCA do obliczenia wektora score [177]. Najpierw wektor jest centrowany

$$\mathbf{t} = \mathbf{t} - \bar{\mathbf{t}} \quad (98)$$

a następnie zostaje zortogonalizowany w stosunku do \mathbf{Y}

$$\mathbf{t}^* = (1 - \mathbf{Y}(\mathbf{Y}'\mathbf{Y})^{-1}\mathbf{Y}')\mathbf{t} \quad (100)$$

jak też następuje dodanie jego średniej wartości

$$\mathbf{t}^* = \mathbf{t}^* - \bar{\mathbf{t}} \quad (101)$$

Kolejno obliczany jest wektor \mathbf{w} , który następnie podlega skalowaniu

$$\mathbf{w} = \mathbf{t}^{*T} \mathbf{X} / (\mathbf{t}^{*T} \mathbf{t}^*) \quad (102)$$

$$\mathbf{w} = \mathbf{w} / \sqrt{\mathbf{w}^T \mathbf{w}} \quad (103)$$

powyższe kroki należy wykonywać do momentu ustalenia się wartości \mathbf{t}^* .

Podczas konstrukcji praktycznych algorytmów OSC pojawiają się dwa główne problemy. Są nimi brak pełnej ortogonalności oraz położenie komponentów OSC poza przestrzenią \mathbf{X} [178].

Metoda bezpośredniej ortogonalnej korekcji sygnałów (ang. *Direct - OSC*) powstała z myślą o poprawie właściwości podstawowej metody OSC [179, 178]. W pierwszym kroku obliczane są dwie części, na które dekomponuje się macierz \mathbf{Y} . Są to odpowiednio część, która dokonuje projekcji \mathbf{Y} na \mathbf{X} oraz pozostała część, która jest ortogonalna do \mathbf{X} .

$$\mathbf{Y} = \mathbf{P}_X \mathbf{Y} + \mathbf{A}_X \mathbf{Y} = \hat{\mathbf{Y}} + \mathbf{F} \quad (104)$$

$$\hat{\mathbf{Y}} = \mathbf{P}_X \mathbf{Y} \quad (105)$$

$$\mathbf{F} = \mathbf{A}_X \mathbf{Y} \quad (106)$$

gdzie \mathbf{F} jest częścią ortogonalną do \mathbf{X} , natomiast $\hat{\mathbf{Y}}$ jest projekcją \mathbf{Y} w \mathbf{X} .

Macierz \mathbf{P}_X , która wykonuje projekcje na przestrzeń \mathbf{X} jest obliczana następująco:

$$\mathbf{P}_X = \mathbf{X}(\mathbf{X}\mathbf{X})^{-1} \mathbf{X}^T \quad (107)$$

Macierz \mathbf{A}_X jest projekcją ortogonalną, która stanowi dopełnienie \mathbf{P}_X

$$\mathbf{A}_X = \mathbf{I} - \mathbf{P}_X \quad (108)$$

Następnie dokonuje się rozkładu macierzy \mathbf{X} na dwie ortogonalne części

$$\mathbf{X} = \mathbf{P}_{\hat{\mathbf{Y}}} \mathbf{X} + \mathbf{A}_{\hat{\mathbf{Y}}} \mathbf{X} \quad (109)$$

W typowych danych w spektroskopii liczba analizowanych kanałów spektralnych jest większa od liczby badanych próbek, co sprawia, że macierz \mathbf{X} może być rzutowana bezpośrednio na \mathbf{Y} , zamiast na $\hat{\mathbf{Y}}$ [179]

$$\hat{\mathbf{Y}} = \mathbf{Y} \quad (110)$$

$$\mathbf{X} = \mathbf{P}_Y \mathbf{X} + \mathbf{A}_Y \mathbf{X} \quad (111)$$

Nieco inne rozwiązanie jest zaproponowane w [181], w którym poszukiwane są wagi \mathbf{w} o następujących właściwościach:

$$\max(\mathbf{w}^T \mathbf{X}^T \mathbf{X} \mathbf{w}) \quad (112)$$

po założeniu ortogonalności między usuwaną częścią \mathbf{X} oraz \mathbf{Y}

$$\mathbf{w}^T \mathbf{X}^T \mathbf{Y} = 0 \quad (113)$$

oraz jednostkowej długości wektora \mathbf{w} otrzymujemy

$$\mathbf{w}^T \mathbf{w} = 1 \quad (114)$$

dotatkową macierzą, potrzebną do obliczenia wektora \mathbf{w} , jest następująca:

$$\mathbf{M} = \mathbf{I} - \mathbf{X}^T \mathbf{Y} (\mathbf{Y}^T \mathbf{X} \mathbf{X}^T \mathbf{Y})^{-1} \mathbf{Y}^T \mathbf{X} \quad (115)$$

Wektor \mathbf{w} jest pierwszym wektorem własnym (ang. *eigenvector*) następującej macierzy

$$\mathbf{M} \mathbf{X}^T \mathbf{X} \quad (116)$$

Inną metodą EOSC (ang. *Emphatic* – OSC), oblicza się następująco [182]

$$\mathbf{M} = (\mathbf{Y} - \bar{\mathbf{Y}})^T (\mathbf{X} - \bar{\mathbf{X}}) \quad (117)$$

$$\mathbf{B} = \text{null}(\mathbf{M}) \quad (118)$$

gdzie macierz \mathbf{B} jest jądrem macierzy inaczej zwana przestrzenią zerową.

$$\mathbf{Q} = (\mathbf{X} - \bar{\mathbf{X}})\mathbf{B} \quad (119)$$

W dalszej kolejności dokonuje się dekompozycji singularnej

$$\mathbf{Q} = \mathbf{USV}^T \quad (120)$$

Następnie można już obliczyć macierz korekcji

$$\mathbf{D} = \mathbf{I} - \mathbf{BVS}^{-1}\mathbf{U}^T(\mathbf{X} - \bar{\mathbf{X}}) \quad (121)$$

Korekcję wykonuje się następująco:

$$\mathbf{E} = (\mathbf{X} - \bar{\mathbf{X}})\mathbf{D} \quad (122)$$

Dla nowych danych korekcję przeprowadza się w podobny sposób

$$\mathbf{X}_{\text{OSC-test}} = (\mathbf{X}_{\text{test}} - \bar{\mathbf{X}}_{\text{model}})\mathbf{D}_{\text{model}} \quad (123)$$

Metoda OSC ma również swoje modyfikacje, takie jak ROSC (ang. *Region Orthogonal Signal Correction* – OSC) [183]. Metoda ta polega na użyciu określonego fragmentu widma do obliczenia ortogonalnych komponentów *scores and loading weights*, które następnie stosuje się dla pozostałych części widma. Ważnym aspektem tej metody jest wybór odpowiedniego fragmentu widma.

W metodzie nazwanej REOSC (ang. *Ridge Estimated Orthogonal Signal Correction* – REOSC) modyfikacje stosuje się przez wprowadzenie współczynnika λ , o wartości pomiędzy zerem a jednością, który ogranicza usuwane fragmenty macierzy danych [184]

$$\mathbf{X}_{\text{REOSC}} = \mathbf{X} - \lambda \sum_{i=1}^n \mathbf{t}_{\text{OSC}-i} \mathbf{p}_{\text{OSC}-i}^T \quad (124)$$

Modyfikacje takie jak REOSC są wynikiem pewnych praktycznych analiz, z których wynika, że ortogonalność macierzy dla zbioru uczącego nie musi jednakowo przekładać się do zbiorów testowych oraz dalszych, w których wykorzystywany będzie model OSC. Dodatkowo w pomiarach mogą pojawić się błędy wielkości zależnych, które spowodują brak prawdziwej ortogonalności dla modelu OSC. Jeśli występuje taka sytuacja, to usunięcie komponentu według skonstruowanego modelu może spowodować usunięcie części użytecznej informacji.

Inną metodą modyfikacji jest POSC (ang. *Piecwise Orthogonal Signal Correction* – POSC), który skupia się na usunięciu lokalnych zmienności w danych [185]. Lokalność polega na zastosowaniu OSC osobno dla poszczególnych fragmentów widma. Dodatkowo w algorytmie obliczania wektorów OSC wykonano pewne zmiękczenie

kryterium ortogonalności przez zrezygnowanie z procedury iteracyjnej. Kolejne kroki algorytmu zostaną przedstawione zgodnie z [185]. W pierwszym kroku klasycznie obliczana jest pierwsza główna składowa oraz wektor ładujący, na podstawie którego obliczony zostaje wektor wag

$$\mathbf{w} = \mathbf{p} / (\mathbf{p}^T \mathbf{p}) \quad (125)$$

Wektor score zostaje poddany przekształceniu, dzięki któremu staje się ortogonalny do zmiennej zależnej

$$\mathbf{t}_{\text{osc}} = (1 - \mathbf{Y}(\mathbf{Y}^T \mathbf{Y})^{-1} \mathbf{Y}^T) \mathbf{t} \quad (126)$$

po czym może zostać obliczony wektor ładujący

$$\mathbf{p}_{\text{osc}} = \mathbf{X}^T \mathbf{t}_{\text{osc}} / (\mathbf{t}_{\text{osc}}^T \mathbf{t}_{\text{osc}}) \quad (127)$$

Umożliwia to wyznaczenie macierzy po usunięciu pierwszej składowej ortogonalnej oraz wykonanie przedstawionego obliczenia, jeśli wymagane jest usunięcie kolejnych komponentów ortogonalnych. Mając obliczone macierze wag oraz macierz wartości ładujących, korekcję dla nowych danych przeprowadza się następująco [185]:

$$\mathbf{X}_{\text{OSC-test}} = \mathbf{X}_{\text{test}} - \mathbf{X}_{\text{test}} \mathbf{W} (\mathbf{P}^T \mathbf{W})^{-1} \mathbf{P}^T \quad (128)$$

Podobną do OSC jest metoda DO (ang. *Direct Orthogonalization* – DO), której zadaniem jest pozostawienie informacji zawartych w \mathbf{X} , które są skorelowane z y podczas operacji wstępnego przetwarzania danych [186]. Dzięki usunięciu części danych niezwiązanych z modelowaniem zależności regresji, same modele PLS będą zawierały mniej parametrów. Podstawową różnicą między OSC a DO jest to, że DO nie służy do usuwania niepotrzebnych komponentów z danych przed użyciem ich do predykcji, lecz do budowania lepszych modeli. Model kalibracji nie jest budowany na podstawie oryginalnej macierzy kalibracji \mathbf{X} , lecz z wykorzystaniem macierzy po określonych przekształceniach. Metodę można przedstawić następująco [186]:

- wykonanie centrowania danych kalibracyjnych $\tilde{\mathbf{X}}_c$ oraz $\tilde{\mathbf{y}}_c$,
- wykonanie ortogonalizacji między macierzami

$$\mathbf{w} = \tilde{\mathbf{X}}_c^T \tilde{\mathbf{y}}_c (\tilde{\mathbf{y}}_c^T \tilde{\mathbf{y}}_c)^{-1} \quad (129)$$

$$\hat{\mathbf{X}} = \tilde{\mathbf{X}}_c - \tilde{\mathbf{y}}_c \mathbf{w}^T \quad (130)$$

- wykonanie PCA dla macierzy $\tilde{\mathbf{X}}_c$, obliczenie $\hat{\mathbf{T}}_c \hat{\mathbf{P}}_c$ dla określonej liczby składowych głównych

$$\hat{\mathbf{T}}_c \hat{\mathbf{P}}_c^T \approx \hat{\mathbf{X}}_c \quad (131)$$

- usunięcie danych niezależnych z macierzy kalibracji

$$\tilde{\mathbf{T}}_c = \tilde{\mathbf{X}}_c \hat{\mathbf{P}}_c \quad (132)$$

$$\tilde{\mathbf{X}}_c^{\text{DO}} = \tilde{\mathbf{X}}_c - \tilde{\mathbf{T}}_c \hat{\mathbf{P}}_c^T \quad (133)$$

- obliczenie modelu z wykorzystaniem macierzy $\tilde{\mathbf{X}}_c^{\text{DO}}$ oraz $\tilde{\mathbf{y}}_c$.

Metoda DO znacząco zmniejsza liczbę czynników PLS niezbędnych do uzyskania poprawności na określonym poziomie [187]. Inną podobną metodą jest projekcja ortogonalna (ang. *orthogonal projection*) [188].

6. OPTYCZNA ANALIZA GAZÓW O OTWARTEJ ŚCIEŻCE POMIAROWEJ

6.1. TECHNOLOGIA ANALITYCZNA PROCESÓW

Chemia analityczna procesów – znana również pod nazwą technologia analityczna procesów – wykorzystuje kombinację chemii analitycznej, inżynierię procesów oraz wielowymiarową analizę danych. Umożliwia ona pozyskanie lepszej wiedzy odnośnie do stanu procesu, co ułatwia poprawę efektywności procesu, zmniejszenie kosztów oraz lepsze ich zrozumienie. Jej celem jest stworzenie możliwości ciągłej i nieinwazyjnej analizy on-line procesów w celu kontroli jakości produktów końcowych. Rozwój technologii analitycznej procesów został zapoczątkowany w przemyśle petrochemicznym, a obecnie jest ona stosowana w wielu innych dziedzinach. Podstawową różnicą między analizą procesów a pomiarami laboratoryjnymi jest brak zarówno możliwości kontroli warunków pomiaru, jak też użycia urządzeń przez wyspecjalizowanych operatorów. Instrumenty pomiarowe powinny być w stanie wykonywać wiele czynności samoczynnie, w tym autodiagnostykę, kalibrację oraz walidację. Nie bez znaczenia jest konieczność ciągłej pracy. Ogólnie instrumenty pomiarowe stosowane w pomiarach procesów muszą być solidniejsze pod względem budowy fizycznej. Solidność odnosi się również do części softwareowej, od której wymaga się dostarczania wartościowych informacji na podstawie danych o zmiennych właściwościach.

System służący do analizy procesu składa się z kilku komponentów: podsystemu pobierania próbek, podsystemu analizy próbek, podsystemu przetwarzania i zarządzania danymi. Jeżeli próbki są analizowane poza procesem, to powinny być pobierane w sposób reprezentatywny dla całego procesu bez wprowadzania zauważalnych opóźnień. Pobrana próbka nie może zmienić swoich właściwości, ponieważ nie będzie wtedy umożliwiała otrzymania adekwatnych informacji o procesie. Nie bez znaczenia jest możliwość automatyzacji działania systemu pobierania próbek oraz ewentualna możliwość przeprowadzania automatycznej kalibracji. Kalibracja całego systemu analizy musi zostać wykonana po jego zainstalowaniu oraz okresowo podczas jego pracy.

W niektórych przypadkach wymagane jest dostarczenie odpowiednich standardów. Alternatywną metodą jest analiza porównawcza wykonywana w laboratorium. Okresowa weryfikacja kalibracji musi być przeprowadzana ze względu na możliwe zmiany zachodzące w analizowanych próbkach oraz w samym instrumencie pomiarowym.

Właściwe pomiary wykonywane w technologii PAT dotyczą chemicznych parametrów i właściwości procesów, takich jak skład mieszaniny. Jednym z takich procesów jest proces gazyfikacji [189], w którym zawartość produktów gazowych może być monitorowana za pomocą spektrometru FTIR [190]. Spektroskopia FTIR jest również używana do monitorowania olejów transformatorowych [191, 192]. Sprawdza się także w monitorowaniu reakcji zachodzących podczas produkcji biopaliw [193]. Jest wiele prac dotyczących nowych czujników, które mogą być zastosowane w monitorowaniu procesów [194]. Technologia analityczna procesów PAT to jednak głównie spektroskopia w bliskiej podczerwieni NIR. Zaletą ogólną metod spektroskopowych on-line są niewielkie wymagania odnośnie do przygotowania próbek. W przemyśle farmaceutycznym wykorzystuje się NIR ze względu na jej prostotę, brak inwazyjności, szybki czas analizy, możliwość analizy złożonych substancji [195].

Pomiary spektroskopowe on-line w warunkach przemysłowych narażone są na fluktuacje różnorodnych wielkości: temperatury, turbulencje przepływu, zmiany długości ścieżki pomiarowej. Oddziaływania te mogą być nieliniowe. Jeśli zbudowany został odpowiednio odporny model kalibracji lub stosowane są metody transferu kalibracji, to modele muszą być okresowo sprawdzane. Ponieważ procesy przemysłowe są dosyć złożone, proponowane są grafy kontrolne służące jako podpowiedź, jaką akcję podjąć w danej sytuacji. Aby można było wykonać korekcję modelu, najpierw należy wykryć nieprawidłowość. Jeśli wykryta nieprawidłowość nie może zostać skorygowana automatycznie, możliwe są dwie opcje [196]:

- ponowna kalibracja wymagająca znacznych nakładów eksperymentalnych zbudowania nowego modelu oraz jego walidacji,
- standaryzacja przez korekcję modelu, która może być wykonana za pomocą mniejszej liczby próbek niż pełna ponowna kalibracja.

Samą kalibrację sprawdzać można za pomocą protokołu monitoringu, który powinien być przeprowadzany w czasie rzeczywistym [197].

6.2. OPTYCZNA ANALIZA GAZÓW W OTWARTEJ ŚCIEŻCE POMIAROWEJ

Optyczne metody pomiaru zawartości gazów mają wiele zalet, które ujawniają się szczególnie w przypadku monitorowania środowiska oraz procesów i instalacji przemysłowych. Podstawowe zalety to możliwość pomiaru bez kontaktu z badanym środowiskiem, pomiary mogą być przeprowadzone w pewnej odległości od badanego obiektu, są to również badania nieniszczące. Spektroskopia w podczerwieni polega na pomiarze absorpcji lub emisji promieniowania. Większość składników ma widmo charakterystyczne odróżniające go od innych składników, co ułatwia jego identyfikację.

cję, jak też określenie zawartości ilościowej. Pomiar spektroskopowy są nieniszczące. Dodatkowo mogą to być zarówno pomiary zdalne, jak też pomiary *in situ*. Optyczne metody w podczerwieni umożliwiają bezpośrednią i jednoczesną analizę wielu składników również w przypadku pomiarów przepływów [198].

Otwarta ścieżka detekcji (ang. *open path sensing*) w pomiarach optycznych oznacza, iż promieniowanie docierające do czujnika (spektrometru) nie przebiega w zamkniętej kuwecie. Cechą charakterystyczną jest tutaj pomiar zintegrowanej koncentracji na długości ścieżki pomiarowej. Rozwój tej techniki zapoczątkowany został w monitorowaniu środowiska atmosferycznego. Stąd często jako otwartą ścieżkę rozumie się długą ścieżkę pomiarową niezbędną do wyznaczania śladowej ilości niektórych substancji w atmosferze. Jednakże długie ścieżki pomiarowe mogą być również osiągnięte w zamkniętych kuwetach z wielokrotnymi odbiciami.

Ponieważ analizowane substancje znajdują się w pewnej odległości od elementów systemu pomiarowego, nazywane są także pomiarami zdalnymi (ang. *remote sensing*). Mogą to być również pomiary miejsc niedostępnych za pomocą metod bezpośrednich. Są to więc również pomiary nieinwazyjne, co eliminuje potrzebę posiadania systemu pobierania próbek. Jeśli emisja gazów odbywa się z kilku miejsc lub też ich zawartość jest niejednorodna na dużym obszarze, to pobieranie próbek musiałoby się odbywać w wielu miejscach. Pomiar dużego obszaru z wykorzystaniem metod punktowych nie jest wtedy opłacalny lub wręcz jest niemożliwy.

W konfiguracji aktywnej, w której analizowana jest ścieżka między źródłem promieniowania a detektorem, możliwy jest pojedynczy, podwójny lub wielokrotny przebieg promieniowania przez badane środowisko. Przykładem są optyczne pomiary zdalne wykorzystujące pomiary absorpcyjne za pomocą spektrometrów na długości drogi 8,8 m z wykorzystaniem wielokrotnych odbić, co daje długość drogi pomiaru 70 m [199]. Wyróżnić można dwa rodzaje układów służących do konfiguracji aktywnej. Detektor (spektrometr) może być wyposażony jednocześnie w źródło promieniowania, a po drugiej stronie ścieżki znajdować się musi wtedy lustro. Promieniowanie od źródła do spektrometru przebiega wtedy dwukrotnie przez badaną ścieżkę. Źródło może być również elementem niezależnym i powinno znajdować się po drugiej stronie ścieżki niż spektrometr. Przebieg promieniowania przez środowisko jest wtedy jednokrotny. W niektórych przypadkach za źródło promieniowania użyć można naturalnych źródeł mających wyższą temperaturę niż badana atmosfera. Najczęściej wykorzystywanym źródłem jest wtedy promieniowanie słoneczne, choć mogą to być również nagrzane elementy konstrukcji czy budynków. Mierzone widmo można dla tej sytuacji nazwać pasywnym widmem transmisyjnym. Jeśli tak jak w przypadku gazów wydobywających się z kominów chmura gazowa ma wyższą temperaturę niż otoczenie, to emituje ona określone promieniowanie. Mierzone widmo jest wtedy pasywnym widmem emisyjnym. Pomiary pasywne odbywają się z wykorzystaniem spektrometrów, z których najbardziej popularnymi są spektrometry FTIR. Ich największą zaletą jest uniwersalność pomiaru szerokiej gamy składników.

Optyczne techniki aktywne o otwartej ścieżce pomiarowej mogą wykorzystywać spektroskopię FTIR, ale także źródła laserowe TDLAS i filtry optyczne. Pomiary składu powietrza są często wymagane ze względów prawnych. Substancje toksyczne w powietrzu mogą zwiększać ryzyko wystąpienia określonych chorób. Oczywiście zawartość szkodliwych składników w przypadku normalnej działalności przemysłowej jest ściśle określona, jednakże taka ścisła kontrola nie jest możliwa w przypadku wystąpienia zdarzeń losowych lub wypadków. Jeśli już takie zdarzenie nastąpi, niezbędna jest szybka i efektywna metoda on-line, taka jak [200]: OP-FTIR lub DOAS (ang. *Differential Optical Absorption Spectrometer*). Metoda open-path umożliwia również pomiary gazów cieplarnianych z rozproszonych źródeł emisji [201]. Pomiary otwartej ścieżki, jak wszystkie pomiary spektroskopowe, mogą odnieść korzyści w sytuacji, gdy wykorzystamy metody jednoczesnego użycia wielu długości fali [202].

Dodatkowe możliwości pomiaru w trudnych warunkach pojawiają się, gdy wykorzystane zostaną światłowody jako fragment toru pomiarowego. Zarówno źródło promieniowania, jak też odbiornik może zostać wtedy odsunięty na bezpieczną odległość od niebezpiecznego żrącego środowiska. Taki system z kuwetą, do której doprowadzone jest promieniowanie za pomocą światłowodów, określane jest mianem kuwety o otwartej ścieżce pomiarowej (ang. *open-path gas cell*), a metoda pomiaru spektroskopii absorpcyjnej o zewnętrznej otwartej ścieżce pomiarowej (*extrinsic open-path absorption spectroscopy*) [203]. Przykładowe zastosowanie światłowodów w technice NDIR to pomiary dwutlenku węgla z silników diesla, gdzie możliwe są pomiary gazów do temperatury 100 °C [204]. Technika NDIR stosowana jest również w pomiarach metanu w kopalniach [205, 206]. Zupełnie nowe możliwości pomiarowe pojawiają się dzięki możliwości budowy spektrometrów FTIR w technologii MEMS [207].

Pomiary nieekstrakcyjne, służące do diagnostyki procesów spalania, również mierzą iloczyn długości drogi oraz zawartości. Muszą one wykonywać pomiary w warunkach, jakie panują w przepływających gazach. Cechą charakterystyczną pomiarów różnorodnych gazów wylotowych jest niejednorodność temperatury spotykana na długości drogi pomiarowej. Nawet specjalnie zbudowane kuwety, służące do symulacji warunków panujących w gazach wylotowych z elektrowni konwencjonalnych, wykazują znaczne niejednorodności temperaturowe [208]. Wiele pomiarów tego typu zakłada osiową symetryczność badanej ścieżki pomiarowej [199]. Jest to niezbędne ze względu na dużo mniejszą liczbę danych dostępnych do jej odtworzenia niż w przypadku klasycznej tomografii komputerowej.

7. PODSUMOWANIE I WNIOSKI

Obecnie bardzo popularnymi metodami pomiaru zanieczyszczeń gazowych są metody optyczne oraz systemy czujników typu elektroniczny nos. W warunkach

przemysłowych do monitorowania gazów często wykorzystywana jest spektroskopia laserowa lub spektrometria FTIR. Wnioskowanie ilościowe na podstawie różnorodnych pomiarów w takich warunkach ma wiele cech wspólnych. Przykładem jest zmiana warunków powodujących dryft, której należy odpowiednio przeciwdziałać. Najczęściej wykonuje się to w sposób algorytmiczny. Korekcja dryftu i adaptacja do zmieniających się warunków jest cechą charakterystyczną wielu rozwiązań czujników inteligentnych. Za przykład wymienić można algorytm wielomianów progresywnych.

Dla zmieniających się warunków pomiaru modele odwrotne, szczególnie dla spektroskopii, najczęściej budowane są za pomocą metod regresji PLS lub też z wykorzystaniem sieci neuronowych. Wspólną cechą rozpatrywanych zagadnień jest także zarówno używanie fuzji danych czujników, zespołów modeli ensemble, jak też wielowymiarowej korekcji sygnałów. Metody te mają zastosowanie także w czujnikach wirtualnych, technologii analitycznej procesów oraz w pomiarach optycznych o otwartej ścieżce pomiarowej.

LITERATURA

- [1] CIĘSZCZYK S., *A multi-band integrated virtual calibration- inversion method for open path FTIR spectrometry*, Metrology and Measurement Systems, 2013, Vol. 20, 287–298.
- [2] CIĘSZCZYK S., *A local model and calibration set ensemble strategy for open-path FTIR gas measurement with varying temperature*, Metrology and Measurement Systems, 2013, Vol. 21, 513–524.
- [3] MROCZKA J., SZCZUCZYŃSKI D., *Improved regularized solution of the inverse problem in turbidimetric measurements*, Applied Optics, 2010, Vol. 49, 4591–4603.
- [4] MROCZKA J., SZCZUCZYŃSKI D., *Simulation research on improved regularized solution of inverse problem in spectral extinction measurements*, Applied Optics, 2012, Vol. 51, 1715–1723.
- [5] MROCZKA J., SZCZUCZYŃSKI D., *Improved technique of retrieving particle size distribution form angular scattering measurements*, Journal of Quantitative Spectroscopy & Radiative Transfer, 2013, Vol. 129, 48–59.
- [6] MROCZKA J., *The cognitive process in metrology*, Measurement, 2013, Vol. 46, 2896–2907.
- [7] PALLAS-ARENY R., WEBSTER J.G., *Sensors and Signal Conditioning*, 2nd Edition, John Wiley & Sons, Inc., 2000.
- [8] Van PUTTEN M.J.A., Van PUTTEN M.H.P., *Facing drift: a comparison of three methods*, Sensors and Actuators A, 2001, Vol. 90, 172–180.
- [9] ZAKRZEWSKI J., *Czujniki i przetworniki pomiarowe*, Wydawnictwo Politechniki Śląskiej, Gliwice 2004.
- [10] CUADROS-RODRIGUEZ L., GAMIZ-GRACIA L., ALMANSA-LOPEZ E., BOSQUE-SENDRA J.M., *Calibration in chemical measurement processes: II. A methodological approach*, Trends in analytical chemistry, 2001, Vol. 20, No. 4, 195–206.
- [11] SISK B.C., LEWIS N.S., *Comparison of analytical methods and calibration methods for correction of detector response drift in arrays of carbon black-polymer composite vapour detectors*, Sensors and Actuators B, 2005, Vol. 104, 249–268.
- [12] HOLMBERG M., ARTURSSON T., *Drift Compensation, Standards, and Calibration Methods*, Handbook of Machine Olfaction: Electronic Nose Technology, edited by T.C. Pearce, S.S. Schiffman, H.T. Nage, J.W. Gardner, Wiley-VCH Verlag GmbH, Co., KGaA, Weinheim 2003, 325–346.

- [13] PAUL O., RAZ R., KAUFMANN T., *Analysis of the offset of semiconductor vertical Hall devices*, Sensors and Actuators A, 2012, Vol. 174, 24–32.
- [14] ALIAGHASRGHAMISH M., EBRAHIMI S., *Recursive Least Squares Fuzzy Modeling of Chemoresistive Gas Sensors for Drift Compensation*, International Symposium on Innovations in Intelligent Systems and Applications (INISTA), 2011, 1–5.
- [15] HONG S.K., *Compensation of nonlinear thermal bias drift of Resonant Rate Sensor using fuzzy logic*, Sensors and Actuators, 1999, Vol. 78, 143–148.
- [16] SCOTT S.M., JAMES D., ALI Z., *Data analysis for electronic nose systems*, Microchimica Acta, 2007, Vol. 156, 183–207.
- [17] ROMAIN A.C., NICOLAS J., *Long term stability of metal oxide-based gas sensors for e-nose environmental applications: An overview*, Sensors and Actuators B, 2010, Vol. 146, 502–506.
- [18] PADILLA M., PERERA L., MONTOLIU I., CHAUDRY A., PERSAUD K., MARCO C., *Drift compensation of gas sensor array data by Orthogonal Signal Correction*, Chemometrics and Intelligent Laboratory Systems, 2010, Vol. 100, 28–35.
- [19] VERGARA A., VEMBU S., AYHAN T., RYAN M.A., HOMER M.L., HUERTA R., *Chemical gas sensor drift compensation using classifier ensembles*, Sensor and Actuators B, 2012, Vol. 166–167, 320–329.
- [20] Di CARLO S., FALASCONI M., *Drift Correction Methods for Gas Chemical Sensors in Artificial Olfaction Systems: Techniques and Challenges*, Wen Wang (Ed.), Advances in Chemical Sensors, InTech, 2012, 305–326.
- [21] MARCO S., ORTEGA A., PARDO A., SAMITIER J., *Gas Identification with Tin Oxide Sensor Array and Self-Organizing Maps: Adaptive Correction of Sensor Drifts*, IEEE Transactions on Instrumentation and Measurement, 1998, Vol. 47, No. 1, 316–321.
- [22] BALABAN M.O., KOREL F., ODABASI A.Z., FOLKES G., *Transportability of data between electronic noses: mathematical methods*, Sensors and Actuators B, 2000, Vol. 71, 203–211.
- [23] TOMIC O., EKLOV T., KVAAL K., HAUGEN J.K., *Recalibration of a gas-sensor array system related to sensor replacement*, Analytica Chimica Acta, 2004, Vol. 512, 199–206.
- [24] IRAQI A., MORAWSKI R.Z., BARWICZ A., BOCK W.J., *Static calibration of a thermally self-compensated fiber-optic sensor for high-pressure measurements*, IEEE Instrumentation and Measurement Technology Conference, Ottawa, Canada, 1997, 814–818.
- [25] BEN SLIMA M., MORAWSKI R.Z., KRASZEWSKI A.W., BARWICZ A., NELOSNI S.O., *Calibration of a Microwave System for Measuring Grain Moisture Content*, IEEE Transaction on Instrumentation and Measurement, 1999, Vol. 48, No. 3, 778–783.
- [26] DELPHA C., SIADAT M., LUMBRERAS M., *Discrimination of a refrigerant gas in a humidity controlled atmosphere by using modelling parameters*, Sensors and Actuators B, 2000, Vol. 62, 226–232.
- [27] SOHN J.H., ATZENI M., ZELLER L., PIOGGIA G., *Characterisation of humidity dependence of a metal oxide semiconductor sensor array using partial least squares*, Sensor and Actuators B, 2008, Vol. 131, 230–235.
- [28] ARTURSSON T., EKLOV T., LUNDSTROM I., MARTENSSON P., SJOSTROM M., HOMBERTG M., *Drift correction for gas sensors using multivariate methods*, Journal of Chemometrics, 2000, Vol. 14, 711–723.
- [29] ZIYATDINOV A., MARCO S., CHAUDRY A., PERSAUD K., CAMINAL P., PETERA A., *Drift compensation of gas sensor array data by common principal component analysis*, Sensors and Actuators B, 2010, Vol. 146, 460–465.
- [30] LAGHROUCHE M., IDJERI B., HAMMOUCHE K., TAHANOUT M., BOUSSEY J., AMEUR S., *Temperature compensation of micromachined silicon hot wire sensor using ANN technique*, Microsyst. Technol., 2012, Vol. 18, 237–246.
- [31] KOUDA S., DIBI Z., MEDDOUR F., *Modeling of a Smart Humidity Sensor*, International Conference on Microelectronics ICM, IEEE, 2008, 135–138.

- [32] FUTANE N.P., CHOWDHURY S.R., CHOWDHURY C.R., SAHA H., *ANN based CMOS ASIC design for improved temperature-drift compensation of piezoresistive micro-machined high resolution pressure sensor*, Microelectronics Reliability, 2010, Vol. 50, 182–291.
- [33] PRAMANIK C., ISLAM T., SAHA H., *Temperature compensation of piezoresistive micro-machined porous silico n pressure sensor by ANN*, Microelectronics Reliability, 2006, Vol. 46, 343–351.
- [34] BARBOSA C.H., KUHNER G.S., LIMA E.A., VELLASCO M., PACHECO M., *Application of a MLP Neural Network for Compensation of Distortions in Strain Gauges Measurements Due to Temperature Variations in Offshore Operation of a Flexible Pipe-Lay Vessel*, IEEE, Proceedings of the 2002 International Joint Conference on Neural Networks, 2002, 92–96.
- [35] CHEN T., MARTIN E., *The Impact of Temperature Variations on Spectroscopic Calibration Modelling: A Comparative Study*, Journal of Chemometrics, 2007, Vol. 21, Iss. 5–6, 198–207.
- [36] SINGH A.P., KUMAR S., KAMAL T.S., *Virtual compensator for correcting the disturbing variable effect in transducers*, Sensors and Actuators A, 2004, Vol. 116, 1–9.
- [37] PATRA J.C., ANG E.L., DAS A., CHAUDHARI N.S., *Auto-compensation of nonlinear influence of environmental parameters on the sensor characteristics using neural networks*, ISA Transactions, 2005, Vol. 44, 165–176.
- [38] PATRA J.C., CHAKRABORTY G., MEHER P.K., *Neural Network-based Robust Linearization and Compensation Technique for Sensors under Nonlinear Environmental Influences*, IEEE Transactions on Circuits and Systems I: Regular Papers, 2008, Vol. 55, Iss. 5, 1316–1327.
- [39] BAHA H., DIBI Z., *A Novel Neural Network-Based Technique for Smart Gas Sensors Operating in a Dynamic Environment*, Sensors, 2009, Vol. 9, 8944–8960.
- [40] HAJMIRZAHEYDARALI M., GHAFARINA V., *A Smart Gas Sensor Insensitive to Humidity and Temperature Variations*, IOP Conf. Series: Materials Science and Engineering, 2011, Vol. 17, 1–8.
- [41] ISLAM T., PRAMANIK C., SAHA H., *Modeling, simulation and temperature compensation of porous polysiloxon capacitive humidity sensor using ANN technique*, Microelectronics Reliability, 2005, Vol. 45, 697–703.
- [42] ZUPPA M., DISTANTE C., SICILIANO P., PERSAUD K.C., *Drift counteraction with multiple self-organising maps for an electronic nose*, Sensors and Actuators B, 2004, Vol. 98, 305–317.
- [43] HOLMBERG M., WINQUIST F., LUNDSTROM I., DAVIDE F., DiNATALE C., D'AMICO A., *Drift counteraction for an electronic nose*, Sensors and Actuators B, 1996, Vol. 35–36, 528–535.
- [44] WENZEL M.J., MENSAH-BROWN A., JOSSE F., YAS E., *Online Drift Compensation for Chemical Sensors Using Estimation Theory*, IEEE Sensors Journal, 2011, Vol. 11, No. 1, 225–232.
- [45] LO Y.L., CHU C.S., YUR J.P., CANG Y.C., *Temperature compensation of fluorescence intensity-based fiber-optic oxygen sensors using modified Stern-Volmer model*, Sensors and Actuators B, 2008, Vol. 131, 479–488.
- [46] MISEJE M., STRUCEL J., *Signal Processing in Smart Sensor Systems 1*, Journal of Electrical Engineering, 2003, Vol. 54, 154–159.
- [47] MISEJE M., STRUCEL J., *Signal Processing in Smart Sensor Systems 2*, AT & P Journal, Intelligent Sensoric Systems, 2006, Vol. 2, 7–12.
- [48] SINGH A.P., KAMAL T.S., KUMAR S., *Development of a virtual curve tracer for estimation of transducer characteristics under the influence of a disturbing variable*, Sensors and Actuators A, 2005, Vol. 120, 518–526.
- [49] SINGH V.R., *Smart sensors: Physics, technology and applications*, Indian Journal of Pure & Applied Physics, 2005, Vol. 43, 7–16.
- [50] YAMASAKI H., *What are the Intelligent Sensors*, Handbook of Sensors and Actuators, 1996, Vol. 3, 1–18.
- [51] RIVERA J., HERRERA G., CHACON M., ACOSTA P., CARRILLO M., *Improved Progressive Polynomial Algorithm for Self-Adjustment and Optimal Response in Intelligent Sensors*, Sensors, 2008, Vol. 8, 7410–7427.

- [52] GAURA E. NEWMAN R.M., *Smart, intelligent and cogent MEMS based sensors*, Intelligent Control, Proceedings of the 2004 IEEE International Symposium, 2004, 432–436.
- [53] MORÓN Z., *Can analog circuits still increase sensors' smartness?*, Sensors and Actuators A, 1998, Vol. 68, 474–479.
- [54] PASIC A., DOWLING J., *Linearising Calibration Methods for a Generic Embedded Sensor Interface (GESI)*, 1st International Conference on Sensing Technology, Palmerston North New Zealand, 2005, 185–190.
- [55] ERDEM H., *Implementation of software-based sensor linearization algorithms on low-cost micro-controllers*, ISA Transactions, 2010, Vol. 49, 552–558.
- [56] SINGH A.P., KAMAL T.S., KUMAR S., *Development of a virtual linearizer for correcting transducer static nonlinearity*, ISA Transactions, 2006, Vol. 45, No. 3, 319–328.
- [57] KLUK P., MISIURSKI G., MORAWSKI R.Z., *Total Least Squares versus RBF Neural Networks in Static Calibration of Transducers*, IEEE Instrumentation and Measurement Technology Conference, Ottawa, Canada, 1997, 424–426.
- [58] KHAN S.A., SHAHANI D.T., AGARWALA A.K., *Sensor calibration and compensation using artificial neural network*, ISA Transaction, 2003, Vol. 42, 337–352.
- [59] WANG X., YE M., *Hysteresis and nonlinearity compensation of relative humidity sensor using support vector machines*, Sensors and Actuators B, 2008, Vol. 129, 274–284.
- [60] SARMA U., CHAKRABORTY D., BORUAH P.K., *Design of a Smart and High Precision Industrial Temperature Measurement and Monitoring System Using K-type Thermocouple and SPI-compatible Temperature Sensor*, Sensors & Transducers Journal, 2009, Vol. 102, Iss. 3, 1–9.
- [61] DANISMAN K., DALKIRAN I., CELEBI F.V., *Design of a high precision temperature measurement system based on artificial neural network for different thermocouple types*, Measurement, 2006, Vol. 39, 695–700.
- [62] KOUIDER M., NADI M., KOURTICHE D., *Sensors Auto-calibration Method – Using Programmable Interface Circuit Front-end*, Sensors, 2003, Vol. 10, 491–497.
- [63] RIVERA-MEJIA J., CARRILLO-ROMERO M., HERRERA-RUIZ G., *Quantitative Evaluation of Self Compensation Algorithms Applied in Intelligent Sensors*, Instrumentation and Measurement Technology Conference (I2MTC), IEEE, 2010, 1116–1120.
- [64] LYAHOU K.F., Van der HORN G., HUIJSING J.H., *A Non-Iterative, Polynomial, 2-Dimensional Calibration Method Implemented in a Microcontroller*, IEEE Instrumentation and Measurement Technology Conference, 1996, 62–67.
- [65] Van der HORN G., HUIJSING, *Integrated Smart Sensors Design and Calibration*, Norwell, MA, Kluwer, 1998.
- [66] DIAS PEREIRA J.M., POSTOLACHE O., GIRAO S., *PDF-Based Progressive Polynomial Calibration Method for Smart Sensors Linearization*, IEEE Transactions On Instrumentation And Measurement, 2009, Vol. 58, No. 9, 3245–3552.
- [67] KJELDAHL K., BRO R., *Some common misunderstandings in chemometrics*, Journal of Chemometrics, 2010, Vol. 24, 558–564.
- [68] RAHILI S., GHASARI J., GOLFAR A., *Intelligent Selection of Calibration Points Using a Modified Progressive Polynomial Method*, IEEE Transaction On Instrumentation and Measurement, 2012, 2519–2523.
- [69] GUERRERO E., SANZ-PASCUAL M.T., MEDRANO N., CALVO B., *A Digitally Programmable Calibration Circuit for Smart Sensors*, Third Latin American Symposium on Circuits and Systems (LASCAS), IEEE, 2012, 1–4.
- [70] POLAK A.G., MROCZKA J., *Pośrednie pomiary właściwości obiektów złożonych*, [w:] Problemy metrologii elektronicznej i fotonicznej, J. Mroczka (red.), Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 2008, 15–78.

- [71] BAKSHI B.R., UTOJO U., *A common framework for the unification of neural, chemometric and statistical modeling methods*, Analytica Chimica Acta, 1999, Vol. 384, 227–247.
- [72] BOKOV V., *Multivariate multiplicative composite modelling techniques*, Measurement, 2006, Vol. 39, 153–168.
- [73] KADLEC P., GABRYS B., *Soft Sensor Based on Adaptive Local Learning*, Advances in Neuro-Information Processing, 2009, Vol. 5506, 1172–1179.
- [74] LIN B., JORGENSEN S., *Soft sensor design by multivariate fusion of image features and process measurements*, Journal of Process Control, 2011, Vol. 21, 547–553.
- [75] SLISKOVIC D., GRABIC R., HOCENSKI Z., *Methods for Plant Data-Based Process Modeling in Soft-Sensor Development*, Automatica, 2011, Vol. 52 (4), 306–318.
- [76] KADLEC P., GABRYS B., *Local learning-based adaptive soft sensor for catalyst activation prediction*, AIChE Journal, 2011, Vol. 57, 1288–1301.
- [77] KADLEC P., GRBIC R., GABRYS B., *Review of adaptation mechanisms for data-drive soft sensors*, Computers and Chemical Engineering, 2011, Vol. 35, 1–24.
- [78] RUHM K., *Sensor fusion and data fusion – Mapping and reconstruction*, Measurement, 2007, Vol. 40, 145–157.
- [79] CIĘSZCZYK S., *Sensors signal processing under influence of environmental disturbances*, Przegląd Elektrotechniczny, 2013, Vol. 89, 129–131.
- [80] SAPONARA S., PETRI E., FANUCCI L., TERRENI P., *SENSOR MODELING, Low-Complexity Fusion Algorithms, and Mixed-Signal IC Prototyping for Gas Measures in Low-Emission Vehicles*, IEEE Transactions on Instrumentation and Measurement, 2011, Vol. 60, No. 2, 372–384.
- [81] WECKENMANN A., JIANG X., SOMMER K., NEUSCHAEFER-RUBE U., SEEWING J., SHAW L., ESTLER T., *Multisensor data fusion in dimensional metrology*, CIRP Annals – Manufacturing Technology, 2009, Vol. 58, 701–721.
- [82] KHALEGHI B., KHAMIS A., KARARY F., RAZAVI S., *Multisensor data fusion: A review of the state-of-art*, Information Fusion, 2012, Vol. 14, 28–44.
- [83] WEIMAR U., GOPEL W., *Chemical imaging: II. Trends in practical multiparameter sensor systems*, Sensors and Actuators B, 1998, Vol. 52, 143–161.
- [84] LUO R.C., KAY M.G., *Multisensor Integration and Fusion in Intelligent Systems*, IEEE Transactions on systems, man, and cybernetics, 1989, Vol. 19, No. 5, 901–931.
- [85] ELMENREICH W., *An introduction to sensor fusion*, Research Report, 47/2001, 2002.
- [86] STEINER G., *Model-based Data Fusion in Industrial Process Instrumentation*, in Sensor and Data Fusion, edited by: Nada Milisavljevi, I-Tech, Vienna, Austria, 2009, 201–224.
- [87] LEE P.S., DEXTER A.L., *A fuzzy sensor for measuring the mixed air temperature in air-handling units*, Measurement, 2005, Vol. 37, 83–93.
- [88] MOHAMMAD-DJAFARI A., GIOVANNELLI J.F., DEMOMENT G., IDIER J., *Regularization, maximum entropy and probabilistic methods in mass spectrometry data processing problems*, International Journal of Mass Spectrometry, 2002, Vol. 215, 175–193.
- [89] LIU Y., BROWN S.D., *Wavelet multiscale regression from the perspective of data fusion: new conceptual approaches*, Anal. Bioanal. Chem., 2004, Vol. 380, 445–452.
- [90] GAO L., REN S., *Multivariate calibration of spectrophotometric data using a partial least squares with data fusion*, Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 2010, Vol. 76, 363–368.
- [91] BROWN C.D., VEGA-MONTOTO L., WENTZELL P.D., *Derivative Preprocessing and Optimal Corrections for Baseline Drift in Multivariate Calibration*, Applied Spectroscopy, 2000, Vol. 54, No. 7, 1055–1068.
- [92] TAN C., WANG J., QIN X., LI M., *Ensemble multivariate calibration based on mutual information for food analysis using near-infrared spectroscopy*, Analytical Letters, 2010, Vol. 43, 2640–2651.
- [93] FRAYMAN Y., ROLFE B.F., WEBB G.I., *Solving Regression Problems Using Competitive Ensemble Models*, in Advances in artificial intelligence: proceedings of the 15th Australian Joint Confer-

- ence on Artificial Intelligence, R.I. McKay, Slaney J. (eds.), Springer-Verlag, Berlin, Heidelberg, 2002, 511–522.
- [94] PARTALAS I., TSOUMAKAS G., HATZIKOS E.V., VLAHAVAS I., *Greedy regression ensemble selection: Theory and an application to water quality prediction*, Information Sciences, 2008, Vol. 178, 3867–3879.
- [95] RE M., VALENTINI G., *Ensemble methods: a review*, [In:] *Advances in Machine learning and Data Mining for Astronomy*, Ed: Way M.J. , Scargle J.D., Field M., Ali K.M., Srivastava A.N., Chapman & Hall Data Mining and Knowledge Discovery Series, 2012, Chap. 26, 563–594.
- [96] BREIMAN L., *Bagging Predictors*, Machine Learning, 1996, Vol. 24, 123–140.
- [97] OPITZ D., MACLIN R., *Popular Ensemble Methods: An Empirical Study*, Journal of Artificial Intelligence Research, 1999, Vol. 11, 169–198.
- [98] KOTSIANTIS S., KANELLOPOULOS D., *Combining bagging, boosting and random subspace ensembles for regression problems*, International Journal of Innovative Computing, Information and Control, 2012, Vol. 8, No. 6, 3953–3961.
- [99] PARDO M., SBERVEGLIERI G., *Learning From Data: A Tutorial With Emphasis on Modern Pattern Recognition Methods*, IEEE Sensors Journal, 2002, Vol. 2, No. 3, 203–216.
- [100] DIETTERICH T.G., *Ensemble Methods in Machine Learning*, Lecture Notes in Computer Science, 2000, Vol. 1857, 1–15.
- [101] KOTSIANTIS S., PINTELAS P., *Combining Bagging and Boosting*, International Journal of Computational Intelligence, 2004, Vol. 1, No. 4, 324–333.
- [102] DIETTERICH T.G., *An Experimental Comparison of Three Methods for Constructing Ensembles of Decision Trees: Bagging, Boosting, and Randomization*, Machine Learning, 2000, Vol. 40, 139–157.
- [103] DAQI G., WEI C., *Function Approximation Model Ensembles for the Simultaneous Determinations of Odor Classes and Concentrations*, Proceedings of the 2006 IEEE International Conference on Networking, Sensing and Control, IEEE, 2006, 428–433.
- [104] UKIL A., BERNASCONI J., BRAENDLE H., BUIJS H., BONENFANT S., *Improved Calibration of Near-Infrared Spectra by Using Ensembles of Neural Network Models*, IEEE Sensors Journal, 2010, Vol. 10, No. 3, 578–584.
- [105] ZHENQIANG S., WEIDA T., LEMING S., XUEGUANG S., WENSHENG C., *A Partial Least Squares-Based Consensus Regression Method for the Analysis of Near-Infrared Complex Spectral Data of Plant Samples*, Analytical Letters, 2006, Vol. 39, 2073–2006.
- [106] TAN C., QIN X., LI M., *Subspace Regression Ensemble Method Based on Variable Clustering for Near-Infrared Spectroscopic Calibration*, Analytical Letters, 2009, Vol. 42, 1693–1710.
- [107] TAN C., LI M., QIN X., *Random Subspace Regression Ensemble for Near-Infrared Spectroscopic Calibration of Tobacco Samples*, Analytical Sciences, 2008, Vol. 24, 647–653.
- [108] CONLIN A.K., MARTIN E.B., MORRIS A.J., *Data augmentation: an alternative approach to the analysis of spectroscopic data*, Chemometrics and Intelligent Laboratory Systems, 1998, Vol. 4, 161–173.
- [109] SAIZ-ABAJO M.J., MEVIK B.H., SEGTMAN V.H., NAES T., *Ensemble methods and data augmentation by noise addition applied to the analysis of spectroscopic data*, Analytica Chimica Acta, 2005, Vol. 533, 147–159.
- [110] MEVIK B.H., SEGTMAN V.H., NAES T., *Ensemble Methods and Partial Least Squares Regression*, Journal of Chemometrics, 2004, Vol. 18, 498–507.
- [111] ŚWITAJ-ZAWADKA A., KONIECZKA P., PRZYK E., NAMIEŚNIK J., *Calibration in Metrological Approach*, Analytical Letters, 2005, Vol. 38, 535–376.
- [112] *International Vocabulary of Basis and General Terms in Metrology (VIM)*, ISO, Geneva 1993.
- [113] CUADROS-RODRIGUEZ L., GAMIZ-GRACIA L., ALMANSA-LOPEZ E., *Calibration in chemical measurement processes: I. A metrological approach*, Trends in Analytical Chemistry, 2001, Vol. 20, No. 11, 620–636.

- [114] BARWICZ A., *A system approach to electrical measurements*, Measurement, 1996, Vol. 19, No. 3/4, 131–138.
- [115] TRUCANO T.G., SWILER L.P., IGUSA T., OBERKAMPF W.L., PILCH M., *Calibration, validation, and sensitivity analysis: What's what*, Reliability Engineering and System Safety, 2006, Vol. 91, 1331–1357.
- [116] SOUNDBERG R., *Multivariate Calibration – Direct and Indirect Regression Methodology*, Scandinavian Journal of Statistics, 1999, Vol. 26, 161–207.
- [117] CIOSEK P., WRÓBLEWSKI W., *Potentiometric Electronic Tongues for Foodstuff Biosample Recognition-An Overview*, Sensors, 2011, Vol. 11, 4688–4701.
- [118] BRO R., *Multivariate calibration. What is chemometrics for the analytical chemist?*, Analytica Chimica Acta, 2003, Vol. 500, 185–194.
- [119] ESCANDAR G., DAMIANI P., GOICOECHEA H., OLIVIERI A., *A review of multivariate calibration method applied to biomedical analysis*, Microchemical Journal, 2006, Vol. 82, 29–42.
- [120] BRERETON R.G., *Introduction to multivariate calibration in analytical chemistry*, Analyst, 2000, Vol. 125, 2125–2154.
- [121] RAJALAHTI T., KVALHEIM O.M., *Multivariate data analysis in pharmaceuticals: A tutorial review*, International Journal of Pharmaceutics, 2011, Vol. 417, 280–290.
- [122] WOLD S., SJOSTROM M., ERIKSSON L., *PLS-regression: a basic tool of chemometrics*, Chemometrics and Intelligent Laboratory System, 2001, Vol. 58, 109–130.
- [123] SOUNDBERG R., *Aspects of statistical regression in sensometrics*, Food Quality and Preference, 2000, Vol. 11, 17–26.
- [124] HOSSEIN-BABAEI F., GHAFARINIA V., *Compensation for the drift-like terms caused by environmental fluctuations in the responses of chemoresistive gas sensors*, Sensors and Actuators B, 2010, Vol. 143, 641–648.
- [125] CHAIYBOUN A., TRAUTE R., HASS T., KIESEWETTER O., DOLL T., *A logarithmic multi-parameter model using gas sensor main and cross sensitivities to estimate gas concentrations in a gas mixture for SnO₂ gas sensors*, Sensors and Actuators B, 2007, Vol. 123, 1064–1070.
- [126] HIROBAYASHI S., KADIR M.A., YOSHIZAWA T., YAMABUCHI T., *Verification of a logarithmic model for estimation of gas concentration in a mixture for a tin oxide gas sensor response*, Sensors and Actuators B, 2003, Vol. 92, 269–278.
- [127] GARDNER J.W., BARLETT P.N., *Performance definition and standardization of electronic nose*, Sensors and Actuators B, 1996, Vol. 33, 60–67.
- [128] VERDU-ANDRES J., MASSART D.L., MENARDO C., STERNA C., *Correction of non-linearities in spectroscopic multivariate calibration by using transformed original variables. Part II. Application to principal component regression*, Analytica Chimica Acta, 1999, Vol. 389, 115–130.
- [129] ROBERTSSON G., *Contribution to the problem of approximation of non-linear data with linear PLS in an absorption spectroscopic context*, Chemometrics and Intelligent Laboratory Systems, 1999, Vol. 47, 99–106.
- [130] BALABIN R.M., SAFIEVA R.Z., LOMAKINA E.I., *Comparison of linear and nonlinear calibration models based on near infrared (NIR) spectroscopy data for gasoline properties prediction*, Chemometrics and Intelligent Laboratory Systems, 2007, Vol. 88, 183–188.
- [131] YANG H., GRIFFITHS P.R., TATE J.D., *Comparison of partial least squares regression and multi-layer neural networks for quantification of nonlinear systems and application to gas phase Fourier transform infrared spectra*, Analytica Chimica Acta, 2003, Vol. 489, 125–136.
- [132] HADJIISKI L., GELADI P., HOPKE P., *A comparison of modelling nonlinear systems with artificial neural networks and partial least squares*, Chemometrics and Intelligent Laboratory Systems, 1999, Vol. 49, 91–103.

- [133] MAKOWSKI P., ROJ J., *Dobór funkcji aktywacji sieci neuronowej realizującej odtwarzanie wielkości wejściowej przetwornika pomiarowego*, *Pomiary Automatyka Robotyka*, 2004, Vol. 7–8, 120–124.
- [134] GUJRAL P., AMRHEIN M., BONVIN D., *Drift correction in multivariate calibration models using on-line reference measurements*, *Analytica Chimica Acta*, 2009, Vol. 642, 27–36.
- [135] GUJRAL P., AMRHEIN M., BONVIN D., *Measurement-Based Drift Correction in Spectroscopic Calibration Models*, 11th Conference on Chemometrics in Analytical Chemistry, Montpellier, France, 2008, Vol. 1, 117–121.
- [136] BROWN C.D., VEGA-MONTOTO L., WENTZELL P.D., *Derivative Preprocessing and Optimal Corrections for Baseline Drift in Multivariate Calibration*, *Applied Spectroscopy*, 2000, Vol. 54, No. 7, 1055–1068.
- [137] BERTAN E., ITURRIAGA H., MASPOCH S., MONTOLIU I., *Effect of orthogonal signal correction on the determination of compounds with very similar near infrared spectra*, *Analytica Chimica Acta*, 2001, Vol. 431, 303–311.
- [138] FEUDALE R.N., WOODY N.A., TAN H., MYLES A.J., BROWN S.D., FERRE J., *Transfer of multivariate calibration models: a review*, *Chemometrics and Intelligent Laboratory Systems*, 2002, Vol. 64, 181–192.
- [139] GEMPERLINE P.J., *Rugged spectroscopic calibration for process control*, *Chemometrics and Intelligent Laboratory Systems*, 1997, Vol. 39, 29–40.
- [140] IGNE B., ROGER J., ROUSSEL S., BELLON-MAUREL V., HURBURGH C., *Improving the transfer of near infrared prediction models by orthogonal methods*, *Chemometrics and Intelligent Laboratory Systems*, 2009, Vol. 99, 57–65.
- [141] CHEN Z., LOVETT D., MORRIS J., *Process analytical technologies and real time process control a review of some spectroscopic issues and challenges*, *Journal of Process Control*, 2011, Vol. 21, 1467–1482.
- [142] LIANG X.Y., LI X.Y., LEI T.W., WANG W., GAO Y., *Study of sample temperature compensation in the measurement of soil moisture content*, *Measurement*, 2011, Vol. 44, 2200–2204.
- [143] ZHANG X., CHANG M., XING L., HU J., *Influence and correction of temperature on optical measurement for fat and protein contents in a complex food model system*, *Infrared Physics & Technology*, 2010, Vol. 53, 177–181.
- [144] PEIRS A., SCHEERLINCK N., NICOLAI B.M., *Temperature compensation for near infrared reflectance measurement of apple fruit soluble solids contents*, *Postharvest Biology and Technology*, 2003, Vol. 30, 233–248.
- [145] SWIERENGA H., WULFERT F., De NOORD O.E., De WEIJER A.P., SMILDE A.K., BUYDENS L.M.C., *Development of robust calibration models in near infra-red spectrometric applications*, *Analytica Chimica Acta*, 2000, Vol. 411, 121–135.
- [146] HAALAND D.M., *Synthetic Multivariate Models to Accommodate Unmolded Interfering Spectral Components during Quantitative Spectral Analyses*, *Applied Spectroscopy*, 2000, Vol. 54, Iss. 2, 246–254.
- [147] GUJRAL P., AMRHEIN M., WISE B.M., BONVIN D., *Framework for explicit drift correction in multivariate calibration models*, *Journal of Chemometrics*, 2010, Vol. 24, 534–543.
- [148] XIE Y., HOPKE P.K., *Calibration transfer as a data reconstruction problem*, *Analytica Chimica Acta*, 1999, Vol. 384, 193–205.
- [149] NI W., BROWN S., MAN R., *Data fusion in multivariate calibration transfer*, *Analytica Chimica Acta*, 2010, Vol. 661, 133–142.
- [150] ABDELKADER M.F., COOPER J.B., LARKIN C.M., *Calibration transfer of partial least squares jet fuel property models using a segmented virtual standards slope-bias correction method*, *Chemometrics and Intelligent Laboratory Systems*, 2012, Vol. 110, 64–73.

- [151] BOUVERESSE E., HARTMANN C., MASSART D.L., LAST I.R., PREBBLE K.A., *Standardization of near-infrared spectrometric instruments*, Anal. Chem., 1996, Vol. 68, 982–990.
- [152] DU W., CHEN Z., ZHONG L., WANG S., YU R., NORDON A., LITTLEJOHN D., HOLDEN M., *Maintaining the predictive abilities of multivariate calibration models by spectral space transformation*, Analytica Chimica Acta, 2011, Vol. 690, 64–70.
- [153] PARK K.S., KO Y.H., LEE H., JUN C.H., CHUNG H., KU M.S., *Near-infrared spectral data transfer using independent standardization samples: a case study on the trans-alkylation process*, Chemometrics and Intelligent Laboratory Systems, 2001, Vol. 55, 53–65.
- [154] WANG Y., KOWALSKI B.R., *Calibration Transfer and Measurement Stability of Near-Infrared Spectrometers*, Appl. Spectroscopy, 1992, Vol. 46, 764–771.
- [155] GRIFFITHS M.L., SVOZIL D., WORSFOLD P., EVANS E.H., *The application of piecewise direct standardization with variable selection to the correction of drift in inductively coupled atomic emission spectrometry*, Journal of Analytical Atomic Spectrometry, 2006, Vol. 21, 1045–1052.
- [156] ALAM T.M., ALAM M.K., NCLNTRE S.K., VOLK D.E., NEERATHILANGAM M., LUXON B.A., *Investigation of chemometric instrumental transfer methods for high-resolution NMR*, Anal. Chem., 2009, Vol. 81, 4433–4443.
- [157] FAN W., LIANG Y., YUAN D., WANG J., *Calibration model transfer for near-infrared spectra based on canonical correlation analysis*, Analytica Chimica Acta, 2008, Vol. 623, 22–29.
- [158] WANG Y., VELTKAMP D.J., KOWALSKI B.R., *Multivariate instrument standardization*, Anal. Chem., 1991, Vol. 63, 2750–2756.
- [159] WULFERT F., KOK W.T., De NOORD O.E., SMILDE A.K., *Correction of temperature-induced spectral variation by continuous piecewise direct standardization*, Anal. Chem., 2000, Vol. 72, 1639–1644.
- [160] CHEN Z.P., MORRIS J., MARTIN E., *Modeling temperature-induced spectral variations in chemical process monitoring*, Proc. of the Seventh IFAC Symposium DYCOPS-7, Cambridge, Massachusetts, USA, 2004.
- [161] CHEN Z.P., MORRIS J., MARTIN E., *Correction of Temperature-Induced Spectral Variations by Loading Space Standardization*, Anal. Chem., 2005, Vol. 77, 1376–1384.
- [162] PEINADO A.C., Van den BERG F., BLANCO M., BRO R., *Temperature-induced variation for NIR tensor-based calibration*, Chemometrics and Intelligent Laboratory Systems, 2006, Vol. 83, 75–82.
- [163] HAGEMAN J.A., WESTERHUIS J.A., SMILDE A.K., *Temperature robust multivariate calibration: an overview of methods for dealing with temperature influences on near infrared spectra*, J. Near Infrared Spectroscopy, 2005, Vol. 13, 53–62.
- [164] GELADI P., *Chemometrics in spectroscopy*. Part 1. *Classical chemometrics*, Spectrochimica Acta, Part B, 2003, Vol. 58, 767–782.
- [165] WOLD S., ANTTI H., LINDGREN F., OHMAN J., *Orthogonal signal correction of near-infrared spectra*, Chemometrics and Intelligent Laboratory Systems, 1998, Vol. 144, 175–185.
- [166] BLANCO M., CASTILLO M., PEINADO A., BENEYTO R., *Determination of low analyte concentrations by near-infrared spectroscopy: Effect of spectral pretreatment and estimation of multivariate detection limits*, Analytica Chimica Acta, 2007, Vol. 581, 318–323.
- [167] POLLANEN K., HAKKINEN A., REINIKAINEN S., LOUHI-KULTANEN M., NYSTROM L., *ATR-FTIR in monitoring of crystallization processes: comparison of indirect and direct OSC methods*, Chemometrics and intelligent Laboratory Systems, 2009, Vol. 76, 25–35.
- [168] ERIKSSON L., TRYGG J., JOHANSSON E., BRO R., WOLD S., *Orthogonal signal correction, wavelet analysis, and multivariate calibration of complicated process fluorescence data*, 2000, Vol. 420, 181–195.
- [169] NIAZI A., GOODARZI M., *Orthogonal signal correction – partial least squares method for simultaneous spectrophotometric determination of cypermethrin and tetramethrin*, Spectrochimica Acta Part A, 2008, Vol. 69, 1165–1169.

- [170] GHASEMI J., NIAZI A., *Spectrophotometric simultaneous determination of nitroaniline isomers by orthogonal signal correction – partial least squares*, Talanta, 2005, Vol. 65, 1168–1173, 2005.
- [171] NIAZI A., AZIZI A., RAMEZANI M., *Simultaneous spectrophotometric determination of mercury and palladium with Thio-Michler's Ketone using partial least squares regression and orthogonal signal correction*, Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy A, 2008, Vol. 71, 1172–1177.
- [172] NIAZI A., YAZDANIPOUR A., *Spectrophotometric simultaneous determination of nitrophenol isomers by orthogonal signal correction and partial least squares*. Journal of Hazardous Materials, 2007, Vol. 146, 421–427.
- [173] BLANCO M., CASTILLO M., PEINADO A., BENEYTO R., *Determination of low analyte concentrations by near-infrared spectroscopy: Effect of spectral pretreatment and estimation of multivariate detection limits*, Analytica Chimica Acta, 2007, Vol. 581, 318–323.
- [174] ZHANG L., ZHANG L., YAN L. WANG J., *Orthogonal Signal Correction Used for Noise Elimination of Open Path Fourier Transform Infrared Spectra*, Journal of Environmental Science and Health, Part A, 2005, Vol. 40, 1069–1079.
- [175] PREYS S., ROGER J.M., BOULET J.C., *Robust calibration using orthogonal projection and experimental design. Application to the correction of the light scattering effect on turbid NIR spectra*, Chemometrics and Intelligent Laboratory Systems, 2008, Vol. 91, 28–33.
- [176] LI B., MORRIS J., MARTIN E., *Orthogonal signal correction: algorithmic aspects and properties*, Journal of Chemometrics, 2002, Vol. 16, 556–561.
- [177] SJOBLÖM J., SVENSSON O., JOSEFSON M., KULLBERG H., WOLD S., *An evaluation of orthogonal signal correction applied to calibration transfer of near infrared spectra*, Chemometrics and Intelligent Laboratory Systems, 1998, Vol. 44, 229–244.
- [178] ZHU D., JI B., MENG C., SHI B., TU Z., QING Z., *The application of direct orthogonal signal correction for linear and non-linear multivariate calibration*, Chemometrics and Intelligent Laboratory Systems, 2008, Vol. 90, 108–115.
- [179] WESTERHUIS J.A., De JONG S., SMILDE A.K., *Direct orthogonal signal correction*, Chemometrics and Intelligent Laboratory Systems, 2001, Vol. 56, 13–25.
- [180] LUYPART J., HEUERDING S., De JONG S., MASSART D.L., *An evaluation of direct orthogonal signal correction and other preprocessing method for the classification of clinical study lots of a dermatological cream*, Journal of Pharmaceutical and Biomedical Analysis, 2002, Vol. 30, 453–466.
- [181] FEARN T., *On orthogonal signal correction*, Chemometrics and Intelligent Laboratory Systems, 2000, Vol. 50, 47–52.
- [182] ZHANG J., ZHANG Z., XIANG Y., DAI Y., HARRINGTON P., *An emphatic orthogonal signal correction-support vector machine method for the classification of tissue section of endometrial carcinoma by near infrared spectroscopy*, Talanta, 2011, Vol. 83, 1401–1409.
- [183] KASEMSUMRAN S., PING DU L., MARUO K., OZAKI Y., *Selective removal of interference signals for near-infrared spectra of biomedical samples by using region orthogonal signal correction*, Analytica Chimica Acta, 2004, Vol. 526, 193–202.
- [184] SHEN Q., JIANG J., SHEN G., YU R., *Ridge estimated orthogonal signal correction for data preprocessing prior to PLS modelling: QSAR studies of cyclooxygenase-2 inhibitors*, Chemometrics and Intelligent Laboratory Systems, 2006, Vol. 82, 44–49.
- [185] FEUDALE R.N., TAN H., BROWN S.D., *Improved Piecewise Orthogonal Signal Correction Algorithm*, Applied Spectroscopy, 2003, Vol. 57, No. 10, 1201–1206.
- [186] ANDERSSON C.A., *Direct orthogonalization*, Chemometrics and Intelligent Laboratory Systems, 1999, Vol. 47, 51–63.

- [187] PIERNA J.A.F., MASSART D.L., De NOORD O.E., RICOUX Ph., *Direct Orthogonalization: some case studies*, Chemometrics and Intelligent Laboratory Systems, 2001, Vol. 55, 101–108.
- [188] PEREZ-MARTIN D., FEARN T., GUERRERO J.E., GARRIDO-VARO A., *Robustness in pig fat NIRs calibrations by orthogonal projection*, Chemometrics and Intelligent Laboratory Systems, 2010, Vol. 100, 36–40.
- [189] BAGGIO P., BARATIERI M., FIORI L., GRIGIANTE M., AVI D., TOSI P., *Experimental and modeling analysis of a batch gasification/pyrolysis reactor*, Energy Conversion and Management, 2009, Vol. 50, 1426–1435.
- [190] GRANADA E., EQUIA P., VILAN J.A., COMESANA J.A., COMESANA R., *FTIR quantitative analysis technique for gases. Application in a biomass thermochemical processes*, Renewable Energy, 2012, Vol. 41, 416–421.
- [191] GEORGIEV A., KARAMANCHEVA I., TOPALOVA L., *Evaluation of the antioxidation activity of ionol and piperidone towards transformer oil using FT-IR spectroscopy*, Journal of Molecular Structure, 2006, Vol. 787, 25–33.
- [192] ZAKHARICH M.P., ZAITSEV I.I., KOMAR V.P., NIKONOVICH F.N., RYZHKOV M.P., SKORNYAKOV I.V., *Analysis of transformer oil using IR analyzers*, Journal of Applied Spectroscopy, 2001, Vol. 68, No. 1, 61–65.
- [193] KILLNER M.H.M., ROHWEDDER J.J.R., PASQUINI C., *A PLS regression model using NIR spectroscopy for on-line monitoring of the biodiesel production reaction*, Fuel, 2011, Vol. 90, 3268–3273.
- [194] SHUK P., BAILEY E., ZOSEL J., GUTH U., *New advanced in situ carbon monoxide sensor for the process application*, Ionics, 2009, Vol. 15, 131–138.
- [195] WU Y., JIN Y., LI Y., SUN D., LIU X., CHEN Y., *NIR spectroscopy as a process analytical technology (PAT) tool for on-line and real-time monitoring of an extraction process*, Vibrational Spectroscopy, 2012, Vol. 58, 109–118.
- [196] LARRECHI M.S., CALLAO M.P., *Strategy for introducing NIR spectroscopy and multivariate calibration techniques in industry*, Trends in Analytical Chemistry, 2003, Vol. 22, No. 10, 634–640.
- [197] COGDILL R.P., ANDERSON C.A., DRENNEN J.K., *Process Analytical Technology Case Study, Part III: Calibration Monitoring and Transfer*, AAPS PharmSciTech, 2005, Vol. 6, E284-E297.
- [198] GALLIGNANI M., Del ROSARIO BRUNETTO M., *Infrared detection in flow analysis – developments and trends (review)*, Talanta, 2004, Vol. 64, 1127–1146.
- [199] BAUM R. T., MCGRATTAN K.B., NYDEN M.R., *An Examination of the Applicability of Computed Tomography for the Measurement of Component Concentration in Fire-Generated Plumes*, Combustion and Flame, 1998, 113, 358–372.
- [200] SHIE R.H., CHAN C.C., *Tracking hazardous air pollutants from a refinery fire by applying on-line and off-line air monitoring and back trajectory modeling*, Journal of Hazardous Materials, 2013, 261, 72–82.
- [201] RO K.S., JOHNSON M.H., VARMA R.M., HASHMONAY R.A., HUNT P., *Measurement of greenhouse gas emissions from agricultural sites using open-path optical remote sensing method*, Journal of Environmental Science and Health Part A, 2009, Vol. 44, 1011–1018.
- [202] LIUDCHIK A.M., *Multiwavelength technique for open-path measurements of the concentrations of atmospheric constituents*, Journal of Applied Spectroscopy, 2012, Vol. 79, 477–483.
- [203] MUDA R., LEWIS E., O'KEEFFE S., DOOLY G., CLIFFORD J., *A Compact Optical Fibre Based Mid-Infrared Sensor System for Detection of High Level Carbon Dioxide Emissions in Exhaust Automotive Applications*, Procedia Chemistry, 2009, Vol. 1, 593–596.
- [204] MULROONEY J., CLIFFORD J., FITZPATRICK C., LEWIS E., *Detection of carbon dioxide emissions from a diesel engine using a mid-infrared optical fiber based sensor*, Sensors and Actuators A, 2007, Vol. 136, 104–110.

- [205] LIANG Y., RONG X. LI B., LIU C., *Study on Intelligence Detection of Methane Gas Based on NDIR Technology*, Deng et al. (eds.): Emerging Research in Artificial Intelligence and Computational Intelligence AICI, 2011, 454–460.
- [206] CHENG W., HAN J., WU Y., *Design of a Methane Concentration Detector Based on Spectrum Absorption*, 2nd International Symposium on Computer, Communication, Control and Automation, 2013, 423–426.
- [207] KRAFT M., KENDA A., SANDNER T., SCHENK H., *MEMS-based Compact FT-Spectrometers – A Platform for Spectroscopic Mid-Infrared Sensors*, Sensors, 2008 IEEE, 2008, 130–133.
- [208] SPELMAN J., SKRIEN S., PARKER T.E., *Power plant exhaust gas simulator for the development of optical diagnostics*, Review of Scientific Instruments, 2001, 72 (9), 3699–3705.

SOME PROBLEMS OF GASES CONCENTRATION MEASUREMENT IN CHANGING CONDITIONS

The paper presents the problem of gases measurement, in which there is a change of environmental conditions. Interference and changing influence factors appear. The problem is shown in terms of smart sensors, virtual sensors and data fusion. The methods of building inverse models for gas concentration determination are presented. These models have the properties of robustness to the influence of the measurement conditions. This occurs especially in open path measurement, when uncontrolled temperature changes are often present. Next, multivariable calibration, calibration transfer and spectrum correction for spectroscopic measurement are shown.



Wydawnictwa Politechniki Wrocławskiej
są do nabycia w księgarni
plac Grunwaldzki 13, 50-377 Wrocław
budynek D-1 PWr., tel. 71 320 29 35
Prowadzimy sprzedaż wysyłkową
zamawianie.ksiazek@pwr.edu.pl

ISBN 978-83-7493-867-9