

INSTYTUT ŁĄCZNOŚCI
WARSZAWA-MIEDZESZYN

BIULETYN
Instytut Łączności

BIULETYN

INFORMACYJNY

11-12 (285-286)

1990

A-WB

BIULETYN INFORMACYJNY

BIBLIOTEKA
Instytutu Łączności

ROK 30

WARSZAWA 1990

NR 11-12(285-286)

INSTYTUT ŁĄCZNOŚCI
Brzozowy Ośrodek
Informacji Naukowej, Technicznej i Ekonomicznej

Redakcja Biuletynu Informacyjnego

Redaktor Naczelny - dr inż. Krystyn Plewko

Z-ca Redaktora Naczelnego - doc. dr inż. Stanisław Sońta

Redaktorzy działów:

doc. dr inż. Alina Karwowska-Lamparska

mgr inż. Mirosław Żurawski

Adres Redakcji:

Instytut Łączności

Branżowy Ośrodek

Informacji Naukowej, Technicznej i Ekonomicznej

Warszawa-Miedzeszyn, ul. Szachowa 1

ISSN 0209-1048

Redaktor: mgr Krystyna Juskiewicz

Montaż tekstu: Barbara Skwara

Dział Ogólnotechniczny Instytutu Łączności
Format B5. Nakład 625. Wpłynęło do Działu Ogólnotechnicznego
1990.08.16. Druk ukończono w styczniu 1991 r.

NIEPEWNOŚCI WYNIKÓW POMIARU

(Wolny przekład z języków angielskiego i niemieckiego dokonany przez prof. dr Jerzego Dudziewicza (pkt. 3, 4, 5, 6) i inż. Leszka Chodakowskiego (pkt. 2) z Instytutu Łączności w Warszawie)

SPIS TREŚCI

	Str.
1. Słowo wstępne	1
2. Wyznaczanie niepewności wyników pomiarów eksperymentalnych	4
2.1. Wprowadzenie	4
2.2. Sprawy ogólne	5
2.3. Spojrzenie na niektóre główne problemy	7
2.4. Pojęcia maksymalnej granicy	11
2.5. Sytuacja obecna	13
2.6. Uwagi końcowe	19
2.7. Dodatki	21
Bibliografia	24
3. Rozważania o niepewności - pytania i odpowiedzi	26
3.1. Uwagi wstępne	26
3.2. Pleliminaria	27
3.3. Istota zalecenia BIPM	30
3.4. Sytuacja obecna i przyszła	35
3.5. Uwagi końcowe	38
Bibliografia	40
4. Niepewność pomiaru - osobisty rzut oka na znaczenie tego pojęcia w przeszłości i przyszłości	42
4.1. Uwagi wstępne	42
4.2. Spojrzenia tradycyjne	44
4.3. Niejasna sytuacja	47
4.4. Pierwsze promyki nadziei	49
4.5. Rola "Code of Practice"	50

	Str.
4.6. Czy rzeczywiście potrzebne są klasy błędów?	52
4.7. Na drodze do zalecenia BIPM	54
4.8. Obecna sytuacja i pewne problemy aktualne	57
4.9. Rzut oka w przyszłość	61
Bibliografia	62
Dodatek: Ocena niepewności	65
1. Klasyfikacja różnych niepewności	66
2. Prawo propagacji niepewności	70
Bibliografia	72
5. Zalecenie 1 (CI-1986) CIPM - wyrażanie niepewności w pracach wykonywanych pod auspicjami CIPM	73
6. Zalecenie 2 (CI-1986) - analiza wyników komparacji międzynarodowych	74

NIEPEWNOŚCI WYNIKÓW POMIARU¹⁾

1. SŁOWO WSTĘPNE

Materiał niniejszego Biuletynu Informacyjnego dotyczy zagadnienia metod oszacowania i wyrażania niepewności wyniku pomiarów eksperymentalnych. Ta tematyka nawiązuje ściśle do materiału zawartego w Biuletynie Informacyjnym nr 1-2 (223 - - 224) z roku 1984, w którym zamieściliśmy teksty tłumaczone z dokumentów BIPM pod tytułem: "Określenia i oceny niepewności pomiaru w świetle praktyki wielkich ośrodków metrologicznych". Od czasu wydania oryginałów tych dokumentów upłynęło już około 10 lat i w tym okresie ukazało się wiele prac, które stanowią cenne komentarze do nowych tendencji dotyczących klasyfikacji błędów pomiaru. Uznaliśmy, że do najciekawszych z nich należą referaty i artykuły głównego współtwórcy tych nowych trendów, a mianowicie materiały opracowane przez naukowca z BIPM, doktora Jorga W. Müllera. Wszystkie te trzy prace (por. pkt. 2, 3 i 4 poniżej) zostały opublikowane w trzech kolejnych tomach (vol. 9, 10 i 11) przeglądu prac BIPM²⁾ pt. "Recueil de Travaux du BIPM", które ukazały się odpowiednio w latach 1983-1984, 1985-1986 i 1987-1988. Oddając do rąk polskiego czytelnika niniejsze materiały tłumacze sądzą, że ułatwią w ten sposób przyswojenie nowego sposobu traktowania niepewności pomiarów eksperymentalnych przez liczne rzesze praktyków i teoretyków związanych z techniką pomiarową i z ogólnymi podstawami metrologii.

1) Wolny przekład z języków angielskiego i niemieckiego dokonany przez prof. dr Jerzego Dudziewicza i inż. Leszka Chodakowskiego z Instytutu Łączności w Warszawie.

2) BIPM - Bureau International des Poids et Mesures (Międzynarodowe Biuro Miar).

Warto dodać, że te trzy materiały zawierają podobne treści i mogą wywołać obawy o zbytłą redundancję. Jednak zdecydowano się świadomie na ten eksperyment i zamieszczono w tym Biuletynie wszystkie trzy artykuły, uważając, że każdy z nich przedstawia tę problematykę z innego punktu widzenia. I tak pierwszy z nich (pkt. 2) zawiera opis istoty nowego sposobu wyznaczania niepewności wyników pomiarów, obala mit o rzekomej logice dotychczasowego podziału błędów oraz przypomina i właściwie interpretuje prawo propagacji błędów. W drugim (pkt. 3) przytoczono interesującą polemikę między "przekonanym" i "nieprzekonanym" do nowego sposobu oceny wyniku pomiaru, a więc ciekawą rozprawę z typowymi zarzutami wobec nowej idei; materiał ten ma duże wartości dydaktyczne i daje dużo argumentów dla propagatorów nowego podejścia do tej problematyki. Wreszcie trzeci z nich (pkt. 4) odzwierciedla własną drogę przejścia od "starego" sposobu pojmowania tych zagadnień do "nowego" sposobu, przytaczając wiele trudności i wahań, które napotkał sam autor w swojej karierze metrologa i z którymi rzetelnie się uporał. Wszystko to przemawiało za tym, aby przekazać całość materiału polskim metrologom w celu zachęcenia ich do obrania nowej drogi przy ocenie niepewności pomiarów, co stanowi dla nich fundamentalne zadanie.

Materiały niniejsze wreszcie uzupełniono dwoma zaleceniami CIPM¹⁾, które zostały przyjęte w październiku 1986 roku na 75 sesji CIPM w Paryżu (por. pkt. 5 i 6 poniżej). Należy oczekiwać, że zakończeniem okresu wdrożenia do praktyki pomiarowej tych nowych zaleceń będzie opracowanie i publikowanie wytycznych dla ich praktycznego stosowania, jakie mają się ukazać po zakończeniu prac Grupy Roboczej Międzynarodowej Organizacji Normalizacyjnej (ISO), która wraz

1) CIPM - Comité International des Poids et Mesures (Międzynarodowy Komitet Miar).

z OIML¹⁾, IEC²⁾ oraz CIPM podjęła się takiej pracy. Jest rzeczą oczywistą, że upłynie jeszcze wiele czasu, zanim te nowe metody zostaną przyswojone i wdrożone do codziennej praktyki. Tłumacze sądzą, że niniejszym opracowaniem przyczynią się również w pewnej mierze do przyswojenia tych nowych tendencji przez polskich fachowców.

-
- 1) OIML - Organisation Internationale de Métrologie Legale (Międzynarodowa Organizacja Metrologii Prawnej).
 - 2) IEC - International Electrotechnical Commission (Międzynarodowa Komisja Elektrotechniczna).

2. WYZNACZANIE NIEPEWNOŚCI WYNIKÓW POMIARÓW EKSPERYMENTALNYCH¹⁾

"Należy pamiętać, że pewne błędy bywają dla jednego przypadkowe, a dla drugiego systematyczne". P.Vigoureux

2.1. Wprowadzenie

Konferencja o podstawowych stałych fizycznych nie jest właściwym miejscem, gdzie ktoś mógłby oczekiwać, że będą tłumaczył obszernie, dlaczego prezentacja wyników doświadczeń powinna zawsze zawierać informację o domniemanej precyzji lub dokładności. W istocie z naukowego punktu widzenia liczbowy wynik pomiaru bez jakiegokolwiek wskazania odnośnie dokładności jest prawie bezużyteczny, ponieważ nie pozwala na jego sensowne porównanie z innym wyznaczonym wynikiem tej samej wielkości albo z teorią.

W artykule - po krótkim przeglądzie niektórych aktualnych sposobów określanie niepewności pomiarów eksperymentalnych w sposób ilościowy - omówione są główne niedostatki pewnych zwyczajowych praktyk. Obecna sytuacja jest niezadawalająca i objawia się często mylnymi interpretacjami. Tradycyjna praktyka rozróżniania niepewności "przypadkowych" i "systematycznych" jest jedną z głównych przyczyn bałaganu z wyznaczaniem niepewności. Rzeczywiście, ten podział jest problematyczny, gdyż może być zmieniany przez

¹⁾ Oryginał angielski tego artykułu J. W. Müllera pod tytułem "The Assignment of Uncertainties to the Results of Experimental Measurements" został opublikowany po raz pierwszy w "Precision Measurement and Fundamental Constants II", B.N.Taylor and W.D. Philips, Eds., Natl. Bur.Stand.(U.S.), Spec. Publ. 617, 1984, a po raz drugi w "Recueil de Travaux du BIPM", vol. 9. 1983 - 1984, BIPM, Sèvres, 1985.

kolejnych użytkowników "niepewności". Podział taki, jeżeli jest w ogóle potrzebny, powinien zależeć tylko od sposobu uzyskiwania liczbowej wartości niepewności. Co się tyczy wpływu na wynik końcowy, to decyduje właściwie zastosowanie dobrze znanego prawa propagacji błędów, dla którego potrzebne są tylko dane wejściowe w celu uzyskania wariancji i kowariancji mierzonych wielkości.

2.2. Sprawy ogólne

Sprawy wyznaczania niedokładności podlegają od wielu dziesięcioleci niekorzystnym wpływom w związku z problemami, z których pewne są trudne do uniknięcia, ale szczególnie w związku z innymi, które właśnie wynikają z braku zgody na to, co chcemy opisać i na sposób, w jaki chcemy to zrobić. Główną trudnością jest to, że chociaż wielu eksperymentatorów stosuje w praktyce przydzielanie niepewności mierzonym wartościom, to jednak często dla czytelnika dalekie jest od oczywistości, co ma dokładnie oznaczać ta liczba. To stanowi poważne źródło kłopotów dla krytycznie oceniających dane, którzy powinni odnosić różne wskazania do wspólnej podstawy, na przykład w celu przydzielania właściwych wag statystycznych danym, które porównuje on między sobą lub z pewną teorią. Wadliwa interpretacja wyznaczania błędów zdarza się dość często (i często jest nie do uniknięcia wobec braku informacji) i może prowadzić w wyniku do zalecania wątpliwych "najlepszych wartości".

Spróbujmy najpierw zrozumieć, czym są te główne problemy i skąd się one biorą, a gdy już to zrozumiemy, powinniśmy zasugerować ewentualny sposób pokonania tych trudności, przynajmniej dla najczęściej występujących przypadków. Warto wspomnieć od razu na początku, że w takiej dziedzinie jak ta, w której tak wiele słów zostało już wypowiedzianych, jest prawie niemożliwe powiedzieć coś,

co by było rzeczywiście nowego; być może nie jest to nawet pożądane, ponieważ nie brakuje rad i to o najlepszych intencjach. To czego potrzebujemy, polega raczej na wydzielaniu sygnału z szumu, kierując się w miarę możliwości takimi niezawodnymi zasadami jak zdrowy rozsądek, prostota, spójność i użyteczność, uważając przy tym, aby nie naruszać podstawowych praw fizyki i statystyki. Można mieć nadzieję, że postępując w ten sposób, możemy wyraźniej ujrzeć główny problem, a następnie wskazać drogi wyjścia z tej sytuacji, która wydaje się obecnie nie do rozwikłania. Literatura na temat błędów jest ogromna, ale zwykle niezbyt pomocna. Niektóre z tych interesujących opinii można znaleźć w sprawozdaniach pierwszej konferencji [2] PMFC (Precision Measurement and Fundamental Constants - Precyzyjne Pomiary i Podstawowe Stałe Fizyczne).

Nasze zadanie jest utrudnione z powodu wielu okoliczności, o których można by myśleć, że są nieistotne, ponieważ przynależą raczej do psychologii niż do fizyki, ale niestety tak nie jest. Po pierwsze, powinniśmy przeciwstawiać się temu, że każdy uważa się za eksperta. Po drugie, nie lubimy zmieniać opinii i zwyczajów, szczególnie jeżeli są one dostępne w drukowanej formie i są liczne. Wreszcie istnieje całkiem pokaźna liczba recept puszczonej w obieg przez grupę ekspertów, z których większość jest całkiem sprzecznych i nieuzasadnionych, niestety częstokroć popieranych przez wpływowe organizacje krajowe, a nawet międzynarodowe. Wszystkie te recepty łączy to, że uwzględniają wąski i uproszczony punkt widzenia, ignorują podobne problemy spotykane w innych dziedzinach i pozwalają oszacować niepewności z zadziwiającą precyzją (co najmniej trzy cyfry znaczące) nawet dla bardzo kiepskich danych. Ich świat jest niezmiernie gaussowski i ciągły. Pomija się przeważnie fakt, że oszacowana niepewność także ma swoją niepewność.

W tym stanie rzeczy śmiałą inicjatywą dr A. Amblera, członka Międzynarodowego Komitetu Miar, była sugestia

polegająca na tym, aby BIPM postarał się wyjaśnić sytuację i jeśli to możliwe - z pomocą krajowych laboratoriów - dojsć do użytecznych i ogólnie akceptowalnych reguł określania niepewności eksperymentalnych. W celu wysondowania opinii krajowych laboratoriów w tej kontrowersyjnej sprawie niepewności został opracowany i rozesłany w lutym 1978 r. do narodowych laboratoriów odpowiedni kwestionariusz. Analizę odpowiedzi przedstawiono w dokumencie [3].

2.3. Spojrzenie na niektóre główne problemy

Chciałbym wspomnieć o niektórych dyskutowanych problemach, zmieniły się bowiem, w wyniku dyskusji i lepszego rozumienia, szczegóły poruszane w odpowiedziach laboratoriów tak, że nie odpowiadają one już obecnemu punktowi widzenia. Stąd dokument [3], chociaż ma wciąż historyczną wartość, jest teraz wypierany przez późniejsze opracowania. Obecna sytuacja będzie przedyskutowana w pkt. 2.5.

Było jasne od początku, że tradycyjny podział niepewności, zwanych zwyczajowo "błędami", na przypadkowe i systematyczne stanie się jednym z najbardziej kontrowersyjnych tematów. Podział ten jest tak mocno ugruntowany w sposobie myślenia fizyków i inżynierów, że większość z nich może uważać kwestionowanie poprawności ich założeń za stratę czasu, a ponadto staranne dociekanie źródła tak kłopotliwego zamieszania niezmiennie wskazuje na ten właśnie problem jako przyczynę kłopotów.

Zgodnie z opinią większości niepewność systematyczna (albo błąd systematyczny) charakteryzuje się tym, że jej wpływ na wynik końcowy jest niezależny od liczby n przeprowadzonych pomiarów. Kontrastuje to jaskrawo z wpływem niepewności przypadkowej, która zmniejsza się proporcjonalnie do pierwiastka kwadratowego z n . Może się wydawać, że ta uderzająca różnica pomiędzy tymi dwoma typami niepewności jest nieodzowną koniecznością. Z drugiej strony

jednak trudno to pogodzić z faktem, że istnieją przypadki, gdzie typ błędu widocznie zmienia się w zależności od warunków, w jakich został wykorzystany. Jest to tak przykra sytuacja w oczach purysty od klasyfikacji błędów, że jedyną drogą zwalczania jej jest ignorowanie. Jednak fakt ten ma miejsce i całkiem łatwo jest znaleźć sytuację, w której zmiany mogą zachodzić w każdym kierunku. Dla tych, którzy wierzą w konieczność składania niepewności w różny sposób, odpowiednio do ich typu, taka obserwacja jest fatalna. Jak to się mogło zdarzyć i co możemy na to poradzić? Diagnoza kłopotów jest prosta: są one spowodowane faktem, że tradycyjna klasyfikacja niepewności zależy od przyszłego ich wykorzystania i na ogół nie może to być wiadome z góry. Jakakolwiek klasyfikacja, jeśli jest w ogóle potrzebna, powinna jednak zależeć tylko od przeszłości, a nie od przyszłości.

Chociaż wiele osób może akceptować powyższe argumenty, to jednak nie chcą oni zmieniać swoich przyzwyczajeń. Mimo wszystko przecież w większości przypadków praktycznych system pracował całkiem dobrze, i skąd mamy wiedzieć, czy niepewność ma się dzielić, czy nie, przez pierwiastek kwadratowy z n . To jest praktyczne pytanie, na które musimy znaleźć przekonującą odpowiedź. Udzielona ona była już uprzednio [4], ale ponieważ jest to istotny punkt, istnieje uzasadnienie, aby powtórzyć główne argumenty, najlepiej na praktycznym przykładzie. Przypuszczam, że czytelnicy są obeznani z ogólnymi regułami propagacji błędów, których elementy zebrane są w Dodatku A.

Rozpatrzmy szereg wazn, których wyniki liczbowe reprezentowane są przez n wyników x_j . Przypuśćmy, że mamy wątpliwości co do dokładności wzorcowych mas stosowanych w pomiarach. Gdyby taki błąd wystąpił, wpłynąłby oczywiście na wszystkie pomiary i powinniśmy rozważyć zamiast zmierzonych wartości x_j - poprawione wartości:

$$y_j = x_j + \alpha, \quad (1)$$

gdzie α jest wspólnym przesunięciem (inny rodzaj poprawek mógłby być bardziej realistyczny, a tu przedstawiono sytuację uproszczoną). Ponieważ zakładamy, że wszystkie znane poprawki zostały już uwzględnione, najbardziej prawdopodobną wartością α jest zero; pozostaje jednak niepewność $s_\alpha > 0$.

Oznaczmy przez s_x eksperymentalną niepewność pojedynczego pomiaru x_j i założmy, że wszystkie pomiary x_j są wzajemnie niezależne.

Dla "funkcji f " (patrz Dodatek A) wybieramy średnią wartość, a więc:

$$\bar{y} = \frac{1}{n} \sum_j y_j = \frac{1}{n} \sum_j x_j + \alpha. \quad (2)$$

Wielkością poszukiwaną jest niepewność wartości \bar{y} . Można ją oszacować w różny sposób, zależnie od wyrobu zmiennych. Oznaczmy najpierw

$$\bar{y} = f_1(x_1, x_2, \dots, x_n; \alpha). \quad (3)$$

Ponieważ x_j i α nie są skorelowane, wobec tego wszystkie kowariancje znikają. Ponieważ:

$$\frac{\partial f_1}{\partial x_1} = \frac{1}{n} \quad \text{i} \quad \frac{\partial f_1}{\partial \alpha} = 1,$$

od razu uzyskamy z równania (A3) wariancję wartości \bar{y} :

$$s_y^2 = \sum \left[\frac{1}{n} s_x \right]^2 + (1 s_\alpha)^2 = \frac{1}{n} s_x^2 + s_\alpha^2 \quad (4)$$

Przy innym możliwym wyborze zmiennych uzyskuje się:

$$\bar{y} = f_2(y_1, y_2, \dots, y_n), \quad (5)$$

W tym przypadku wielkościami "potrzebnymi" w równaniu (A3) są wariancje

$$s_j^2 = \text{var}(y_j) = s_x^2 + s_{\alpha}^2$$

i kowariancje

$$s_{jk} = \text{cov}(y_j, y_k).$$

Dla tych ostatnich można wykazać (porównaj [5]), że ponieważ $E(\alpha)$ i $\text{cov}(x_j, x_k)$ znikają, mamy po prostu:

$$s_{jk} = s_{\alpha}^2.$$

Ponieważ poza tym $\frac{\partial f_2}{\partial y_j} = \frac{1}{n}$ i uwzględniając (A3) otrzymamy, że:

$$s_{\bar{y}}^2 = \frac{1}{n^2} \left[n(s_x^2 + s_{\alpha}^2) + n(n-1) s_{\alpha}^2 \right] = \frac{1}{n} s_x^2 + s_{\alpha}^2,$$

a więc wynik identyczny z (4).

Wynik odpowiada dokładnie temu, czego oczekiwaliśmy. Jest to godne uwagi z wielu powodów. Po pierwsze: wynik nie zależy od wyboru zmiennych i nie trzeba było decydować, jakiego typu jest niepewność. Wszystko co musimy zrobić, to opisać dokładnie sytuację i następnie zastosować ogólne reguły propagacji błędów. Oczywiście można powiedzieć, że s_x zachowuje się jak niepewność "przypadkowa", a s_{α} jak "systematyczna", ale takie zachowanie jest rezultatem użytku, jaki z tego uczyniono.

Gdyby eksperyment polegał na porównaniu n równoważeń różnych typów, składniki s_{α} należałoby traktować jako niepewności przypadkowe. Dodatkowa informacja jest zawarta

w równaniu (4), a mianowicie, że oba składniki "przypadkowe" i "systematyczne" są dodawane "kwadratowo" (nie liniowo lub jeszcze inaczej) i że "błąd systematyczny" s pojawia się w formie odpowiadającej odchyleniu standardowemu. Podczas gdy wszystkie te sprawy są oczywiste i wynikają w sposób naturalny, to stanowiły one przedmiot gorących dyskusji przy takich analizach, których bazą nie było explicite równanie (A3). Powyższa analiza dowodzi ich błędnej podstawy. Rzeczywiście, jeżeli wariancje i ko-wariancje były właściwie wprowadzone do rachunku, to nie ma potrzeby wprowadzania różnych typów niepewności.

2.4. Pojęcia maksymalnej granicy

Można by powątpiewać, czy występują przypadki, kiedy liczbowa ocena niepewności stanowi rzeczywisty problem. Zachodzi to w szczególności wtedy, gdy nie można wyciągać wniosków z powtarzanych pomiarów, ponieważ nie można ich już wykonać. W takiej sytuacji niekompletnej wiedzy - które w rzeczywistości nie stanowi wyjątku, ale raczej normalny przypadek - fizyk powinien zrobić użytek ze swojego doświadczenia, wyobraźni, znajomości teoretycznych podstaw oraz praktycznej realizacji eksperymentu i przeanalizować wszystkie możliwe przyczyny błędów, aby je wyeliminować lub wprowadzić poprawki. Można to zrobić na ogół tylko w sposób przybliżony; resztę należy uznać za niepewność. Zawsze można mieć nadzieję, że nic się rzeczywiście ważnego nie przeoczyło, ale oczywiście nie ma na to gwarancji. Alternatywnym rodzajem klasyfikacji (może trochę żartobliwym) jest rozróżnienie tylko niepewności, które wzięto pod uwagę i niepewności, które zostały pominięte. Gdyby to było realistyczne podejście, to byłoby świetnie. W praktyce, niestety, druga kategoria może być tylko uchwycona pod sam koniec, zwykle jako wynik pewnej niepokojącej rozbieżności. Na razie należy pamiętać o tej idei, ponieważ przypomina

ona, że z naukowego punktu widzenia stanowi najbardziej interesującą kategorię: pozwala to na wykrywanie wcześniej nie podejrzewanych wpływów. W znacznej mierze odkrywanie takich właśnie błędów przyczynia się do postępu eksperymentalnej fizyki. W każdym razie wezwania do rozwikłania niekonsekwencji jest najważniejszym bodźcem do nowych doświadczeń albo prac teoretycznych [6,7]. Zachodzi też niebezpieczeństwo - wynikające z zupełnie przypadkowo stosowanej praktyki - zawyżenia wartości wyznaczonej niepewności w celu uniknięcia późniejszych trudności. Ale oczywiście problemów nie rozwiązuje się ukrywając je.

W ogólnym zamieszeniu wokół wyznaczania niepewności przyjmowano niekiedy za rzecz oczywistą, że "błędy systematyczne" automatycznie należą do kategorii, gdzie obiektywna wartość liczbowa jest nieosiągalna, gdyż w grę wchodzi osobisty sąd eksperymentatora. Ponieważ subiektywne oszacowanie zostało wykluczone jako nienaukowe, stwierdzono, że wszystko co można zrobić, to wyznaczyć "maksimum" niepewności wystarczająca duża, aby zagwarantować, że "prawdziwa wartość" leży praktycznie zawsze w obrębie tych granic. Jest to prosta recepta, ale jej użyteczność jest wątpliwa. Ponieważ "maksymalny błąd" staje się bardziej bezpieczny, gdy jest powiększany, wobec tego jest wadliwie zdefiniowana wielkość, dla której, np. nie można określić użytecznych reguł propagacji. Pomimo to zrobiono próby uratowania tej koncepcji, w szczególności przez skojarzenie jej z rozkładem prawdopodobieństwa. Interpretując górny i dolny kres jako granice, między którymi musi znajdować się "prawdziwa wartość" i zakładając następnie, że w wyniku naszej niewiedzy każdy punkt w tym przedziale ma tę samą szansę być "dobrym", dochodzi się do stałej gęstości prawdopodobieństwa między tymi granicami (rozkład jednostajny). Ponieważ odchylenie standardowe dla rozkładu jednostajnego o szerokości $2a$ wynosi $a/\sqrt{3}$ sugeruje się, aby uważać to za odpowiednią niepewność. Rzeczywiście, niektórzy biorą to przybliżenie na tyle serio, aby wykonywać szczegółowe

tablice liczbowe dla splotu gęstości rozkładu jednostajnego z funkcją Gaussa. Niestety, nie ma tu dostatecznej podstawy do takiego założenia, znanego także jako postulat Bayesa. Przeciwnie, dzięki Jaynesowi [8] znane są teraz niektóre przeciwne przykłady praktyczne dla gęstości prawdopodobieństwa a priori, co odpowiada "zupełnej niewiedzy", a w szczególności przypadek rozkładu dwumianowego z prawdopodobieństwem p jako nieznanym parametrem [9]. Odpowiedni rozkład gęstości jest tak różny od jednostajnego, że najlepiej poniechać takiego sposobu podejścia.

2.5. Sytuacja obecna

Wszystkie laboratoria, które nadesłały odpowiedzi na kwestionariusz BIPM wymieniony w pkt. 2.2, zostały zaproszone na spotkanie Grupy Roboczej, której zadaniem miało być przedyskutowanie najważniejszych problemów związanych z wyznaczeniem niepewności. Eksperti z 11 krajowych laboratoriów zbrali się w Sèvres w dniach 21 - 23 września 1980; przewodniczył dr P. Giacomo, dyrektor BIPM.

Zróbmy szybki przegląd wyników spotkania (więcej szczegółów - w publikacji [10] lub [13]). Sprawa niepewności przypadkowych, które mogą być oszacowane za pomocą serii powtarzanych pomiarów, nie wywołała poważniejszych trudności i Grupa Robocza zaleciła wyrażanie ich w zależności od estymatorów odchylenia w próbie s_i uzupełnionych stopniami swobody ν_i . Nie implikuje to żadnego założenia co do rozkładu prawdopodobieństwa. Dla pomiarów, które mogą być skorelowane, należy określić i podać odpowiednie kowariancje, przyjmując, że duże znaczenie problemu usprawiedliwia dodatkową pracę. Wyczuwało się niewielkie zainteresowanie oszacowaniem przedziałów ufności. Jeżeli można było przyjąć, że rozkład zmiennej losowej w populacji jest normalny, jak to się często zakłada, to odpowiedniość jest dobrze znana i włączanie jej nie daje nowych informacji.

Taka wskazówka może być jednak potrzebna do podejmowania decyzji statystycznych albo przy porównaniach z prawnie narzuconymi granicami tolerancji. Trudniejsza jest sytuacja, gdy nie można oszacować niepewności metodą statystyczną i w tym przypadku estymacja musi opierać się na "innych metodach", które nieuchronnie implikują pewne elementy oceny subiektywnej. Należy jednak wyraźnie podkreślić, iż ten fakt nie jest ani nowy, ani nie ma on jakiegokolwiek związku z poprzednią klasyfikacją błędów: niepewność uzyskana przez takie "oszacowanie na oko" może mieć taki sam wpływ na końcowy wynik, jakie miało - jak dotąd sądzono - oddziaływanie czynników "przypadkowych" i "systematycznych". Zależy to tylko od sposobu, w jaki wielkość rozważana "wchodzi" w nasze równania opisujące sytuację fizyczną. Niepewność wynikowa jest w każdym razie natury przypadkowej.

To doprowadziło Grupę Roboczą do wniosku, aby sugerować dla czysto praktycznych powodów i tymczasowo, szczególnie odnośnie terminologii, dwie grupy lub klasy składanych niepewności, a mianowicie:

- grupę A takich, które wyznacza się za pomocą metod statystycznych,
- grupę B takich, które wyznacza się za pomocą innych metod.

Nie jest to ani bardzo ścisła, ani bardzo głęboka klasyfikacja, ale próba przybliżenia do rzeczywistości. Ponieważ, jak zobaczymy za chwilę, wszystkie składowe począwszy od tej chwili będą jednakowo traktowane, to możliwość ich błędnego podziału nie ma żadnego znaczenia. Nie trzeba więc martwić się o dalsze szczegóły dotyczące tych grup. Rzeczywiście, różnica pomiędzy tymi grupami nie jest "fundamentalna", lecz raczej praktycznej natury i powinno to być pomocne dla eksperymentatora. Podczas gdy dla grupy A zwykle metody oceny prowadzą wprost do wariancji w próbie lub estymatora odchylenia standardowego, to zwykle nie

zachodzi to dla grupy B. Jednakże ta trudność nie jest charakterystyczna dla nowej metody; ona zawsze istniała, ale była maskowana za pomocą takich koncepcji jak "maksymalne granice" albo podobnych, których następnie nie można było użytecznie odnieść do innych składowych klasy A o lepiej zdefiniowanym znaczeniu statystycznym.

Byłoby całkiem mylące sądzić, że nowa klasyfikacja jest w istocie stara, ale prezentowana nowym językiem. Zdziwiałe, że w jednej z odpowiedzi na kwestionariusz sugerowano, iż wymiana słowa "błąd" na "niepewność" automatycznie powinna rozwiązać wszystkie poważniejsze problemy. My mamy trochę mniej wiary w siłę słów i myślimy, że pierwszą sprawą powinny być raczej pojęcia, natomiast słowa związane z nimi są mniej ważne. Stare i nowe traktowanie niepewności jest zupełnie różne z punktu widzenia zasad, a więc przejście nie można dokonać za pomocą słownika.

Jakie znaczenie chcielibyśmy przypisać składowym niepewności z grupy B? Odpowiedź może się zjawiać po uwzględnieniu ogólnego prawa propagacji błędów (równanie A3). Ponieważ jest to jedynie dostępna, silna baza, wybór jest narzucony. Jeżeli chcemy zrobić użytek z tego, to potrzebujemy dla wszystkich niepewności pewnych wielkości, które można by traktować jako przybliżenia odpowiednich odchyłek standardowych. Twierdzenie, że uzyskanie takiej aproksymacji jest niemożliwe, prowadzi do impasu, gdyż takie stanowisko implikuje, że odrzucamy jakąkolwiek możliwość osiągnięcia finalnej wartości niepewności, która zawierałaby składowe z obu grup. To jest oczywiście możliwy punkt widzenia, ale doświadczenie wykazuje, że to daje w rezultacie sytuację, która jest najmniej zadawalająca, gdyż nikt w istocie nie wie co robić, gdy występuje więcej niż jedna niepewność dotycząca wyniku końcowego. Tak więc łączenie w pewien sposób różnych (niekiedy źle zdefiniowanych) składowych niepewności pozostawia się osobom oceniającym, co często powoduje irytację eksperymentatorów. Twierdzimy, że ten ważny krok powinien być wykonywany przez osobę,

która jest najlepiej zaznajomiona ze szczegółami pomiarów, a tą osobą jest oczywiście sam eksperymentator.

Powinno być więc jasne, że jeśli chcemy podać "użyteczne" miary składowych niepewności, to powinny one być w takiej formie, która czyni zadość zwykłemu formalizmowi. Stąd wynika, że mamy wybrać coś, co może być najłatwiej dostępną aproksymacją odpowiedniego odchylenia standardowego. To stanowi istotę trzeciego punktu projektu zaleceń (por. Dodatek B). Zdajemy sobie sprawę, że sam test ogranicza się do wskazania celu, do którego dążymy bez dalszych zaleceń co do najważniejszej kwestii, a mianowicie, jak to w praktyce zrobić. Ta sprawa będzie stanowiła ważny temat dla późniejszego dokumentu, który ma bardziej dotyczyć praktycznych realizacji nowych metod niż ich tła. Ponieważ liczba możliwości jest ograniczona, możemy już naszkicować parę z nich. Wybór zależy od osobistych preferencji eksperymentatora, a w szczególności od jego dotychczasowych zwyczajów. Tak więc ktoś może być dość odważny, aby zgadywać żadaną wielkość (oznaczymy ją przez u_j) wprost; wymaga to wyczucia w odniesieniu do czegoś, co mogłoby odpowiadać np. 2/3 albo 70% przedziału ufności. Inny znowu może będzie spokojniejszy, próbując oszacować przedział ufności na 50% (znany także jako "błąd prawdopodobny"); wtedy u_j może wynosić około 1,5 raza tej wartości. Ostatecznie dla osób, które mają zwyczaj zaczynać od "bezpiecznego" przedziału ufności, problem polega na tym, aby zdecydować, jakiej to odpowiada wielokrotności odchylenia standardowego i odpowiednio ją podzielić.

Można by temu zarzucać, że ta zgodność zależy od przyjętego rozkładu prawdopodobieństwa w nieznannej populacji. W zasadzie jest to słuszne, oczywiście, ale dla przewidywanej niezbyt wielkiej dokładności założenie to zwykle nie jest krytyczne. W każdym razie nie ma sensu zastępować naszej ignoracji sztywnym przepisem. Wydaje się, że lepiej pozwolić w tej sprawie na osobiste preferencje eksperymentatorów, z których pewni mogą sobie wybrać - jako prze-

ciętą wskazówkę - rozkład gęstości prawdopodobieństwa Gaussa lub prostokątny (jednostajny); podczas gdy inni będą mogli być zwolnieni od "budowania" takich konstrukcji.

Po uzyskaniu oszacowania dla obydwu klas można je łatwo przekształcić do postaci potrzebnych w równaniu (A3) przez wymnożenie ich przez odpowiednie pochodne cząstkowe. To są właśnie kwadraty tych nowych wielkości, które ze względu na prostotę oznaczone są w projekcie zaleceń przez s_j^2 i u_j^2 odpowiednio dla klas A i B. Do tego można dodać wyrażenia na estymatory kowariancji, gdy zachodzi taka potrzeba i możliwość. Jesteśmy teraz w stanie wstawić te wartości do ogólnego wzoru na propagację błędów. Przy braku korelacji pomiędzy n różnymi niepewnościami sprowadza się to do następującej sumy:

$$\sum_{j=1}^{n_1} s_j^2 + \sum_{j=1}^{n_2} u_j^2 \quad (n_1 + n_2 = n)$$

dla uzyskania aproksymowanej "wariancji" funkcji $y = f(x_1, x_2, \dots, x_n)$. Pierwiastek kwadratowy tej wielkości został nazwany "łączną¹⁾ niepewnością". Ta prowizoryczna terminologia może prowokować sprzeciw, ale przyjęcie bardziej odpowiedniej nazwy odkłada się na później. Jednakże Grupie Roboczej radzono, aby nie używała takich wyrażen jak "odchylenia standardowe" lub "wariancja" na składniki grupy B dla uniknięcia słusznych zastrzeżeń matematyków - - statystyków, którzy przypisują tym nazwom dobrze zdefiniowane znaczenia, które nie są stosowane, jeżeli - jak w naszym przypadku - mają być do nich włączone elementy subiektywne.

Ważny punkt dotyczy relacji pomiędzy tą łączną niepewnością i pojęciem przedziału ufności. Dla składowych

1) Dosłownie "złożoną" - "kombinowaną" (combined).

z grupy A taka zależność jest zwykle ustalona dzięki przyjęciu założenia (które nawet tu często jest trudno umotywić), że wszystkie składowe mogą być uważane za próbki pochodzące z populacji normalnej, co pozwala nam na użycie współczynnika Studenta w celu przejścia od odchylenia standardowego do przedziału ufności przy określonym poziomie ufności. Jednak dla grupy B, a więc i dla łącznej niepewności, znacznie trudniej to usprawiedliwić. Jednakże naiwnością byłoby twierdzenie, że problem ten został stworzony poprzez nowe podejście zarysowane powyżej. Trudności zawsze istniały, ale były ukryte w mętnych pojęciach, podczas gdy teraz problem jest "wystawiony na światło". Jeżeli chcemy być całkiem uczciwi, to powinniśmy przyznać, że nie można ustalić określonego powiązania z poziomem ufności. Oczywiście można zakładać, albo raczej mieć nadzieję, że dla składowych należących do grupy B, zależności w przybliżeniu są takie same jak dla populacji normalnej, ale to nie jest pewne, a jest raczej pobożnym życzeniem. Odwoływanie się do centralnego twierdzenia granicznego nie wiele pomaga i rzadko jest umotywowane, przynajmniej dla sytuacji, w których występuje niewiele składników.

Wobec tych trudności Grupa Robocza uznała za konieczne przyjęcie trochę odmiennego punktu widzenia. Gdy ocena granic ufności w zwykłym znaczeniu jest niewykonalna, możemy zawsze wybrać arbitralny mnożnik łącznej niepewności, aby uzyskać szerszy przedział, a więc i wyższy poziom ufności. Nowa procedura różni się tym od konwencjonalnej, że taki mnożnik (współczynnik), oznaczany symbolem k , przyjmuje zazwyczaj takie wartości: np. $k=2$ albo 3 . Nie ma raczej wątpliwości, iż to już było w rzeczywistości często robione do tej pory i to nie z większym uzasadnieniem niż teraz; rzeczywista różnica polega tylko na otwartym przyznaniu się, że ta procedura jest w pewnej mierze arbitralna. Wartość w ten sposób uzyskana nazwana jest niepewnością całkowitą, którą to nazwę należy traktować jako tymczasową, a zalecenie 5 mówi, że ilekroć tylko jej zasto-

sowanie uważa się za niezbędne, należy w sposób wyraźny zaznaczać współczynnik. Potrzeba użycia takich mnożników, które oczywiście nie wnoszą więcej informacji niż niepewność łączna, szczególnie ostro występuje w pomiarach związanych z badaniami naukowymi, ale także często w metrologii przemysłowej lub prawnej, gdzie takie wymagania stawiają osoby, które wierzą, że wszystkie przypadkowe odchylenia mają rozkład Gaussa.

2.6. Uwagi końcowe

Problemy dotyczące wyznaczania niepewności wyników doświadczeń, jak widzieliśmy w powyższej analizie, są tego typu, który czyni ich obróbkę jednocześnie prostą i trudną. Ich dwuznaczna natura, wydaje mi się, jest spowodowana tym, że z jednej strony mają one oczywiste pokrewieństwo z matematyką i statystyką, gdzie używa się teoretycznie dobrze zdefiniowanych pojęć, ale z drugiej strony są one ściśle związane z praktycznymi zadaniami laboratoriów i pracowni; z całym ich ryzykiem i kompromisami, które są typowe dla codziennego życia praktyka. Ponieważ niepewności mogą być wykazywane albo używane przez każdego i w bardzo różnych dziedzinach, ich określanie powinno być akceptowalne w daleko szerszym zakresie, niż jest to wymagane dla bardziej specjalistycznych wielkości. Poza tym różni użytkownicy narzucają pewne ograniczenia odnośnie formatu i treści, które powinny być w miarę możliwości wzięte pod uwagę. Nie jest więc niespodzianką, że istnieje szeroki wachlarz propozycji dla osiągnięcia tego celu.

Zamiast uwzględniania bardzo różnych i często sprzecznych wymagań, można próbować rozważać je jako wytyczne pozwalające nam zmniejszać liczbę użytecznych propozycji. Pamiętając, że możemy wybrać tylko te wielkości, które są zgodne z warunkami narzuconymi przez ogólne prawo propagacji błędów, trzeba stwierdzić, iż możliwości te są tak

mocno ograniczone, że nie pozostaje praktycznie żadna możliwość wyboru. Wynikiem tego - z czym zgadza się, mam nadzieję, większość czytelników - jest propozycja naszkicowana wyżej i opisana w bardziej zwartej postaci w pięciu zaleceniach Grupy Roboczej. Nie ma wątpliwości, że różne szczegóły nadają się do ulepszenia i mogą być ewentualnie zmienione (np. oznaczenia), ale ta propozycja rozważana jako całość, co do której można sądzić, że stwarza racjonalnie logiczną i spójną całość, nie daje wiele, albo i żadnych możliwości wprowadzenia fundamentalnych zmian; gdyby to miało być nie do przyjęcia to, powinno zostać całkowicie usunięte i do problemu należałoby podejść od nowa. Mogę tylko mieć nadzieję, że to się nie wydarzy, ale rozstrzygnięcie należy do tych, którzy będą próbowali zastosować ten prosty schemat w swoich pomiarach. Według mnie kwestie związane z wyznaczeniem niepewności nie dostarczają użytecznego pola dla purystów lub fundamentalistów. Nie jest to ten rodzaj problemu, który można rozwiązać - odpowiadając tak lub nie. Niektóre kompromisy i arbitralne konwencje nie są do uniknięcia i powinny być akceptowane w interesie szerszej stosowności. Nowy schemat opisany w zaleceniach roboczych jest już stosowany w paru laboratoriach krajowych i jego praktyczna użyteczność jest właśnie badana. Jeżeli rezultaty będą zadowalające, to procedury powyższe będą mogły być zalecane do powszechnego użytku. W ten sposób jest nadzieja dojścia we właściwym czasie do akceptowalnego, jednolitego sposobu wyznaczania niedokładności pomiarów doświadczalnych.

W dziedzinie podstawowych stałych fizycznych, głównego tematu tej konferencji, niepewność zajmuje bardzo skromne miejsce. Bardziej aktualne i interesujące problemy znajdują się na innym poziomie i dotyczą najważniejszej drogi przeprowadzania koniecznej aproksymacji metodą najmniejszych kwadratów, jeżeli sobie uprzytomnić, że zarówno podana przez eksperymentatora wartość średnia jak i związana z nią niepewność, mogą być błędne. Ten ważny temat nie będzie tu w ogóle dyskutowany (patrz [7] i [11]), ale trzeba pamiętać,

że istotne w takich pracach to dysponowanie niezawodnym i realistycznym oszacowaniem niepewności. Można więc stwierdzić, że w postaci naszkicowanej powyżej ich obróbka wymaga zastosowania jedynie elementarnej i klasycznej matematyki i według wszelkiego prawdopodobieństwa nie ma potrzeby czegoś bardziej zaawansowanego (choćby propozycji przeciwnych nie brakuje). Jeżeli chodzi o resztę, to trzeba się zdać na zdrowy rozsądek. Należy specjalnie zwrócić uwagę, aby nie przyjmować jakiegokolwiek hipotezy, która nie byłaby rzeczywiście potrzebna, gdyż prowadziłoby to do mniej ogólnego rezultatu, a więc zawęziłoby to zakresów zastosowań.

Jeżeli nawet przyjmiemy optymistyczny punkt widzenia, że przyszły raport o niepewności będzie wynikał w istotny sposób z propozycji naszkicowanych powyżej, to nie należy oczekiwać szybkiego osiągnięcia jednolitości. Przyczyną są zwykle uporczywe i długowieczne. Przypomina mi się historia o filozofie, który obchodził 90 urodziny. Jeden z jego młodych zwolenników, wyrażając podziw dla faktu, iż jego opinie niegdyś ostro atakowane, są teraz ogólnie akceptowane, zapytał uczzonego, jak udało mu się przekonać wszystkich jego opozycjonistów. Starszy pan uśmiechnął się miło i odpowiedział: "Prawda jest taka, że nikt nigdy nie zmieniał opinii, ale tak się zdarzyło, że przeżyłem ich wszystkich".

Są chwile, gdy zastanawiam się, czy nie jesteśmy trochę w podobnej sytuacji.

Autor jest bardzo wdzięczny wielu członkom personelu BIMP za pomocne uwagi dokonane wg pierwotnej wersji tego artykułu. On jednak sam jest odpowiedzialny za ewentualnie kontrowersyjne stwierdzenia wyrażone w tym przeglądzie.

2.7. Dodatki

A. Przypomnienie prawa propagacji błędów

Ponieważ stanowi to wciąż bezsporną bazę każdej poważnej dyskusji na temat niepewności, warto przytoczyć w postaci

pisemnej pewne proste fakty dotyczące dobrze znanego "ogólnego prawa propagacji błędów", jak to nazywano w tradycyjnej terminologii. Wydaje się to być również usprawiedliwione ze względu na wniosek, dla którego zostało przytoczone w § 2.3 powyżej podstawowe rozumowanie, że jego skrupulatne stosowanie rozwiązuje w zasadzie wszystkie często dyskutowane w tej dziedzinie problemy, które były sztucznie stworzone przez proste pomijanie pewnych elementów podstawowych.

Jeżeli y będzie znaną funkcją n zmiennych x_j , a więc $y = f(x_1, x_2, \dots, x_n)$, to małe zmiany x_j powodują zmianę y o wielkość:

$$\Delta y \approx \sum_{j=1}^n \frac{\partial f}{\partial x_j} \Delta x_j,$$

jeżeli można pominąć człony wyższego rzędu w rozwinięciu funkcji w szereg Taylora. Kwadrat tej wielkości wynosi (również w pierwszym przybliżeniu):

$$\begin{aligned} (\Delta y)^2 \approx & \sum_{j=1}^n \left(\frac{\partial f}{\partial x_j} \Delta x_j \right)^2 + \\ & + \sum_{j=1}^n \left(\frac{\partial f}{\partial x_j} \Delta x_j \right) \cdot \sum_{\substack{k=1 \\ k \neq j}}^n \left(\frac{\partial f}{\partial x_k} \Delta x_k \right). \end{aligned} \quad (A1)$$

W praktyce często charakteryzuje się "przesunięcia losowe" lub "błędy" przez ich wariancje lub kowariancje, które można uzyskać podstawiając (E oznacza wartość oczekiwaną):

$$E\{(\Delta x_j)^2\} = \sigma_j^2 \quad \text{i} \quad E\{\Delta x_j \Delta x_k\} = \sigma_{jk}. \quad (A2)$$

Wtedy równanie (A1) można zapisać w bardziej zwartej postaci:

$$\sigma_y^2 \approx \sum_j (f'_j \sigma_j)^2 + \sum_{j \neq k} f'_j f'_k \sigma_{jk}, \quad (\text{A3})$$

przy czym $f'_j = \partial f / \partial x_j$, a sumowania należy dokonywać od 1 do n .

Równanie (A3) stanowi dobrze znany wzór ogólny na "propagację błędu". Identyfikując σ_{jj} z σ_j^2 można równanie (A3) zapisać krócej, jako:

$$\sigma_y^2 \approx \sum_{j,k} f'_j f'_k \sigma_{jk}. \quad (\text{A4})$$

Dla większości przypadków praktycznych należy zastąpić (nieznane) parametry populacji σ_j^2 i σ_{jk} , pojawiające się w równaniu (A3), przez ich odpowiednie (nieobciążone) estymatory w próbie. Tak więc jeżeli założymy, że dla wielkości x_j uzyskujemy m wyników $x_{j1}, x_{j2}, \dots, x_{jm}$ i podobnie dla x_k , to zamiast σ_j^2 można stosować:

$$s_j^2 = \frac{1}{m-1} \left[\sum_i x_{ji}^2 - \frac{1}{m} \left(\sum_i x_{ji} \right)^2 \right],$$

a zamiast σ_{jk} można stosować:

$$s_{jk} = \frac{1}{m-1} \left[\sum_i x_{ji} x_{ki} - \frac{1}{m} \left(\sum_i x_{ji} \right) \left(\sum_i x_{ki} \right) \right],$$

gdzie wszystkie sumowania względem i należy przeprowadzać od 1 do m . Istnieją również równoważne postacie, które

niekiedy mogą być bardziej praktyczne; można je znaleźć w dowolnym podręczniku statystyki.

Bardzo rozpowszechnione, choć wyraźnie nieusprawiedliwione, jest przekonanie, że w większości "przypadków praktycznych" można uważać większość składowych niepewności jako niezależnych i (konsekwentnie) pomijać wyrażenia zawierające kowariancje. Ten niefortunny przesąd doprowadził do raczej bałamutnej klasyfikacji błędów, ukrywającej ten fakt, że np. tzw. niepewności "systematyczne" nie są w żadnym razie różne od niepewności "przypadkowych", z wyjątkiem tego, że są one skorelowane z pomiarami w inny sposób. Ta sprawa jest również dyskutowana w pkt. 2.3 powyżej.

B. Projekt zalecenia w sprawie niepewności

(Przytoczono tu pełny tekst Zalecenia INC - 1 1980, opracowanego przez Grupę Roboczą ds. Wyrażenia Niepewności Pomiaru, który został już opublikowany w Biuletynie Informacyjnym, nr 1-2 (223-224) w roku 1984 [13]).

BIBLIOGRAFIA

1. Van der Waerden B.L.: Science Awakening II. The Birth of Astronomy. Noordhoff, Leyden, 1974.
2. Precision Measurement and Fundamental Constants, wyd. przez D.N. Langenberg and B.N. Taylor. Natl. Bur. Stand. (U.S.) Spec. Publ. 343, August 1971, 493.
3. The BIPM enquiry on error statements, Rapport BIMP-80/3, 1980. Przekład francuski opublikowano w Proces-Verbaux Séances Com. Int. Poids Mes. 48, 69th session, 1980.
4. Müller J.W.: Nucl. Instrum. Methods 163, 241, 1979.
5. Les incertitudes de mesures, w La Physique, tom 4 dzieła Encyclopedie Scientifique de l'Univers. Gauthier-Villars, Paris, 1981, 11.

6. Taylor B.N.: w Precision Measurement and Fundamental Constants, wyd. przez D.N. Langenberg and B.N. Taylor, Natl. Bur. Stand. (U.S.) Spec. Publ. 343, August 1971, 495.
7. Cohen E.R.: w Metrology and Fundamental Constants, Proceedings of the International School of Physics "Enrico Fermi", Course LXVIII, wyd. przez A.Ferro Milone, P. Giacomo and S. Leschiutta, North-Holland, Amsterdam, 1980, 581.
8. Jaynes E.T.: IEEE Trans. Syst. Sci. Cyb. SSC-4, 227, 1968.
9. Müller J.W.: Un nouveau regard sur les probabilités a priori. Rapport BIMP-80/6 (1980).
10. Raport z posiedzenia Grupy Roboczej ds. Niepewności Pomiaru BIMP. Ostateczna wersja francuska została opublikowana w Proces-Verbaux Séances Com. Int. Poids Mes. 49, 70th session, 1981.
11. Wöger W.: PTB-Mitteilungen 89, 401, 1979.
12. Giacomo P.: Metrologia 17, 69, 1981.
13. Określanie i oceny niepewności pomiaru w świetle praktyki wielkich ośrodków metrologicznych. Biuletyn Informacyjny, It, nr 1-2 (223-224), 1984. Instytut Łączności, Warszawa.

3. ROZWAŻANIA O NIEPEWNOŚCI - PYTANIA I ODPOWIEDZI¹⁾

"Przeczytałem, zrozumiałem, odrzuciłem".

"Przeczytałeś, ale nie zrozumiałeś; gdy-
byś bowiem rozumiał, to byś nie odrzucił".

Fragm. korespondencji między
Cesarzem Julianem II a pewnymi
biskupami [1], około AD 360.

3.1. Uwagi wstępne

W ostatnich czasach tak często i długo dyskutowano o problemach związanych z oceną i określeniem niepewności, jakie należy przypisać wynikom pomiarów, że można by przypuszczać, iż wszystko już na ten temat powiedziano i że powinno to już wystarczyć na długo. W pewnym sensie jest to słuszny wniosek; jednak pozostaje faktem, że mimo pewnych wartościowych opracowań - jesteśmy, jak się wydaje, daleko od osiągnięcia consensusu w tej sprawie, bez czego nie może być mowy o rzeczywistej i trwałej unifikacji.

Ponieważ dotychczas napisano na ten temat tak dużo, w szczególności w ciągu ostatnich pięciu lub dziesięciu lat, wydaje się już niemożliwe dodanie czegokolwiek nowego i słusznego. Tak więc wszystko co możemy zrobić, to tylko powtórnie przemyśleć te sprawy i to w sposób selektywny.

¹⁾ Oryginał angielski tego artykułu przeglądowego J.W. Müllera pod tytułem: "Ruminations on Uncertainties - Questions and Answers" został przedstawiony 2 września 1986 roku na 8 Międzynarodowym Sympozjum Metrologicznym INSYMET '86 w Bratysławie i opublikowany w "Recueil de Travaux du BIMP", vol. 10, 1985 - 1986, BIMP, Sèvres 1987.

Aby uczynić tę analizę bardziej żywą, niż można by to osiągnąć przez prezentację w sposób ściśle logiczny, wybrałem formę dialogu. Nie zastosowano tu formy czysto literackiej; w istocie, pewne najbardziej istotne uwagi - jak również sporo często spotykanych nieuzasadnionych obiekcji - jakie sły-
szałem, nie mają postaci pisemnych stwierdzeń, lecz spotykało się je w sposób spontaniczny w toku ożywionych dyskusji. Chociaż problemy omawiane poniżej nie zostały przedstawione dosłownie, to jednak stanowią ich duże przybliżenie. Jeżeli pewne z tych pytań mogą się wydawać naiwne, to proszę pamiętać, że jeszcze gorsze zostały już wyeliminowane.

Będzie można tylko poruszyć pewne wybrane problemy. Dostępne są bardziej kompletne przeglądy i zakłada się, że czytelnik jest z nimi zaznajomiony (por. np. [2 do 4] i zaznaczone tam pozycje bibliograficzne). Podobnie zaprezentował to Giacomo [5], co może służyć jako istotne uzupełnienie. Ostatnio Cohen [6] napisał również doskonały artykuł na ten temat.

Nasz dialog pogrupowano w pewien sposób, stosując dodatkowe tytuły paragrafów, lecz mają one tylko za zadanie ułatwić lekturę całości. Skład tylko dwóch osób jest wystarczająco prosty, nie wymagający dalszych wyjaśnień - z wyjątkiem może tego, że "X" (jak zwykle) oznacza osobę nieznaną, podczas gdy "Y" (dlaczego nie?) można identyfikować z autorem tej gry, którego można - abstrahując od miejscami heretyckich punktów widzenia - uważać ogólnie za reprezentanta opinii BIMP, o ile ona w ogóle istnieje.

3.2. Preliminaria

X: Zawsze uważałem, że trud, jaki ponosi eksperymentator dla oceny błędu swego pomiaru, stanowi czas stracony. Czy nie należałoby raczej doradzić mu, aby zamiast tego poświęcił go na wykonanie większej liczby i lepszych pomiarów?

Y: Pańska rada na pierwszy rzut oka może wydawać się rozsądna, ale w istocie nią nie jest. Skąd może Pan, na przykład, wiedzieć, czy zwiększając liczbę pomiarów polepsza Pan wynik? Przede wszystkim musi Pan zanalizować, w jakim sposób, co Pan usiłuje akurat zmierzyć. To na ogół doprowadzi do zestawienia krótkiej listy głównych przyczyn niepewności, które mogą wpływać na Pański pomiar. W rezultacie może się okazać, że główny składnik niepewności pochodzi od czynnika, co do którego zakłada się, że jest znany, lecz który nie jest mierzony. W tym przypadku byłoby bezcelowe dokonywanie większej liczby odczytów; wtedy opłaca się zwrócić baczniejszą uwagę na to główne źródło niepewności i usiłować zmniejszyć ją, być może dzięki specjalnie planowanemu eksperymentowi. W każdym razie dla wszelkich danych wejściowych trzeba dysponować niezawodnymi estymatorami ich odpowiednich niepewności. Rzeczywiście, jedyną obiektywną miarą do wyrokowania o jakości wyniku pomiaru jest właśnie jego niepewność. Tak więc przed jej oszacowaniem wynik pomiaru ma minimalne znaczenie; wszystko zależy od jego niepewności. Jeżeli ona nie jest wyznaczona lub jest podana w dwuznaczny sposób, jak się to często zdarza, to odpowiedni wynik pomiaru nie może być oceniony obiektywnie. Jest to często przyczyną błędnej interpretacji i w konsekwencji pewnego trendu (przesunięcia - polaryzacji). Dalej zakładamy milcząco, że zastosowano już wszystkie znane poprawki; tak więc pozostałości stanowią niepewności wyników.

X: Zatem uważa Pan, że jasne określenie niepewności jest rzeczą istotną i że należy tego dokonywać na wspólnej podstawie?

Y: Należałoby niekiedy podkreślać, że unifikacja ustaleń o niepewnościach przypomina wprowadzenie międzynarodowo akceptowanych jednostek przez zastosowanie układu SI i intencją było prawdopodobnie podkreślenie jej wagi. W istocie porównanie implikuje pewne niedomówienie. Tak więc, czy mierzyć odległość do księżyca w centymetrach, czy

w calach, jest to zupełnie równoważne, ponieważ stosunek tych jednostek jest dokładnie znany (z definicji). Podobnie czy pisać s lub sec na oznaczenie sekundy jest sprawą, która obchodzi tylko purystów, lecz poza tym nie ma żadnego znaczenia. Jeżeli natomiast niepewność precyzyjnego pomiaru jest podana w niejasny sposób (jak to się stale zdarza), to wartość całego pomiaru jest bardzo problematyczna. Jest to, oczywiście, ważniejszą sprawą niż kwestia, czy litr powinien być oznaczany symbolem l, czy też L.

X: Teraz już nie mam hic przeciwko temu, aby zaakceptować, że wypowiedź o niepewności stanowi istotną część pomiaru. Jednak było to ogólną praktyką przez długi czas i nie mogę zrozumieć, co tu jest nowego - abstrahując od dodatkowego bałaganu. Osobiście byłem zawsze bardzo zadowolony ze starego systemu, który był łatwy do zrozumienia i prosty w praktyce. Po co narzucać nowy system wyrażania niepewności?

Y: Pańskie pytanie wykazuje, że albo nie jest Pan zbyt wymagający, albo jest niezbyt dobrze poinformowany. Mówiąc, że stary system dobrze funkcjonował, implikuje się, że w ogóle istniał jakiś system, a to nie jest prawdą. Już dyskusje, jakie miały miejsce na konferencji w Waszyngtonie w 1970 roku na temat stałych fizycznych [7] wykazały szereg poważnych niezgodności. Jednak głównie ankieta zorganizowana przez BIMP w 1978 roku [8] pokazała wyraźnie, że każdy stosuje swój własny system - często nie przestrzegając go i nie wskazując, w jaki sposób dochodzi się do tej niepewności. Dla każdego, który ma analizować dane pochodzące z różnych źródeł - a dla BIMP zachodzi to przy każdym porównaniu międzynarodowym - ten niedostatek stanowi dokuczliwą przeszkodę, uniemożliwiającą jakąkolwiek rozsądną analizę.

Tak więc nie było starego systemu, lecz duża różnorodność metod. Większość z nich była zupełnie dowolna i nie miała żądanej obiektywnej podstawy. Propozycja opracowana przez specjalnie powołaną grupę roboczą, znana jako zalecenie BIMP, stanowi próbę przezwyciężenia większości poważnych

niedociągnięć poprzednich metod, możliwego uniknięcia arbitralnych (a więc i dyskusyjnych) konwencji i oparcia całości na głębszej podstawie naukowej. Poza tym reguły powinny być łatwe do stosowania w praktyce. Ponieważ zna Pan zalecenie BIPM, które było szeroko publikowane (np. w [3 do 6]), nie będę go tu powtarzał.

3.3. Istota zalecenia BIPM

X: Tak, przeczytałem tekst, ale nie mogę zrozumieć, dlaczego to należy uznać za postęp. Co się zyskuje, na przykład, przez nazwanie błędów, które każdy łatwo rozpoznaje jako przypadkowe i systematyczne, a teraz nazwane jako typu A i typu B?

Y: Gdyby to co Pan myśli było prawdziwe, a mianowicie, że zaleca się zamienić tylko pewne znane wyrażenia przez nowe, to rzeczywiście byłoby to bezużytecznym ćwiczeniem. Nie można zaprzeczyć, że problemy terminologii mają pewne znaczenie. Dobrze dobrane terminy są podobne do symboli matematycznych, towarzyszących nam w naszym rozumowaniu. Jednakże język jest jak żywa istota i stale się rozwija. W ogóle naukowcy mało troszczą się o oficjalne nazewnictwo. Oni chcą być zrozumiani i dlatego stosują terminy narzucające się same w praktyce. Poza tym niebezpieczne jest wybierać słowa zanim nie zostaną w pełni wyjaśnione pojęcia, do których się one odnoszą. Niestety, taka jest właśnie sytuacja dotycząca terminów statystycznych, które stosujemy do opisu niepewności.

Gdyby Pan przeczytał staranniejsze zalecenie BIPM - a w tej sprawie cytat zamieszczony na początku tego artykułu może być znamieny - to przekonałby się Pan, że dotyczy on w istocie pojęć, a nie terminów. Podczas gdy dawny podział (przypadkowy, systematyczny) zależał od sposobu, w jaki miała być stosowana w dalszym ciągu wielkość mierzona, to zalecana obecnie klasyfikacja na typy oznacza po prostu,

sposób, w jaki uzyskuje się te niepewności. Tak więc są to dwa zupełnie różne schematy.

Klasyfikacje są prawie zawsze zabiegami nieco sztucznymi, niezależnie od tego, czy są dobre czy złe, różnią się od siebie ich praktyczną przydatnością. Jak długo wierzone, że niepewności systematyczne i przypadkowe wymagają różnej obróbki przy ich propagacji, tak długo wydawała się nie tylko pożyteczną, ale i konieczną klasyfikacją. Obecnie przyjmuje się ogólnie, że jedyna podstawa naukowa jest dana przez ogólne prawo propagacji, a ono nie rozróżnia składników przypadkowych od systematycznych.

X: Ja jednak utrzymuję, że taka separacja jest istotna. Tak więc, na przykład, błędy przypadkowe można zmniejszać zwiększając liczbę pomiarów, podczas gdy błędy systematyczne pozostają stałe. W jaki sposób może Pan dojść do rozsądnego systemu, jeżeli pomija Pan to podstawowe rozróżnienie?

Y: My niczego nie pomijamy; przeciwnie, spoglądając na te problemy w nieco bardziej ogólny sposób możemy dojść do zupełnie zadowalającego i zrozumiałego opisu, gdzie to co Pan nazwał podstawowym rozróżnieniem okaże się raczej trywialną konsekwencją. Istotą zagadnienia, którą należy zapamiętać, jest to, że wielkości rozważane są nie zawsze od siebie niezależne. Ta zależność wzajemna lub korelacja jest na ogół zbyt skomplikowana (lub w ogóle niezrozumiała), aby dokonać rozsądnej analizy. Jednakże można to opisać statystycznie. Najprostszym i najbardziej użytecznym parametrem, opisującym w sposób formalny taką korelację, jest tak zwana kowariancja (przy dwóch wielkościach oznaczana przez $\text{Cov}(x_1, x_2)$). Nie jest to przypadek, że właśnie ta wielkość pojawia się razem z wariancjami we wzorze (prawie) na propagację błędu.

Interesujący jest przypadek szczególny, mianowicie wtedy, gdy na wszystkie mierzone wielkości wpływa ta sama niepewność s . Może się to zdarzyć, przy pomiarach czasu za pomocą

zegara, który został niedokładnie nastawiony lub przy ważeniu za pomocą wagi, jeżeli pozycja zerowa jest wadliwa. W obu tych przypadkach pomiary są skorelowane i można wykazać, że w tej sytuacji wszystkie kowariancje par pomiarów są te same i wynoszą s^2 , która jest wariancją przyczyny wspólnego rozrzutu. Podstawienie tego wyniku do prawa propagacji błędów pokazuje od razu, że składnik s^2 pozostaje stały niezależnie od liczby pomiarów. Przy metodzie tradycyjnej taki składnik byłby nazywany niepewnością systematyczną i wymagałby zastosowania specjalnej procedury.

X: Jeżeli wszystko to jest prawdziwe, to wynika z tego, że stara procedura była bezpieczna. Dlaczego więc zmieniać i stosować skomplikowane pojęcie kowariancji?

Y: Ponieważ pojęcie kowariancji jest bardziej elastyczne i pozwala nam rozpatrywać również wszystkie przypadki pośrednie. Poza tym rozwiązuje to automatycznie wszystkie wadliwie postawione problemy dotyczące propagacji błędów (składanie liniowe lub kwadratowe).

X: Słyszałem, że niepewności systematyczne nie powinny być nigdy mieszane z niepewnościami przypadkowymi, lecz trzymane oddzielnie i specyfikowane. Zdaje się, że było to też wymagane w pewnym "code of practice". Na ile jest to zgodne z procedurą proponowaną przez Pana, gdzie wszystkie składniki systematyczne i przypadkowe są traktowane na równej stopie?

Y: Oczywiście nie ma tu zgodności! Jednak publikacja [9], do której Pan nawiązał, została tymczasem uznana jako nie-realistyczna, nawet przez pewnych jej autorów. Istnieje ku temu kilka powodów. Po pierwsze, łatwo zauważyć, że niepewność przypadkowa, na przykład, gdy wykorzystywana jest w kolejnym etapie, staje się na ogół "systematyczną", a może i zejść przypadkiem odwrotny. Już samo to wystarczy, aby wykazać, że warunek konieczny programu, aby utrzymywać te dwie

kategorie zupełnie oddzielnie, nie może być w rzeczywistości realizowany. A poza tym, co mamy robić z wynikiem zawierającym dwie niepewności? W wielu przypadkach o wielkim znaczeniu praktycznym, a w szczególności we wszystkich aproksymacjach metodą najmniejszej sumy kwadratów, nie ma sposobu (i nie ma aktualnie potrzeby) łączenia ze sobą tych dwóch niepewności dla każdej wartości wejściowej; a dla aproksymowanej wartości wyjściowej nikt nie może rzeczywiście twierdzić, czy ich odpowiednie niepewności, jak to podają dobrze znane procedury obliczeniowe, są przypadkowe czy systematyczne. Tradycyjna klasyfikacja po prostu traci jakiegokolwiek znaczenie. Dlatego też wydaje się rozsądne unikać od początku takiego systemu i to właśnie proponujemy.

X: Ale dlaczego wobec tego proponuje się podział na "typ A" i "typ B"?

Y: No właśnie, głównie ze względów praktycznych i pedagogicznych. Mówi nam to, czy wartość liczbowa przyporządkowaną niepewności uzyskano przez pewną procedurę statystyczną (np. przez powtarzanie pomiarów lub także przez aproksymację), czy też przez wyczucie oparte na doświadczeniu lub innych pośrednich informacjach. Tak więc na ogół niepewność "typu A" będzie lepiej znana niż niepewność "typu B", przy czym ta ostatnią będzie nieuchronnie obarczona oceną subiektywną. Jak długo składniki "typu B" nie będą mogły być pomijane tak długo nie będzie można oczekiwać, że ostateczna niepewność będzie zupełnie obiektywna. Pamiętajmy jednak o tym, że jakakolwiek niepewność oparta na skończonej próbie n pomiarów ma sama nieuniknioną niepewność statystyczną i że nawet dla $n = 10$ wynosi ona około 24% (dla rozkładu normalnego, por. tablicę 1 w [2]). Tak więc nie należy zapominać, że nawet estymator typu A może mieć małą precyzję. Istotną różnicą jest jednak to, że w przypadku typu B nie ma redundancji w estymacji lub innymi słowami: nie pozostawiono żadnego stopnia swobody. Implikuje to dla niepewności typu B, że nie możemy w żaden obiektywny sposób dojść do prze-

działów ufności. O tym fakcie nie należy zapominać. Tyle jeżeli chodzi o stronę pedagogiczną. Poza tym podział na typy nie ma żadnych konsekwencji i dlatego można go często pominąć.

Ponieważ przez cały okres dyskusji w Grupie Roboczej BIMF nie poruszano aspektów terminologicznych, należy uważać wyrażenia "typ A" i "typ B" jako prowizoryczne.

X: Dlaczego nie stosować "maksymalnych granic", jak to się na ogół dzieje z błędami systematycznymi? One są tak wygodne.

Y: Może to jest i prawdziwe, że maksymalną granicę łatwo znaleźć i w przypadku wątpliwości powiększamy ją nieco. A jednak nietrudno jest pojąć, że taka wartość jest raczej bezużyteczna, jeżeli mamy ją wyrażać ilościowo albo w zależności od prawdopodobieństwa błędu, albo jako wielokrotność odpowiedniego odchylenia standardowego. Taka "transformacja" jest wtedy potrzebna, gdy próbujemy stosować prawo propagacji, ponieważ niezbędnymi wielkościami są odchylenia standardowe (lub wariancje), a nie maksymalne granice. Jest rzeczą ważną uprzytomnić sobie, że nie istnieje ogólna prawo propagacji dla przedziałów ufności.

X: Ale nie można przecież bezpośrednio estymować odchylenia standardowego, podczas gdy mogą łatwo znaleźć maksymalną granicę.

Y: Jeżeli na przykład Pańska maksymalna granica może być identyfikowana z przedziałem ufności 99%, to dzieli ją Pan przez trzy i będzie Pan miał grube oszacowanie szukanego odchylenia standardowego. W sprawie innych możliwych "przepisów" porównać należy [2] lub [5]. Ujmując to krótko, można powiedzieć, że rzeczywiście nie trudniej jest uzyskać nową wielkość niż starą; w dużym stopniu jest to sprawą jej pozyskiwania. Istotnie, próba bezpośredniego oszacowania odchylenia standardowego zamiast (szerokiego) przedziału

ufności może się okazać bardziej niezawodna w ostatecznym rezultacie. Warto pomyśleć o problemie oceny (po pobieżnym spojrzeniu) wielkości, powiedzmy, słonia. Łatwiej będzie podać wartość grubości "głównego ciała" lub jego "grubości w połowie" niż jego "całkowitej długości" wraz z jego trąbą i ogonem. Zwróćmy uwagę, że dinozaur służył jako bardziej przekonująca ilustracja, ale jest dzisiaj rzadko spotykany. W każdym razie jest rzeczą osobliwą, że ludzie wolą oceniać część skrajną niż centralną empirycznego rozkładu prawdopodobieństwa, a więc właśnie ten rejon, który jest najmniej znany przy skończonej liczbie obserwacji.

X: W końcu niepewność, reprezentowana w postaci estymatora odchylenia standardowego, odpowiada przedziałowi, który w wielu zastosowaniach (np. przemysłowych) nie jest dostatecznie bezpieczny: uważa się za niedopuszczalne, aby jedna trzecia wszystkich przypadków przekraczała wartość ustaloną. Co więc należy zrobić?

Y: Ostatni punkt zalecenia BIMP wyraźnie stwierdza, że złożona niepewność może być wymnożona przez pewien współczynnik k , przez co uzyskuje się ogólną niepewność. Dokładna wartość współczynnika k nie jest zazwyczaj znana. Pewną inspirację można uzyskać, porównując z przypadkiem rozkładu normalnego, dla których łatwo znaleźć odpowiednie współczynniki. Jeżeli rzeczywisty rozkład nie jest znany, wystarczy podawać zaokrąglone wartości współczynnika k , ale należy je wyraźnie zaznaczać.

3.4. Sytuacja obecna i przyszła

X: Jaki jest obecny status zalecenia BIMP? Czy na przykład laboratoria narodowe stosują je?

Y: Należy przede wszystkim podkreślić dwa punkty. Pierwszy polega na prostym fakcie, że Grupa Robocza BIMP

ds. Niepewności doszła do porozumienia i była zdolna do wydania zalecenia. Jest to wynik, którego niewiele oczekiwało - po niepowodzeniu wielu podobnych prób w przeszłości. Drugi punkt stanowi to, że - że o ile możemy to ocenić - zalecenie zostało w całości przyjęte z zadowoleniem przez laboratoria narodowe, które są naszymi bezpośrednimi korespondentami. To również mogłoby być bardzo różnie.

Chociaż wiemy już, że zalecenie zostało przyjęte z powodzeniem we wszystkich porównaniach organizowanych przez BIMP od roku 1980, mamy jednak mało informacji o rzeczywistej sytuacji wewnętrznej w różnych laboratoriach. W roku 1981 BIMP zaleciło, aby dalej badać tę metodę w celu ujawnienia ewentualnych trudności praktycznych i możliwie bardziej podstawowych problemów, które, być może, były przeoczone i BIMP było proszone o przedstawienie sprawozdania dla CIMP po próbnym okresie dwóch, trzech lat [5]. Ponieważ ten plan nie został wykonany i ponieważ nie odbyło się nowe zebranie Grupy Roboczej, nie można obecnie przedstawić wiarygodnego przeglądu całej sytuacji. Wszystkie odpowiedzi, które do nas dotarły, były pozytywne. W RFN na przykład wydano ostatnio przepisy normalizacyjne na temat niepewności [10]; podobne wytyczne istnieją również we Francji [11]. W obu przypadkach dokumenty zostały oparte na zaleceniu BIMP. Z drugiej strony, w pewnych dziedzinach (np. w fizyce wielkich energii) można się było przekonać o raczej niefortunnym zwyczaju podawaniu w publikacjach dwóch, a nawet trzech niepewności eksperymentalnych, pozostawiając już czytelnikowi dalszy ich los. Jeżeli trzeba przekazać bardziej szczegółową informację, to może się okazać przydatna specyfikacja liczbowa różnych składników.

X: A co stanie się w najbliższej przyszłości?

Y: Na swojej ostatniej sesji w 1985 r. CIMP sugerowało - mając na uwadze bardzo szerokie praktyczne implikacje jakiegokolwiek próby międzynarodowej normalizacji wyrażenia na niepewność pomiaru - aby BIMP nawiązało kontakt z ISO w tej

sprawie w celu wdrożenia tych nowych pojęć w skali światowej. W tym czasie ISO utworzyło Grupę Roboczą (ISO TAG-4, z R.Collé z NBS jako jej przewodniczącym) i jej pierwsze zebranie przewidywane jest w następnym miesiącu (październik 1-3, 1986) w Paryżu. Może tam zapaść ważna decyzja i wszyscy oczekujemy tego z dużym zainteresowaniem.

X: A jeżeli chodzi nie o bezpośrednią przyszłość, jakie jest Pańskie osobiste przekonanie? Czy sądzi Pan, że na dłuższą metę może rzeczywiście nastąpić ujednoczenie w skali całego świata?

Y: "Jest rzeczą nierozsądną prorokować, szczególnie jeżeli chodzi o przyszłość" jak jesteśmy pouczani przez stare chińskie przysłowie. Jest mi oczywiście bardzo przyjemnie, że do obecnej chwili nie wykryto poważniejszej wady w zaleceniu BIMP, chociaż pewne szczegóły można by na pewno poprawić. Jest także rzeczą zupełnie jasną, że praktyczne wdrożenie będzie wymagało opracowania bardziej szczegółowego dokumentu (z przykładami zaczerpniętymi z różnych dziedzin); nie było zamiaru, aby to zalecenie spełniało takie zadanie.

Poza tym nie powinniśmy zupełnie zapominać we wszystkich tego typu dyskusjach, które często dotyczą drobniejszych punktów, że pozostaje bardziej istotny problem ogólny dotyczący wszystkich zastosowań, a mianowicie chodzi o to, aby być pewnym, że nie przeoczyło się najważniejszego składnika niepewności wynikowej - właśnie z tego powodu, że nikt nie myślał o nim. To zdarzyło się nie raz i niekiedy staje się to widoczne, gdyż występuje poważna statystyczna niezgodność. Zrozumienie przyczyny rozbieżności może się stać rzeczywistym krokiem do przodu.

Ocena niepewności eksperymentalnej nie jest odpowiednim polem dla ludzi zagorzałych i jakakolwiek próba, aby oszacować ją dokładniej niż około 20% jest prawdopodobnie iluzoryczna. W ogólnym przypadku ocena niepewności jest w większym stopniu ćwiczeniem zdroworozsądkowym niż obiektem rozważań filozoficznych.

Jeżeli chodzi o końcowy sukces tego zadania to sądzę, że nie powinienem przed Panem skrywać, że moje osobiste doświadczenie z problemami tu poruszonymi, jak również z naturą ludzką w ogólności, znacznie osłabia mój naturalny optymizm. Ilekroć próbujemy zmienić nabyte przyzwyczajenia ludzi - a wyrażanie niepewności należy oczywiście pod względem psychologicznym właśnie do tej kategorii - wtedy zasięg czasowy obejmuje już nie lata, ale pokolenia. Zmiana reguł i wytycznych jest wprawdzie prosta, ale zmiana ludzkich opinii jest właściwie niemożliwa. Przypomina mi to pewien fragment w liście Cesarza Juliana [1], gdzie pisze on.

"Chociaż wszyscy oklaskują moje słowa, to jednak tylko niewielu jest przekonanych nimi, a i to są ci, którzy wydawali się już dzielić trafne opinie, zanim do nich przemówiłem, ale nie śmieli pokazać tego w sposób otwarty".

Byłbym bardzo szczęśliwy, gdyby dalszy postęp wykazał mi, że byłem zbytnim pesymistą.

3.5. Uwagi końcowe

Postać dialogu ma zaletę polegającą na konfrontacji opinii (lub uprzedzeń) i na ukazywaniu ich w jasnym świetle. Z drugiej strony nie nadaje się ona zupełnie do jakiegokolwiek rodzaju formalnego opracowania typu rozwojowego. Dlatego też brak w naszej prezentacji wzorów matematycznych, co wcale jednak nie oznacza, że są one nieistotne. Wręcz przeciwnie, dostarczają one koronnych argumentów przemawiających za metodą przedstawioną powyżej słowami. Być może czytelnik będzie teraz miał motywację, aby zapoznać się również z bardziej formalnymi aspektami, szukając tych czy innych publikacji wymienionych powyżej.

Na zakończenie warto wspomnieć dwa problemy, które nie mogły być przedstawione w postaci pytań. Są one zupełnie świeżej daty, ale ich waga może usprawiedliwiać krótką

analizę. Dotyczą one wariancji Allana i problemu wag statystycznych.

Tak zwana wariancja Allana jest wielkością mającą naturę wariancji, ale jest obliczana w inny sposób: zamiast żeby tworzyć, dla danego szeregu pomiarów, wszystkie kwadraty różnic między wynikami indywidualnymi a ich wartością średnią, jak to się zwykle dzieje, stosuje się tutaj tylko kwadraty różnic między kolejnymi pomiarami; wartość średnia nie stanowi już punktu odniesienia. Jej główną zaletą jest prawie zupełny brak wrażliwości na możliwe małe dryfty. Elementarna analiza podana jest w [4]. Ostatnio Allan opublikował [12] artykuł, w którym sugeruje stosowanie tego wyrażenia w sposób zupełnie ogólny, a więc proponuje poniechanie tradycyjnych wzorów, których wszyscy uczyliśmy się. Argumentuje on to w ten sposób, że jeżeli tylko mierzona wielkość podlega zakłóceniom typu szumu migotania o widmowej gęstości mocy o charakterze $1/f$, to wariancja o tradycyjnej postaci jest rozbieżna, podczas gdy nowa pozostaje skończona. Jest to oczywiście bardzo istotny fakt i w dziedzinie pomiarów czasu, gdzie można osiągnąć najwyższą dokładność, stosowanie wariancji Allana stało się normalną praktyką już od wielu lat.

Ale może się pojawić nowy zaskakujący element. W [12] twierdzi się, że w sposób oczywisty istnieje szum migotania również i w innych procesach pomiarowych, w szczególności przy sprawdzianach i ogniwach normalnych. Jest rzeczą jednak rozsądną odczekać pewien czas na niezależne potwierdzenie tej tezy. Sygnalizowana uprzednio obserwacja szumu typu $1/f$ w zjawisku szybkości zaniku radioaktywności, gdy pomiar rozciągał się siłą rzeczy na bardzo długi okres, uznawano od tego czasu za zjawisko bez uzasadnienia [13]. W każdym razie należy bardzo dokładnie badać dalej tę sprawę. Jeżeli okaże się, że rzeczywiście należy stosować wariancję Allana, to nie będzie to powodowało żadnych kłopotów dla prawa propagacji błędów. Oczywiście, należałoby to stosować równoległe z odpowiednio zmodyfikowaną definicją kowariancji.

Może jednak powstać problem właściwej interpretacji, ponieważ wariancja Allena nie może już być uważana jako miara stabilności wartości średniej.

Inny problem, o którym proponowałem wspomnieć, dotyczy wag statystycznych. Chociaż stanowi to bardzo starą i elementarną procedurę, może ona ukrywać niespodzianki. Normalnie bierze się wagę statystyczną pomiaru jako odwrotność jego wariancji, ale tym samym zakłada się, że wyniki są niezależne. W jaki sposób kowariancje wpływają na wagi statystyczne? Ten problem został rozwiązany dopiero ostatnio i mamy nadzieję opublikować go w najbliższej przyszłości. Tu wystarczy może wspomnieć o wyrazistej i raczej zaskakującej właściwości, a mianowicie o fakcie, że wagi mogą przybierać nie tylko wartości dodatnie, ale i ujemne. Przyjemność znalezienia prostej interpretacji tego faktu pozostawiamy na razie czytelnikowi.

Wreszcie trzeba zbliżyć się ku końcowi. Deje mi to ostatnią okazję do przytoczenia znowu słów z listu Cesarza Juliana [1], gdzie stwierdza on:

"Powiedziałem, co było do powiedzenia, jak to mówią oratorzy; teraz możecie sami decydować, co jest potrzebne".

BIBLIOGRAFIA

1. Julian: Briefe. Wydanie grecko-niemieckie, B.K. Weis. Heimeran, Monachium, 1973.
2. Müller J.W.: Some second thoughts on error statements. Nucl. Instr. and Meth., 163, 1979, 241-251.
3. id: The assignment of uncertainties to the results of experimental measurements. W Precision Measurement and Fundamental Constants II, NBS Special Publ., 617, 1984, 375-381.

4. id.: The treatment of measurement uncertainties. W Proceedings of the IMEKO Interregional Training Course on Ensuring Measurement Accuracy, Seibersdorf, 1984, 19.
5. Giacomo P.: On the expression of uncertainties. W Quantum Metrology and Fundamental Physical Constants, Plenum, New York, 1983, 623-630.
6. Cohen E.R.: The unification of random and systematic uncertainties. Measurement Science Conference, Irvine, 1986, 9.
7. Panel discussion: Should least squares adjustments of the fundamental constants be abolished?. W Precision Measurement and Fundamental Constants, NBS Special Publ., 343, 1971, 493-525.
8. The BIMP enquiry on error statements. Raport BIMP-80/3, 1980, 50. Francuski przekład w Proces-Verbaux Séances Com. Int. Poids Mes., 48, 1980, C1-C36.
9. Campion P.J., Burns G.E., Williams A.: A code of practice for the detailed statement of accuracy. NPL, London, 1973.
10. Grundbegriffe der Messtechnik. DIN 1319, 4, Beuth, Berlin, 1985, 18.
11. Bureau National de Métrologie: Demande d'habilitation d'un service de métrologie en métrologie dimensionnelle. dokument BNM, No 88-86, 1986, 48.
12. Allan D.W.: Should the classical variance be used as a basic measure in standards metrology?. Konferencja na temat precyzyjnych pomiarów elektromagnetycznych. NBS, 1986, 29.
13. Winterbon K.B.: 1/f noise in counting statistics. W Progress report PR-P-139, AECL, Chalk River, 1984, 4-5.

4. NIEPEWNOŚĆ POMIARU - OSOBISTY RZUT OKA
 NA ZNACZENIE TEGO POJĘCIA W PRZESZŁOŚCI
 I PRZYSZŁOŚCI ¹⁾

"Freilich ist geschrieben, aber geschrieben ist nicht gedruckt, und gedruckt ist nicht erwiesen". ²⁾

Ulrich Bracker (1780)
 der "Arme Mann im Tockenburg" [1]

4.1. Uwagi wstępne

Temat tego seminarium w PTB ³⁾, a mianowicie eksperymentalna niepewność pomiaru, ma zaskakującą właściwość, polegającą na tym, że jest jednocześnie staromodny i nowoczesny. Aspekt tradycyjny jest oczywisty przynajmniej dla metrologów, jeżeli sobie uprzytomnić, że sam liczbowy wynik pomiaru jest praktycznie bez wartości, jeżeli nie towarzyszą mu dane dotyczące niepewności pomiaru. Przeglądając wcześniejsze roczniki czasopism fizycznych można się łatwo przekonać, że taka praktyka jest raczej świeżej daty; jeszcze w pierwszych dziesięcioleciach naszego wieku wiarygodność wyniku pomiaru była określona przede wszystkim autorytetem autora.

-
- 1) Oryginał niemiecki tego referatu J.W.Müllera pod tytułem "Messunsicherheiten - ein persönlicher Rück-und Ausblick" został przedstawiony 23 września 1987 roku w Brunświku z okazji 79 Seminarium PTB pt. "Posługiwanie się niepewnościami pomiaru przy wykonywaniu pomiarów w miernictwie naukowym, przemysłowym i prawnym" i został po raz pierwszy opublikowany w PTB - Bericht PTB - FMRB, 118, Vorträge 79, PTB - Seminars, wyd. K. Weise, Braunschweig, 1988., a po raz drugi w "Recueil de Travaux du BIMP", vol. 11, 1987-1988, Sèvres, 1989.
- 2) Oczywiście, to jest napisane, ale napisane to nie jest wydrukowane, a wydrukowane to jeszcze nie dowiedzione.
- 3) PTB - Physikalisch - Technische Bundesanstalt (Federalny Instytut Fizyki i Techniki).

To jednak, że wybrany temat jest również bardzo aktualny, wynika stąd, że jeszcze przed 20 laty nikt nie wpadłby na pomysł, aby prowadzić na ten temat dyskusję, ponieważ wydawało się, że na wszystkie związane z tym pytania znaleziono już dawno definitywną odpowiedź. Tak więc naszym zadaniem jest wykazać powody, które doprowadziły do zasadniczej zmiany orientacji w pewnej dziedzinie metrologii, która w przeciwnym przypadku byłaby uznana za ugruntowaną.

PTB odgrywa od dawna bardzo aktywną rolę usiłując znaleźć ogólnie uznaną i praktycznie użyteczną normę dla problemów związanych z podawaniem niepewności pomiaru. Te usiłowania stały się w ostatnich latach coraz bardziej intensywne i znalazły swój wyraźny oddźwięk w przepisach wewnętrznych, publikacjach i aktach normatywnych DIN [2,3]; nasze seminarium znajduje się również w tym nurcie.

Wobec olbrzymiej liczby opracowań, w których zajmowano się w ubiegłych latach interesującymi nas zagadnieniami, nie tylko stało się stratą czasu "przebieranie ziarna od plew", lecz również - dla referenta - znalezienie tematu, który w pewnej mierze będzie wyglądał na nowy i nie będzie wymagał od słuchaczy znajomości zbyt wielu technicznych szczegółów. Może się przy tym wyłonić tendencja, że korzystne będzie rozszerzenie tradycyjnego horyzontu tak pod względem czasowym jak i przestrzennym.

Przy poszukiwaniu przyczyn, które doprowadziły do tych nowych tendencji, pojawi się również okazja wniknięcia w pewne problemy, które są jeszcze stale przedmiotem dyskusji. Przy takim opisie nie uda się, oczywiście, uniknąć pewnego osobistego zabarwienia, tym niemniej BIMP może się okazać odpowiednim do tego punktem obserwacji. Proszę jednak nie oczekiwać teraz ode mnie "historii rachunku błędów", gdyż nie mam do tego ani umiejętności, ani czasu, ani wreszcie ochoty. Jeżeli ten temat miałby być w przyszłości jednak omawiany przez jakiegoś kompetentnego autora, to mogłyby wystąpić pewne niespodzianki. Mogło by się mianowicie okazać, że to co uważało by się później za konieczny rozwój,

stanowi raczej następstwo losowych zdarzeń, pewien rodzaj przedstawienia teatralnego, którego przebieg w istocie określony jest przez kilku głównych aktorów.

4.2. Spojrzenia tradycyjne

Chyba tak jak większość z Państwa również i ja spotkałem się z tematem "błąd pomiaru" po raz pierwszy jako student podczas pierwszych zajęć praktycznych z fizyki. Ponieważ już byłem namiętnie do tego przyzwyczajony, aby możliwie najpilniej słuchać przekazywanych pouczeń, to spotkanie przebiegało bez wielkiego problemu; do krytycznego myślenia brakowało mi jeszcze warunków. Co było podówczas "kanonicznym" mniemaniem, a więc i materiałem nauczania, w sprawie "propagacji błędów"?

Zasadnicze informacje w tej sprawie zestawiono w tabelicy 1 poniżej. Zgodnie z tym "świat błędów" rozpada się na dwie półkule, które są od siebie oddzielone. W pierwszej połowce występują tzw. błędy statystyczne lub przypadkowe (losowe). Są one na ogół "nieszkodliwe", ponieważ można je łatwo wyznaczyć z wielu obserwacji, a poza tym można je dowolnie zmniejszać zwiększając odpowiednio liczbę obserwacji. Dziedzina tzw. błędów systematycznych uchodzi za bardziej niebezpieczną. Ponieważ "wchodzą" one do pomiaru nie przypadkowo, ale właśnie systematycznie, trudniej jest je oszacować, a poza tym mają tę nieprzyjemną właściwość, że nie dają się zmniejszać wskutek zwiększania liczby powtarzania pomiarów. Takie spojrzenie na to zagadnienie wydawało się dosyć jasne, proste i łatwo można było zilustrować je kilkoma przekonującymi przykładami, które nadały temu "pieczęć" - wydawało się - nieodpartej rzeczywistości. Oczywiście, pozostawały niejasne jeszcze pewne zagadnienia szczegółowe i widoczne były pewne "skazy", ale wydało się, że są one o mniejszym znaczeniu.

Tablica 1

Tradycyjny podział typów błędów pomiaru
i ich właściwości

a) Błędy przypadkowe:	b) Błędy systematyczne:
<ul style="list-style-type: none"> - mogą być wyznaczane przez powtarzane (liczba n) pomiary; - rozkład jest znany (rozkład normalny); - błąd maleje proporcjonalnie do czynnika $1/\sqrt{n}$; - prawo propagacji błędów jest znane (dodawanie kwadratowe). 	<ul style="list-style-type: none"> - mogą być zazwyczaj oszacowywane tylko z grubsza (np. wartości graniczne 1); - rozkład jest na ogół nieznan; - błąd nie maleje przy zwiększaniu liczby pomiarów n; - propagacja błędów nie jest jasna; na ogół dodaje się je w sposób liniowy.
<p><u>Wniosek:</u> Oba typy błędów powinno się traktować oddzielnie i jeśli to możliwe - nie składać.</p>	

Tak więc było to dość niespodziewane, że oba typy błędów dały powód do ich zupełnie różnego zachowania się pod względem matematycznym. Według mnie mało przekonujące jest na ogół przyjmowane uzasadnienie dotyczące liniowej propagacji błędów systematycznych ze względu na konieczność zachowania w ten sposób parametru wartości maksymalnej. Również stosowany niekiedy przepis, aby do trzech lub czterech składników dodawać je liniowo, a dalej kwadratowo, było pewną konstrukcją myślową, której bezradność była jeszcze bardziej oczywista. Abstrahując jednak od tych problemów (w praktyce często nieistotnych), wydawało się, że całość jest dość rozsądna. Taki podział był wprawdzie niefortunny, ale jednak nieuchronny, ponieważ odpowiadał w sposób widoczny obiektywnemu stanowi rzeczy. W przypadku, gdy - przykładowo - składniki jednego typu były wyrażane w postaci odchyleń standardowych, a składniki drugiego typu jako wartości maksymalne, wydawało się również zrozumiałe, że nie dodaje się

ich w pewien sposób, lecz wyszczególnia oddzielnie. W wielu praktycznych przypadkach było to wprawdzie nieco niewygodne, ale z braku lepszej propozycji wydawało się konieczne, aby to również akceptować.

Tak mniej więcej widziałem sytuację w tej dziedzinie, gdy wstąpiłem przed 20 laty do BIMP. Jak Państwu zapewne wiadomo, jedno z głównych zadań tej instytucji polega na okresowej organizacji pomiarów porównawczych wśród laboratoriów państwowych. Obejmują one prawie wszystkie ważniejsze dziedziny, o ile tylko są one reprezentowane w BIMP. W przydzielonym mi zakresie działalności cała sprawa była już w znacznym stopniu uzgodniona. Przy oszacowywaniu napływających wyników pomiaru pojawiały się jednak stale pewne problemy związane z niedokładnością pomiaru. Już nawet przy najbardziej elementarnym zadaniu, uśrednianiu, powstawały zasadnicze trudności, gdy chciano wyznaczać średnią ważoną: czy, na przykład, przy wagach statystycznych niepewności systematyczne powinny pozostawać nieuwzględniane? Salomonowym rozwiązaniem praktycznym przez długi czas było to, że rezygnowano z tworzenia wartości średniej - co niektórym uczestnikowi mogło się wydawać słuszne, miał on więc swobodę w wyborze dogodnego dla siebie punktu odniesienia. W rzeczywistości nawet rezygnując z ważenia nie zabezpiecza się bardziej od dowolności postępowania, może się przecież wyłonić problem wyskoków, które musimy eliminować, jeżeli średnia ma nie utracić jakiegś wiarygodności.

Mogę więc powiedzieć, że taka sytuacja, w której przyszło mi działać, niewątpliwie miała na mnie wielki wpływ. Nauczyłem się badać problemy z różnych stron i przekonałem się, że przypisywana im jest duża doza aspektu psychologicznego. Moja wątpliwość wobec "w zasadzie jedynie słusznych" rozwiązań zwiększyła się i cała dziedzina zaczynała - moim zdaniem - wykazywać dosyć skomplikowaną strukturę, gdzie do szlachetnego wina podstaw statystyczno-matematycznych zaczęto domieszać znaczną część wody, aż stanie się

ono przyjemne w smaku dla celów praktycznych. Także wnioski, które postulują formułowanie "w pełni obiektywnych" wypowiedzi, wydaje mi się nie trafną skoro chce się zadowolić faktem, że w ogóle nie do uniknięcia są oceny subiektywne. Jeżeli rozpatrzeć poza tym cel tego działania, polegającego na otrzymywaniu stosownych w praktyce ocen dokładności, to trzeba przyznać, że nie chodzi tu o abstrakcyjny problem statystyki lub logiki, lecz że nie powinny być przy tym przeoczone pewne aspekty praktyczne. To stwierdzenie nie powinno być w żadnym razie traktowane jako zgoda na matematyczne niedbalstwo; powinno ono raczej otworzyć oczy, aby widzieć cały problem w całej jego różnorodności. Próby rozwiązania wymagają więc otwartego i pragmatycznego podejścia.

4.3. Niejasna sytuacja

W takiej trudnej sytuacji nie łatwo jest przewidzieć, skąd można oczekiwać rozwiązania, a przynajmniej postępu. Czy może jakieś założenie wymaga pewnej zmiany lub może szukamy czegoś, czego w ogóle nie ma? Były już wtedy liczne opracowania, szczególnie z NBS¹⁾ (przede wszystkim Eisenharta i Youdena [4-8]), jednak wobec klasyfikacji niepewności wyrażano dziwne ociąganie i niezdecydowanie; problem ten nie znajdował w nich dużego zainteresowania. Niejednokrotnie można było słyszeć opinię, że "błędy systematyczne" są nieważne, ponieważ można je wyeliminować przez proste przesunięcie. Dla fizyków eksperymentalnych takie wnioski musiały wydawać się nieco dziwne.

Zanim będziemy kontynuować te rozważania trzeba koniecz- nie wskazać na pewne zamieszanie, które powstało z biegiem lat w dziedzinie słownictwa. Dotyczy ono w szczególności pojęcia "błędu systematycznego" i nie jest ono ograniczone tylko do języka. Nietrudno rozróżnić tu dwa główne aspekty,

1) NBS - National Bureau of Standards (Narodowe Biuro Standardów) zwane obecnie National Institute of Science and Technology - NIST.

które w praktyce nakładają się na siebie. Pierwsza część zamieszczenia jest szczególnie trywialna, ale wcale jeszcze nie wygasła, a mianowicie pomieszanie pojęć "błąd" i "poprawka". Jeżeli się wie, że pewien wpływ na mierzoną wielkość wywołuje odchylenie, którego wartość i znak jest znany, lub można ją oszacować, to można i powinno się wprowadzić poprawkę. W dalszej części rozważań zakłada się, że poprawka została już wprowadzona.

Jest to w istocie bardzo ważny punkt, ponieważ wynika stąd, że wszystkie pozostałe jeszcze niepewności pomiarowe występują z jednakowym prawdopodobieństwem jako dodatnie lub ujemne, tzn., że wartość oczekiwana jest równa zeru. Spotykane niekiedy twierdzenie, jakoby samo istnienie systematycznych niepewności pomiaru konieczne prowadziło do systematycznego odchylenia wyniku końcowego, należy więc uważać za błędne. Ten wniosek jest niezależny od drugiego zamieszczenia, uwarunkowanego tym, że przymiotnik "systematyczny" stosuje się w wielu, częściowo przeciwstawnych sobie znaczeniach. Obok znaczenia pierwotnego zastosowanego w tablicy 1, określa się niekiedy systematyczną niepewność również taką, co do której zakłada się (słusznie lub niesłusznie), że można dać tylko pewną wypowiedź w sprawie wartości maksymalnej. Chodzi więc o dwa zupełnie różne punkty wyjścia, które nie mają ze sobą prawie nic wspólnego. Że to pomieszanie pojęć w dalszym ciągu występowało i dalej występuje, wynika prawdopodobnie stąd, że w pewnych przypadkach oba założenia okazują się jednocześnie słuszne, a nie każdy postępuje zupełnie logicznie.

Dalsze, często spekulatywne definicje, polegające na (fikcyjnym) zachowaniu się niepewności w przypadku powtórzeń - niezależnie od faktu, że taki eksperyment jest nie-realizowalny - wspomina się tylko na marginesie.

Czy wymienione choroby są uleczalne? W pierwszym przypadku terapia jest bardzo prosta: wystarczy zamienić słowo "błąd" słowem "niepewność", które nie ma posmaku wykroczenia. Propozycja nie jest nowa, a nawet w znacznej mierze

realizowana, jednak trudno jest niekiedy przejść od dobitnych wyrażzeń (jak na przykład "propagacja błędów") do odpowiednich nowych form. Wreszcie metrologia prawna nie jest przygotowana do takiego kroku, ponieważ w jej specyficznym świecie wartości "prawdziwe" są znane, tak więc odchylenia od nich są rzeczywiście "błędami". Droga wyjścia z drugiego nieporozumienia wydaje się być trudniejsza. Zobaczymy jednak później, że można w ogóle zrezygnować z pojęcia "błędu systematycznego", ponieważ okaże się ono jako niepotrzebne. Na razie możemy spokojnie powierzyć te troski komisjom słowniczym. Jeżeli w pewnej dziedzinie będą jasne pojęcia, to znajdują się także i niezbędne terminy; nieodpowiednie propozycje wykruszą się same.

4.4. Pierwsze promyki nadziei

Z powodów omówionych wyżej wynika, że problem "dodawania błędów systematycznych i przypadkowych" stawiany był często w innej postaci, przy czym zadawano sobie pytanie, jak powinno się składać niepewność przypadkową (zakładając na ogół rozkład normalny) z niepewnością "systematyczną" (podanej w postaci "maksymalnie możliwej" wartości). Ten problem rozpatruje się w różny sposób. Przy jednej z pierwszych prób [9] zakładano, że "błąd" systematyczny można utożsamiać z jego wartościami granicznymi, można go więc przedstawiać za pomocą pewnego rodzaju podwójnej "funkcji delta". Taki model można oczywiście stosować do obliczeń, ale jest on tak surowy, że w najlepszym przypadku może być traktowany jako oszacowanie górnej granicy. Znacznie silniej przemawiał do mnie model Wagnera [10], w którym zakładano rozkład równomierny między granicami, co jest bardziej realistyczne. Jeżeli nawet niemożliwe jest ściśle uzasadnienie (np. przez założenie nieznanego parametru - błędu - i zastosowanie postulatu Bayesa), to całości niewiele to szkodzi. Uwzględniając niepewność granic oraz intuicję, że środkowemu obszarowi można przypisać większą wagę niż war-

tościom krańcowym, można by, na przykład, przyjąć również rozkład trójkątowy albo też jakiegokolwiek inne wyrażenie analityczne dla gęstości prawdopodobieństwa w ograniczonym obszarze, które zostało zaproponowane już w XVIII wieku [np. 11].

W dalszym ciągu widzę istotną wartość tych prac, a w szczególności prac prof. Wagnera, może mniej w szczegółach budowy, co do której można mieć różne zdanie, a raczej w tym prostym fakcie, że można wtedy składać niepewności różnych typów bez zastanowienia. Takie postępowanie nie dawało wtedy autorowi żadnych podstaw do wątpliwości; ich zainteresowania koncentrowały się wówczas na praktycznej metodzie postępowania. Trzeba również podkreślić, że od początku było jasne, że należy przy tym zawsze dodawać wariancje [12].

Gdyby chciało się poznać wynikającą stąd funkcję gęstości, to można ją łatwo wyznaczyć, jeżeli jakaś składowa błędu przybierze postać prostokąta. Ponieważ każdą funkcję prostokątną można uważać za różnicę dwóch funkcji skokowych, to można uzyskać np. dla splotu z rozkładem normalnym jako wynikową gęstość różnicę dwóch całek gaussowskich. Już w [10] uznano to za słuszne. Rachunki przeprowadzone później przez Dietricha w jego grubej książce [10] świadczą o uwzględnieniu przestrogi Goethego, zawartej w jego "Dywanie zachodu i wschodu" - "Płaski jest kwargel, kiedy go się zgniecie, ale nie silny!"¹⁾

4.5. Rola "Code of Practice"²⁾

Nie wiem, czy z wami jest podobnie: już niejednokrotnie zdarzyło mi się, że bywałem silnie pobudzany przez te publikacje, które wykazywały oczywiste braki. Jeżeli natomiast

1) Tłumaczenie Jana Sztaudyngera - por. "Dywan zachodu i wschodu". PIW, 1963, oprac. Anna Miłska. Cytowany wyjątek znajduje się w "Księdze Przysłów".

2) Autor ma tu na myśli pozycję bibliograficzną [14].

mam do czynienia z pracą, która prezentuje pewną dziedzinę po mistrzowsku i z rozważą, to wywołuje wprawdzie podziw, ale jednocześnie stawia przed oczyma tę gorzką prawdę, że czegoś takiego nie moglibyśmy sami wykonać. Czujemy się wtedy jak ten amator, który po recitalu wielkiego pianisty wraca głęboko poruszony do domu i przez długi czas nie pozwala się zasiąść do fortepianu. W przeciwieństwie do tego prace, które dają się łatwo skrytykować, pobudzają nas do tego, aby to zrobić lepiej. Z tego względu jestem bardzo wdzięczny autorom pracy pt. "Code of Practice" [14] za to, że opublikowali dzieło, których usterki mogły być im nieobce.

Już w 1972 roku zetknąłem się z jego wstępną odbitką, jednak nie mogłem jeszcze wtedy dać wyrazu różnym mieszanym uczuciom, które mnie ogarniały przy jej czytaniu. Mój wpływ ograniczał się do przykładu liczbowego: stosowane pierwotnie dane Rutheforda i Geigera o cząstkach alfa zostały zastąpione pomiarami NPL cząstek beta (str. 35), ponieważ wykazałem właśnie [15], że cząstki alfa mają rozkład Poissona, który jest wyraźnie zniekształcony.

Znacznie ważniejsze było dla mnie rozpoznanie, które po wielokrotnym czytaniu stawało się dla mnie coraz bardziej oczywiste, iż wymaganie ścisłego podziału niepewności pomiaru na klasy, a mianowicie na "przypadkowe" i "systematyczne", które określają podział, wydaje mi się niejasny i sprzeczny. Całe zło zaczyna się już na pierwszej stronie; gdzie między innymi wymaga się zupełnie bezpodstawnie, aby przypadkowa niepewność pomiaru miała "rozkład normalny lub jakiś inny odpowiedni rozkład". Powiedziane jest tam również, że należy przy niepewności systematycznej oczekiwać, że "wpływa ona na wynik pomiaru lub go przesuwają. Jeżeli wynik eksperymentu będzie wykorzystywany dalej, to - jak jesteśmy pouczani na następnej stronie - jego "błąd" przypadkowy w kolejnym wyniku powinien być znowu sklasyfikowany jako składnik przypadkowy, chociaż jest to "nieco nielogiczne", jak brzmi to demaskująco naiwnie jednak takie postępo-

wanie jest "korzystniejsze", ponieważ prowadzi do "bardziej konsekwentnego sposobu traktowania niepewności pomiaru". W rzeczywistości jest to jedynie nieporadną próbą dopasowania rzeczywistości do z góry wybranego schematu, polegającego mianowicie na tym, że obie kategorie niepewności pomiaru rozdziela się ściśle od siebie. Rzeczywistość powinna się poddać tej zasadzie, w razie potrzeby - choćby i gwałtem. Nie dziwi nas jednak, że przy takiej grze wykorzystuje się wybrane niejasne definicje.

4.6. Czy rzeczywiście potrzebne są klasy błędów?

Po tych doświadczeniach staje się jasne, że moja dotychczasowa wiara w sens i konieczność tradycyjnej klasyfikacji niepewności pomiaru nieco się zachwieła. Również bezsporny fakt, że typ niepewności może się zmieniać w zależności od przykładu zastosowania, odbierał kredyt zaufania przepisom "Code of Practice". Po co w ogóle wydaje się klasyfikacje? Przede wszystkim po to, aby wprowadzić pewien porządek do danej dziedziny. Tak było w przypadku Carla von Linné i właśnie wtedy żywiono nadzieję (przynajmniej skrycie), że takiemu schematowi można przypisać również głębszy sens. W takich dziedzinach jak matematyka lub fizyka takie podejście jest czysto pragmatyczne: wybrany podział ma być pożyteczny, a nawet konieczny. Wydaje się, że to kryterium w przypadku niepewności pomiaru zostało rzeczywiście spełnione, okazało się przecież, że koniecznie trzeba rozpatrywać w różny sposób składniki przypadkowe i systematyczne. Okazuje się więc chyba, że gdzieś tu potrzebny jest wyraźny podział. Z drugiej jednak strony można stwierdzić empirycznie dziwną przemianę typów. Tak więc powstała paradoksalna sytuacja, że już nie wiadomo było, czy polowano na kameleona, czy też chodzi tylko o spór słowny. Rozwiązanie tej zagadki przyszło do mnie w sposób raczej niespodziewany i to bez wysiłku, a mianowicie na podstawie prostego przykładu. Ponieważ był on już dwa razy szczegółowo

omawiany (np. w [16] lub [17]), powiem tu tylko to, co wydaje się niezbędne. Załóżmy, że mamy n wyników pomiarowych x_j , które ze względu na wspólny efekt wymagają poprawki c , tak że poprawione wartości wynoszą $y_j = x_j + c$. Jako szukaną "funkcję f " obieramy wartość średnią:

$$\bar{y} = \frac{1}{n} \sum_j y_j = \frac{1}{n} \sum_j x_j + c.$$

Poprawka c jest obarczona niepewnością s_c . Teraz zaczyna interweniować ogólne prawo propagacji niepewności, które jest - jak miemam - dobrze znane. Zawiera ono pochodne cząstkowe względem kolejno występujących zmiennych oraz ich wariancje i kowariancje. W naszym przykładzie możliwe są dwa proste, ale w zasadzie różne zastosowania tego ogólnego przepisu. W pierwszym traktuje się jako zmienne obserwacje x_j oraz "efekt" c . Ponieważ te wielkości są od siebie niezależne, to kowariancje odpadają i uzyskujemy od razu dla wariancji \bar{y} wyrażenie:

$$s_{\bar{y}}^2 = \frac{1}{n} s_x^2 + s_c^2.$$

W drugim jako wielkości wejściowe stosujemy poprawione wyniki pomiarów y_j . One jednak wskutek wspólnej poprawki c są ze sobą skorelowane, przy czym można wykazać, że $\text{Cov}(y_j, y_k) = s_c^2$. Tak więc również i w tym przypadku zastosowanie prawa propagacji do wariancji \bar{y} prowadzi do wyżej podanego wyniku. Ten wynik jest jednak wyjątkowo godny uwagi, ponieważ wskazuje nam, że teraz w zupełnie oczywisty sposób wynika to, co dotychczas można było uzyskać tylko w sztuczny sposób i przy uwzględnieniu "typów błędów". Składniki całkowitej wariancji uważane dotychczas jako "systematyczne" i "przypadkowe" pojawiają

się tu z oczekiwanymi wartościami współczynników ($1/n$ lub 1). Poza tym zrozumiałe jest teraz dodawanie i staje się oczywiste, że jako wartości s_c nie należy wstawiać maksymalnej granicy, lecz wartość, która odpowiada odchyleniu standardowemu. Po wykonaniu tych prostych rachunków (praca oryginalna ma datę 17.10.1973) stało się dla mnie od razu jasne, że wszystkie prowadzone dotychczas dyskusje na temat "klas błędów" obróciły się w niwecz. Tego rodzaju podział jest mianowicie tylko wtedy potrzebny, jeżeli chce się uniknąć kowariancji. Ponieważ jednak w ogólnym przypadku nie możemy nic zmienić w odniesieniu do korelacji zmiennych, a prawo propagacji błędów w dzisiejszej jego uogólnionej postaci uwzględnia właśnie ten przypadek, wobec tego najprościej jest również to stosować. Zresztą łatwo przekonać się, że prostota wybranego przykładu nie ogranicza ogólności postępowania. Ponieważ obie dotychczasowe klasy odpowiadają tylko przypadkom granicznym braku korelacji lub pełnej korelacji, wobec tego zastosowanie kowariancji stanowi prawdziwe i konieczne uogólnienie.

Tym samym wydawałoby się, że usunięta jest główna przeszkoda w uniformizacji niepewności pomiaru, dla której było możliwe pewnego rodzaju ujęcie bez podziału na klasy. Jednocześnie ujawniły się prawdziwe znaczenie i praktyczna przydatność korelacji; tak więc skończył się okres pomijania kowariancji.

4.7. Na drodze do zalecenia BIPM

Można by już przypuszczać, że moja historia zbliża się ku końcowi, jednak tak nie jest. Jeżeli chodzi o indywidualne pojmowanie w tej dziedzinie, to nie osiągnięto właściwie niczego; jeżeli nie uda się zaznajomić z tym wszystkim i wprowadzić międzynarodowo, to wszystko będzie bezużyteczne. Nastąpił potem dłuższy okres wewnętrznych dyskusji w łonie BIPM. Pełnym zrozumienia doradcą był dla mnie w tej sprawie prof. Allisy, a jego rady, oparte na

bogatym doświadczeniu, okazały się niezwykle wartościowe. Nasz dyrektor, dr Terrien, orientował się bardzo dobrze, że sprawy są dyskusyjne, jednak umiał on skierować uwagę swego następcy, prof. Giacomo, na te zagadnienia, a jego żywe zainteresowania tymi problemami nie osłabło do dzisiaj. Zagadnienia związane z błędami pomiaru dyskutowane są w rozmaity sposób na posiedzeniach sekcji CCEMRI¹⁾. Wreszcie z inicjatywy dr Amblera (z NBS) postawiono ten problem w jego ogólnej postaci na posiedzeniu CIPM we wrześniu 1977 roku. Wtedy to polecono BIPM nawiązać kontakt z laboratoriami państwowymi - jeszcze przed utworzeniem grupy roboczej - i wyjaśnić przez ankietowanie, jakie specjalne problemy mogły by się rzeczywiście nadawać jako tematy do dyskusji. W tej sytuacji wstrzymano się w ogóle od publikowania czegokolwiek, co mogłoby być poczytywane za mieszanie się BIPM w proces badania opinii.

Dalszy rozwój tej sprawy powinien być już powszechnie znany i dlatego będzie przedstawiony tylko w skróconej formie. Próbowaliśmy zainicjować dyskusję na temat istotnych otwartych problemów, stawiając przy tym osiem odpowiednich pytań. Rozesłaną na cały świat ankietę. "Kwestionariusz na temat niepewności" zaopatrzone dłuższym wstępem: w podobny sposób postępowano z poszczególnymi pytaniami. Mają one np. następujące brzmienie: "Czy istnieje istotna różnica między błędami przypadkowymi a systematycznymi?" albo "Czy należy zalecać praktyczną regułę: a) wyrażania błędów systematycznych, b) łączenia błędów przypadkowych i systematycznych i c) wyrażania niepewności wynikowej?". Nadesłane od 21 laboratoriów odpowiedzi zostały starannie przeanalizowane [19]. Wynikało stąd, że jesteśmy bardzo daleko od jednolitego poglądu. Mimo to - a może właśnie dlatego - okazało się, że ankietę była bardzo cenna i bez wątpienia wywołała w całym świecie

1) Comité Consultatif pour les étalons de mesure des rayonnements ionisants - Komitet Doradczy Etalonów Jednostek Miar Promieniowania Jonizującego.

proces pogłębionej analizy tych problemów, które uważane były jako od dawna rozwiązane. Jednocześnie okazało się, że z wielu miejsc strzymano rozsądne propozycje. W ten sposób wytworzyła się korzystna atmosfera dla spotkania ekspertów. Nadszedł również moment, aby opublikować uzyskane wcześniej materiały [16], ponieważ mogły one teraz służyć jako podstawa do ujednoczenia poglądów.

W dodatkach do [19] opublikowano cztery odpowiedzi w pełnej wersji. Wśród nich również odpowiedź PTB, w której zajęto się przede wszystkim zagadnieniami terminologii. Warto wspomnieć jako kuriozum, że z całą powagą proponowano rozróżniać nie dwa, ale sześć rodzajów błędów, co zostało szczegółowo uzasadnione. Co by się miało z tym dalej robić, wszakże nie powiedziano.

Następne spotkanie w Sèvres w październiku 1980 roku, w którym wzięli udział specjaliści z ośmiu krajów, i które prowadził prof. Giacomo [20], doprowadziło do znanego zalecenia BIPM, którego nie potrzebuję Państwu tu przedstawiać w szczegółach, ponieważ jego treść jest ogólnie znana. Już to było właściwie dość dziwne, że spotkanie doprowadziło do jakiegoś namacalnego rezultatu; również niezwykłą rzeczą jest to, że to zalecenie spotkało się z pozytywnym przyjęciem. Niekiedy pytam się samego siebie, czy w sposób niezamierzony nie przyczyniła się do tego również stosunkowo skromna i mało pretensjonalna postać tego zalecenia, bo przecież okazało się wielokrotnie w późniejszych dyskusjach, że właśnie najważniejsza sprawa bywała często błędnie rozumiana. I tak na przykład jeszcze dzisiaj wielu ludzi uważa, że "typ A" i "typ B" stanowią tylko nowe nazwy na "przypadkowy" i "systematyczny", co należy w sposób oczywisty uważać za absolutnie mylną interpretację. Zresztą utworzenie przez grupę roboczą tych nowych typów nie było najszcześniejszym pomysłem, ponieważ u powierzchownego czytelnika mogło powstać fałszywe wrażenie nowej dychotomii. W dalszym ciągu trudno jest zrozumieć, że obecnemu podziałowi przypada raczej funkcja pedagogiczna i terapeutyczna, która w końcu jest zbędna.

Aby jeszcze na krótko powrócić do postawionego na początku pytania w sprawie przyczyny zadziwiającego sukcesu, można by znaleźć teraz prostą odpowiedź: było nią początkowo prawie heretycko brzmiące twierdzenie, że tradycyjne "klasy błędów" są nie tylko niepotrzebne, ale i szkodliwe, ponieważ cała ta dziedzina zeszła na całe dziesięciolecia na manowce. Dopiero po odrzuceniu tego balastu można było odetchnąć. Tym samym kończy się historia, którą chciałem Państwu opowiedzieć. Może będą i inne historie, które pokażą ten okres z innej perspektywy i przy okazji dadzą konieczne uzupełnienia.

4.8. Obecna sytuacja i pewne problemy aktualne

Zalecenie BIPM jest w sposób zamierzony krótkie; ogranicza się ono do rzeczy najbardziej istotnych i nigdy nie było pomyślane jako instrukcja praktyczna. Jeżeli próbowaloby się je zastosować w praktyce, to można by napotkać wiele trudności, których tam nawet nie zaznaczono. Odtąd rozwiązanie całego kompleksu zagadnień podjęła się Komisja ISO (TAG-4/WG-3). Jej główne zadanie polega na tym, aby opracować pewien "Guidance Document", który ma pomóc w posługiwaniu się tym zaleceniem BIPM w jak najszerszym obszarze zastosowań. Znajdzie się w nim w szczególności instrukcja co do liczbowego oszacowywania niepewności typu B, jeżeli trzeba będzie np. przyjąć jako punkt wyjścia maksymalne granice. Ten istotny dokument jest obecnie dopiero w opracowaniu, jednak wybór dr Collé (z NBS) do tego delikatnego zadania pozwala mieć nadzieję na właściwe załatwienie sprawy i na praktycznie użyteczny rezultat.

Można by prawie mniemać, że znajdujemy się w stanie kompletnie "bezproblemowym". Tak jednak nie jest. Wprawdzie nowy system jest w ogólnych zarysach jasny, ale w pewnych szczegółach jest jeszcze "otwarty". Przede wszystkim jest on jeszcze praktycznie mało wypróbowany (poza płaszczyznę międzynarodową); na razie opisano szczegółowo tylko

nieliczne rzeczywiste i przekonywujące przykłady zastosowań [21]. Od niniejszego seminarium można by właśnie oczekiwać szczególnie w tym kierunku dużego postępu. Ważny jest także dalszy rozwój formalizmu matematycznego; wymaga on jeszcze wiele solidnej pracy [22]. Dzisiaj jest już ona w znacznej mierze wykonana i PTB wzięło w tym wiodący udział [3]. Przy każdym zastosowaniu trzeba jednak założyć, że sprawy podstawowe są uzgodnione. Dlatego też uznałem za warte zachodu, aby powtórnie (w sprawie pierwszej próby - por. [23]) umocnić tę podstawę poprzez elementarną dyskusję - mając również nadzieję, że w ten sposób uda się eliminować ewentualne przeszkody, które niejednemu zainteresowanemu utrudniałyby dalszy postęp.

Było by niehonorowo przemilczać, że są jeszcze zdeklarowani przeciwnicy tej doktryny BIPM. Głównym powodem opozycji jest ich przekonanie, że nie można dokonywać obróbki "błędów systematycznych" za pomocą ustalonych reguł propagacyjnych, ponieważ można je stosować tylko do składników przypadkowych. "Randomizacja" składników systematycznych stanowi również zabieg niedozwolony; możliwe i dozwolone jest tylko podawanie obu zasadniczo różnych składników - wbrew wszystkim praktycznym wymaganiom na jedną wartość liczbową (np. w [24] lub [25]). Nie odczuwamy ochoty uczestniczenia w starym sporze, który "okopuje" się niekiedy za barykadami filozoficznymi i nabiera wtedy nieco scholastyczno - średniowiecznego posmaku. Powyżej już zrozumieliśmy, że argument o odchyleniu systematycznym jest fałszywy, wartości oczekiwane wszystkich składników są przecież równe zeru. Nie trzeba również nigdzie dokonywać "randomizacji", ponieważ nasza niewiedza dotycząca wielkości i znaku czyni to zupełnie zbędnym. Tak więc jako jedyną cechą rozróżniająca pozostaje dokładność, z jaką oceniamy - jak mniemamy - wartość liczbową, jednak tak subiektywny element jest nieprzydatny do klasyfikowania. Nie można podać żadnych obiektywnych powodów do twierdzenia, że niepewności systematyczne i przypadkowe są za-

sadniczo różnego rodzaju. W ten sposób wyżej przytoczone argumenty okazują się psychologicznymi przesadami, na które nie ma żadnej rady przy użyciu racjonalnych przesłanek. W ostatnim czasie jednak upewniamy się w poglądzie, że chodzi tu raczej o potyczki pojedynczych bojowników znajdujących się w odwrocie, którzy wiedzą, że wojna jest już rozstrzygnięta. Dlatego też najlepiej teraz myśleć o współżyciu po zawarciu traktatu pokojowego. Prawdopodobnie za kilka dziesiątków lat nie będzie już można dobrze zrozumieć, co było właściwie istotą sporu.

Należy obstawiać przy tym, że wszystkie dotychczas poznane "alternatywne" propozycje obliczeń wykazują tak duże wady, że praktycznie są bezużyteczne; w przypadku gdy wielkości wejściowe są skorelowane, zwykle zawodzą, ponieważ się tego nie przewidywało. Również procedura nie jest na ogół konsekwentna, tzn. uzyskane wielkości wyjściowe nie dają się zastosować jako nowe wielkości wejściowe. W przypadkach, gdy procedura daje dwie dane liczbowe dotyczące niepewności, jest co najmniej niejasne, co ma się z tym stać. Musi to też dziwić, że te propozycje wynikają często z jakiegoś zbyt wąskiego celu stosowania, w którym niepewności przydziela się tylko takie zadanie, aby nas uchronić od sytuacji, w której jakiś zaproponowany zakres wartości mógłby się później okazać zbyt wąski. Oczywiście jest wiele przypadków, w których taka funkcja koła ratunkowego ma swoje znaczenie; oprócz tego można jednak przytoczyć również ważne, zupełnie inne przykłady, jak choćby wyszukiwanie małych, ale istotnych różnic, które przy zbyt szeroko zakrojonych danych dotyczących błędów pozostają niewykryte. Tak więc ogólnie pożyteczne dane dotyczące niepewności powinny być przynajmniej w tym sensie "neutralne", aby nie wykluczały a priori żadnych ewentualnych zastosowań.

W zakończeniu poruszymy krótko pewne jeszcze aktualne problemy. Ważne zagadnienie ogólnej natury, które nie było dotychczas rozważane, dotyczy tzw. przedziału ufności [26].

Powinien on mieć tę właściwość, aby przyporządkować poszukiwanej, jednak nieznaną, prawdziwej wartości pewną zadaną wartość prawdopodobieństwa. Takie wymagania podnoszone są zwłaszcza ze strony organizacji prawnych, które nie zdają sobie z tego sprawy, że takie wymagania nie mogą być ściśle spełnione. W każdym razie byłoby do tego potrzebne pewne założenia dodatkowe, co do prawomocności których można mieć różne poglądy. Poza tym nie można przeoczyć, że dla przedziałów ufności (w przeciwieństwie do odchyleń standardowych) nie ma ogólnego i praktycznie użytecznego prawa propagacji.

Inny krąg zagadnień daje się opisać za pomocą hasła "wariancja Allana". Ponieważ do jej tworzenia z czasowego ciągu danych pomiarowych stosowane są tylko wartości sąsiednie, jest ona w znacznej mierze niezależna od dowolnego, ogólnego przesunięcia. Ta właściwość jest szczególnie pożądana w dziedzinie pomiarów częstotliwości, jednak powstają poważne wątpliwości wobec wymagania, aby stosować taką wielkość w ogóle w metrologii [27]. Wskutek jej szczególnej definicji nowa wariancja nie ma nic wspólnego ze stabilnością wartości średniej. Można się więc narazić na niebezpieczeństwo, że tworzy się wielkość, o której nie można właściwie powiedzieć, co praktycznie oznacza. Poza tym przekonano się również, że nieuzasadnione jest twierdzenie o nieuniknionym szumie $1/f$ - na przykład przy rozpadzie alfa [28].

Ewentualne znaczenie innych nowych kierunków badań w dziedzinie błędów pomiaru jest jeszcze nie zupełnie jasne. Chodzi tu o próbę zastosowania do ich charakterystyk tzw. teorii zbiorów rozmytych [29]. Choć wydaje się, że nie doprowadzi to do jakichś praktycznie istotnych zastosowań, to jednak warto mieć na uwadze ten sposób podejścia.

4.9. Rzut oka w przyszłość

Nie wazę się prorokować na temat przyszłego rozwoju tej dziedziny. To że nowa koncepcja przeżyła już okrągłe dziesięć lat, można ocenić jako sukces, wcześniejsze analogiczne próby popadały często w zapomnienie po krótszym czasie. BIPM nie miało zamiaru uczynienia sobie z tego nigdy sprawy prestiżowej. Wyraża ono raczej pogląd, że do jego zadań należy wspieranie tego rodzaju prac rozwojowych tak długo, jak długo uzna to za potrzebne znaczna większość państwowych laboratoriów. Centralne stanowisko BIPM ułatwia mu taką próbę, jak również pewne doświadczenie, które zostało zdobyte przy podobnych wcześniejszych usiłowaniach unifikacyjnych; nie zamierza ono jednak w żaden sposób odgrywać centralnej roli dłużej niż jest to potrzebne.

Nikt nie stwarza sobie iluzji, że zharmonizowanie rozbieżnych poglądów stanowi łatwe przedsięwzięcie. W najlepszym razie można przyjąć, że rozpoczął się długotrwały proces, którego promocja i rozwój będą wymagały jeszcze wiele wytrzymałości i cierpliwości, ponieważ każda próba zmiany starych przyzwyczajzeń napotyka zrozumiały opór. Jednak nadzieja, że z tego może w końcu coś powstać, co będzie jaśniejsze, bardziej użyteczne i ogólniejsze, a więc stanowiące rzeczywisty postęp, czyni tę próbę wartą wysiłku.

Jesteśmy w środkowym etapie tego procesu, a nasze seminarium stanowi pewien krok na tej drodze. W związku z tym wydaje mi się stosowne wskazać na pewien list Beethovena do jednego ze swych wydawców [30], z którymi miewał on często wiele kłopotu. Wskazuje on nam mianowicie, że gwałtowna niecierpliwość i wzajemne zrozumienie nie muszą się wykluczać. Beethoven pisze w nim najpierw z oburzeniem: "Błędy - błędy! Pan sam jest jednym błędem! Tak więc muszę tam wysłać swego kopistę albo muszę chyba sam tam się udać,

jeśli chcę, aby moje dzieła nie ukazywały się jako same błędy. Trybunał muzyczny w Lipsku - jak się zdaje - nie wyda żadnego porządnego korektora; proszę odesłać dzieła zanim jeszcze otrzyma Pan korektę".

Jeszcze w tym samym liście, znajdujemy jednak niespodziewany, ale charakterystyczny dla Beethovena zwrot (wskutek małej gry słów ¹⁾) w kierunku pojednawczym, gdyż czytamy tam:

"Niech Pan błądzi, ile Pan chce, niech Pan sobie niczego nie żałuje. Zasługuje Pan u mnie na szacunek; tak to już jest u ludzi, że się ich za to ceni, że nie popełnili większych błędów.

Pański oddany sługa Beethoven".

Tak więc żywię niepłoną nadzieję, że i w tym spotkaniu, gdy już zderzą się najrozmaitsze poglądy, w końcu dojdzie, jak u Beethovena, mimo wszystko do harmonijnego pogodzenia i wzajemnego zrozumienia.

BIBLIOGRAFIA

1. Braker U.: W: Etwas über William Shakespeares Schauspiele von einem armen ungelehrten weltbürger, der das glück genoss, denselben zu lesen. Piękne nowoczesne wydanie ukazało się w Insel-Bucherei. No 435, Insel, Leipzig, 1964. Wybrany cytat znajduje się w "Sturm".
2. PTB-Bericht ATWD-17: Beiträge zum Problem der Angabe von Messunsicherheiten. Seminarium PTB, wrzesień, 1980, wraz z przyczynkami opracowanymi przez W. Wögele, K. Weisego i W. Mannharta, 1981.
3. Grundbegriffe der Messtechnik; Behandlung von Unsicherheiten bei der Auswertung von Messungen. Niemiecka norma DIN 1319, część 4, Beuth, Berlin, 1985.

1) Niestety nieprzetłumaczalnej na język polski.

4. Cali J.P., Marsh K.N.: An annotated bibliography on accuracy in measurement. *Poure and Appl. Chem.*, 55, 1983, 907-930. Dołączone komentarze są często warte przeczytania. Przytacza 14 prac W.J. Youdena.
5. Eisenhart Ch.: Some canons of sound experimentation. *Bull. Inst. int. de stat.*, 37, 1960.
6. Id: Realistic evaluation of the precision and accuracy of instrument calibration systems. *J. of Res. NBS*, 67C, 1963, 161-187.
7. Id: Expression of the uncertainties of final results. *Science*, 160, 1969, 1201-1204.
8. Ku H.H.: Notes on the use of propagation of error formulas. *J. of Res. NBS*, 70C, 1966, 263-273.
9. Weichselberger K.: Über die Addition zufälliger und systematischer Fehler. *Operations Res. Verfahren*, 5, 1968, 423-444.
10. Wagner S.: Zur Behandlung systematischer Fehler bei der Angabe von Messunsicherheiten. *PTB-Mitteilungen*, 69, 1969, 343-347.
11. Eisenhart Ch.: Laws of error. W: *Encyclopedia of Statistical Science*, vol. 4, Wiley, New York, 1983, 530-566, specjalnie rys. 1.
12. Wagner S.R.: On the quantitative characterization of the uncertainty of experimental results in metrology. *PTB-Mitteilungen*, 89, 1979, 83-89.
13. Dietrich C.F.: Uncertainty, Calibration and Probability. *The Statistics of Scientific and Industrial Measurement*. Hilger London, 1973, rozdział 4.
14. Campion P.J., Burns J.E., Williams A.: *A Code of Practice for the detailed Statement of Accuracy*. NPL, H.M. Stationary Office, London, 1973.

15. Müller J.W.: A general test detecting dead-time distortions in a Poisson process. Rapport BIPM-72/10, 1972.
16. Id.: Some second thoughts on error statements. Nucl. Instr. and Meth., 163, 1979, 241-251. Praktycznie identyczny wstępny przedruk ukazał się w 1978 r., jako Rapport BIPM-78/4.
17. Id.: The assignment of uncertainties to the results of experimental measurements. W Precision Measurement and Fundamental Constants II. NBS Special Publ., 617, 1984, 375-381.
18. Mannhart W.: A small guide to generating covariances of experimental data. PTB-Bericht FMRB-84, 1981.
19. The BIPM enquiry on error statements. Rapport BIPM-80/3, 1980, 50 str. Francuska wersja w: Procès-Verbaux Séances Com. Int. Poids Mes., 48, 1980, C1-C36.
20. Report on the meeting of the BIPM Working Group on the Statement of Uncertainties (R. Karls). Francuska wersja w: Procès-Verbaux Séances Com. Int. Poids Mes., 49, 1981, A1-A12.
21. Allisy A.: Les incertitudes des mesures. Applications pratiques. Bulletin BNM, No, 53, 1983, 5-15.
22. Id.: Some statistical methods used in metrology. W: Metrology and Fundamental Constants. Wydane przez A. Ferro Milone i P. Giacomo, North Holland, Amsterdam, 1980, 20-37.
23. Müller J.W.: Ruminations on uncertainties - questions and answers. W: 8 Międzynarodowym Sympozjum Metrologii INSYMET 86 w Bratysławie, 1986.
24. Colclough A.R.: Two theories of experimental error. J. of Res. NBS, 92, 1987, 167-185.

25. Grabe M.: Principles of metrological statistics. Metrologia, 23, 1987, 213-219.
26. Wöger W.: Probability assignment to systematic deviations by the principle of maximum entropy. IEEE Trans. Instr. Meas., IM-36, 1987, 655-658.
27. Allan D.W.: Should the classical variance be used as a basic measure in standards metrology? IEEE Trans. Instr. Meas., IM-36, 1987, 646-654.
28. Nieuwenhuizen Th. M., Frenkel D., van Kampen N.G.: Objections to Handel's quantum theory of 1/f noise. Phys. Rev., A 35, 1987, 2750-2753.
29. Hasche K.: Zur quantitativen Beschreibung der Unsicherheit eines Massanschlussergebnisses. ASMW Metrologische Abhandlungen, 5, 1985, 295-316 i 6, 1986, 1-27.
30. List do Breitkopfa i Härtla w Lipsku z 6 maja 1811 r. Nosi on nr 306 w (ang.) pełnym wydaniu E. Andersona: The Letters of Beethoven. 3 tomy, New York, 1961; autograf znajduje się w Domu Beethovena w Bonn (zbiór H.C. Bodmera).

Sèvres, wrzesień 1987

DODATEK: OCENA NIEPEWNOŚCI¹⁾

A. Allisy i J.W. Müller

Zgodnie ze zwyczajem, niepewności pomiaru dzieli się na dwie klasy - nazywane jako przypadkowe i systematyczne. Ten tradycyjny punkt widzenia nie ma jednak obiektywnego

1) Por. Absorbed Dose Determination in Photon and Electron Beams. An International Code of Practice, IAEA Technical Reports Series No. 277, IAEA, Vienna, 1987, 73-77.

uzasadnienia [1]. Tak więc ponieważ niepewność "systematyczna" jest zazwyczaj traktowana jako składnik, który nie może być zmniejszony przez zwiększanie liczby pomiarów, do których wchodzi ten składnik, wobec tego stanowi to pojęcie zależne od sposobu, w jaki pewien wynik (wraz z jego niepewnością) jest dalej wykorzystywany. W rezultacie typ niepewności może się zmieniać (od przypadkowego do systematycznego lub odwrotnie) zależnie od dalszej obróbki danych. Wskazuje to na fakt, że taka klasyfikacja, chociaż niekiedy praktyczna, nie jest zbyt dobrze uzasadniona.

Poza tym można powiedzieć, że nie należy mieszać pojęcia poprawki z pojęciem niepewności; jeżeli tylko istnieje pewne przesunięcie i jego wartość i znak są znane, to należy je skorygować wprowadzając odpowiednią poprawkę. Po wprowadzeniu poprawki przesunięcie jest w zasadzie wyeliminowane, jednak jakaś niepełna wiedza o tym przesunięciu musi być uwzględniona przez wprowadzenie niepewności tej skorygowanej wartości. Obecnie należy rozróżniać między "błędem" a "niepewnością" w taki sposób, aby błąd mógł być (a więc i powinien być) skorygowany, podczas gdy dla niepewności nie można tego zrobić, ponieważ nieznanym jest jego znak. W praktyce jednak nie należy zbyt często polegać na ścisłym przestrzeganiu słownika, ponieważ jest to zawsze trochę wątpliwe.

1. KLASYFIKACJA RÓŻNYCH NIEPEWNOŚCI

Prosta i ustalona klasyfikacja różnych składników niepewności powinna być oparta na sposobie, w jaki uzyskano wartości liczbowe. Tak więc istnieje normalnie możliwość podjęcia decyzji, czy dana wartość pochodzi z pewnej procedury statystycznej (niezależnie od tego, czy prostej czy skomplikowanej), czy też opiera się w głównej mierze na opinii osobistej. To automatycznie nie implikuje, czy wyrobione wyczucie, wykorzystane przez doświadczonego

badacza, jest bezwzględnie mniej niezawodne niż wartości wynikające z pewnego rodzaju oszacowania statystycznego, lecz wskazują że są one w nieunikniony sposób subiektywne.

W zaleceniu BIPM dotyczącym niepewności [2, 3] wprowadziło się to do podziału niepewności na dwie grupy, oznaczone klasą A (oceny statystyczne, obiektywne) i klasą B (przypuszczenia, subiektywne).

W rzeczywistości ta klasyfikacja bywa często niezbyt istotna w praktyce, ponieważ oba typy są następnie traktowane w zupełnie podobny sposób. To rozróżnienie może mieć jednak pewne usprawiedliwienie natury pedagogicznej, ponieważ zmusza eksperymentatora do zastanowienia się nad źródłem jego ocen. Poza tym korzystne może być ostrzeżenie, niezależnie od tego, czy niepewności klasy B odgrywają istotną rolę wobec różnych składników, że należy się powstrzymać od precyzyjnej oceny przedziałów ufności, ponieważ nie można przypisać funkcji gęstości do wartości klasy B z jakąkolwiek wiarygodnością, co byłoby warunkiem wstępnym przy takim rachunku.

Niepewności klasy A. Ponieważ wartości liczbowe tych niepewności oparte są na pewnym wnioskowaniu statystycznym, w większości przypadków są już wyrażane w ich odchyleniach standardowych. Jako przykład możemy rozważyć wyznaczenie dawki dokonane w komorze jonizacyjnej. Należy tu uwzględnić pewien czynnik, polegający na tym, że temperatura komory podczas pomiaru może różnić się od tej, przy jakiej było dokonywane wzorcowanie. Poza tym temperatura może się zmieniać. Jest rzeczą znaną, że ten czynnik jest odwrotnie proporcjonalny do temperatury bezwzględnej, a jego niepewność jest proporcjonalna do niepewności dotyczącej wyznaczenia temperatury. Jedną drogą oceny może polegać na tym, aby dokonać wiele pomiarów temperatury i wyznaczyć odchylenie standardowe jej średniej.

Uzyska się wtedy:

$$s(\bar{T}) = \left\{ \frac{1}{N(N-1)} \sum_{i=1}^N (T_i - \bar{T})^2 \right\}^{1/2}, \quad (1)$$

gdzie T_i jest to wynik indywidualnego pomiaru, a \bar{T} jest to średnia z wykonanych N pomiarów.

Odchylenia standardowe innych istotnych wielkości można wyznaczać w analogiczny sposób; ich złożenie i skutek wywierany na wynik końcowy będzie omówiony później. W specjalnych zastosowaniach (np. przy pomiarach czasu i częstotliwości) może okazać się konieczne stosowanie nieco innej formy wariancji, zwanej wariancją Allana.

Niepewność klasy B. Nie zawsze można ocenić niepewność pomiaru przez powtarzanie pomiarów. Może np. zachodzić konieczność zastosowania wartości podanych w literaturze. Jeżeli niepewności są podane, możliwie w postaci odchyłeń standardowych, to mogą być traktowane jak składniki klasy A i tak dalej przetwarzane. Jednakże taka informacja może być niewystarczająca. Jest więc obowiązkiem eksperymentatora przydzielić im, według swojej najlepszej wiedzy, niepewności, które mogą być uważane jako oszacowane odchylenia standardowe. Nie można tego zrobić w zupełnie obiektywny sposób, ponieważ opiera się to na osobistym doświadczeniu i osądzie; mimo trudności takie niepewności powinny być oszacowywane w postaci odchyłeń standardowych (lub wariancji), ponieważ są to wielkości, które są potrzebne do dalszej propagacji, a przyporządkowanie powinno być dokonane przez eksperymentatora, ponieważ tylko on sam posiada możliwie wszechstronną wiedzę o metodach i procedurach eksperymentalnych.

Niektórzy ludzie wolą w takich przypadkach wyrażać niepewności w postaci "maksymalnych granic" rozumiejąc to w tym sensie, że jest bardzo mało prawdopodobne, aby prawdziwa wartość znajdowała się poza tymi granicami.

Co prawda nie można zaprzeczyć, że takie wartości można łatwo oszacować, ale trzeba również uznać, że one są określone wadliwie i mają ograniczoną wartość praktyczną. W szczególności nie przyczyniają się do formułowania właściwego prawa propagacji błędów. Ponieważ w większości przypadków praktycznych różne składniki przyczyniają się do całkowitej niepewności, jest rzeczą bardzo ważną, aby mieć do dyspozycji ogólne wyrażenie w związku z takim postępowaniem. Zachowując to wymaganie jesteśmy jednak nieuchronnie skazani na stosowanie wariancji (lub ich dodatniego pierwiastka, zwanego odchyleniem standardowym) jako właściwych wielkości. Wszystkie pozostałe miary trzeba będzie - przed ich użyciem - "tłumaczyć" na wariancje.

Kowariancje. W praktycznych pomiarach często występują wielkości, które nie są od siebie niezależne, albo dlatego ponieważ jeden z nich bezpośrednio wpływa na drugi, albo też dlatego, że stanowią one funkcje innej wielkości (na przykład temperatury lub wilgotności). W takich przypadkach mówimy, że są one skorelowane, a stopień korelacji może być formalnie wyrażony przez wielkość zwaną kowariancją. Eksperymentatorzy często uważają, że kowariancje są wielkościami raczej enigmatycznymi i trudno jest nimi operować, ale na szczęście o tak małym wpływie praktycznym, że prawie zawsze mogą być pomijane. Niestety, nie jest to realistyczny punkt widzenia. W sprawie podstawowych pojęć porównać należy np. Allisy [4].

Między wariancją i kowariancją istnieje dość ściśle formalne podobieństwo. Jeżeli mamy N pomiarów wielkości x i y , to jesteśmy w stanie uzyskać nie tylko wariancję, która dla pomiarów x_j przybiera postać:

$$\text{Var}(x) = \frac{1}{N-1} \sum_{j=1}^N (x_j - \bar{x})^2 \quad (2)$$

gdzie \bar{x} oznacza wartość średnią i podobnie dla y , lecz również kowariancję między tymi wielkościami, określoną następująco:

$$\text{Cov}(x, y) = \frac{1}{N-1} \sum_{j=1}^N (x_j - \bar{x})(y_j - \bar{y}) \quad (3)$$

Stąd wynika w sposób oczywisty, że $\text{Cov}(x, x) \equiv \text{Var}(x)$. Jeżeli x i y są od siebie niezależne, to kowariancja jest równa zeru. Należy zauważyć, że pojęcie kowariancji zawsze implikuje istnienie specyficznej pary wielkości (tutaj x i y).

2. PRAWO PROPAGACJI NIEPEWNOŚCI

Wielkość, którą chcemy wyznaczyć za pomocą pomiaru, rzadko można uzyskać przez pomiar pośredni. W większości przypadków żądany wynik uzyskuje się w sposób pośredni, mianowicie przez pomiar kilku wielkości, z których - przez zastosowanie pewnej zależności matematycznej - można dojść do wielkości aktualnie poszukiwanej. Tak więc, jeżeli y jest wynikiem końcowym i jeżeli x_1, x_2, \dots oznaczają wielkości mierzone, to można tę sytuację opisać wyrażeniem $y = F(x_1, x_2, \dots)$, co oznacza, że y jest (znaną) funkcją n wielkości x_1, x_2, \dots, x_n , które można bezpośrednio mierzyć (zauważ, że teraz x_j oznaczają różne zmienne).

Podstawowy wynik statystyki matematycznej gwarantuje nam, że wariancja y z dużym przybliżeniem może być wyrażona następującym równaniem:

$$s_y^2 = \sum_{j=1}^n \left(\frac{\partial F}{\partial x_j} \right)^2 s_j^2 + 2 \sum_{j < k} \left(\frac{\partial F}{\partial x_j} - \frac{\partial F}{\partial x_k} \right) s_{jk}^2, \quad (4)$$

gdzie $s_j^2 \approx \text{Var}(x_j)$ i $s_{jk} \approx \text{Cov}(x_j, x_k)$ są w praktyce estymatorami w próbie tych wielkości. Jest to ogólne prawo propagacji dla niepewności. Należy zwrócić uwagę, że nie rozróżnia się tu między składnikami niepewności klasy A lub klasy B.

Można rozróżniać dwa istotne przypadki szczególne prostych zależności funkcjonalnych. Jeżeli

$$y = F(x_1, x_2, \dots, x_n) = c_0 + \sum_{j=1}^n c_j x_j, \quad (5)$$

gdzie współczynniki c_j są stałymi, uzyskuje się wyrażenie na wariancję y :

$$s_y^2 = \sum_{j=1}^n c_j^2 s_j^2 + 2 \sum_{j < k} c_j \cdot c_k \cdot s_{jk} \quad (6)$$

Inny szczególny przypadek zachodzi wtedy gdy zmienne są czynnikami iloczynu (mnożenia lub dzielenia), a więc gdy:

$$y = F(x_1, x_2, \dots, x_n) = c_0 \prod_{j=1}^n x_j^{c_j} \quad (7)$$

W tym przypadku wygodniej jest stosować pojęcie niepewności względnej, określonej dla wielkości x_j jak następuje:

$$r_j \equiv \frac{s_j}{\bar{x}_j} \quad \text{i analogicznie} \quad r_{jk} \equiv \frac{s_{jk}}{\bar{x}_j \bar{x}_k} \quad (8)$$

dla wariancji względnej.

Po prostych przekształceniach algebraicznych uzyskuje się:

$$r_y^2 = \sum_{j=1}^n c_j^2 \cdot r_j^2 + 2 \sum_{j < k} c_j \cdot c_k \cdot r_{jk} \quad (9)$$

Te dwa szczególne przypadki wyczerpują już dużą część praktycznych sytuacji.

Właściwe zastosowanie pojęcia korelacji pozwala więc osiągnąć rzeczywistą unifikację wypowiedzi o niepewności. Tak więc ogólne możliwe przesunięcie w zbiorze pomiarów można teraz uważać jako wspólną dodatnią kowariancję, a efekt, jaki ma ona na wariancję wynikową, podany jest poprzez prawo propagacji wyrażone wyżej. W rezultacie poprzedni i często nieco dowolny podział niepewności na składowe systematyczne i przypadkowe, wraz z oddzielnymi regułami stosowania do ich propagacji, można obecnie uznać za niepotrzebną komplikację. To konsekwentne stosowanie prawa propagacji w jej ogólnej formie nie tylko pozwala na uproszczenie praktycznej oceny, ale również wyjaśnia stosowane pojęcia. Dalsze szczegóły można znaleźć u Müllera [5] lub w DIN [6].

BIBLIOGRAFIA

1. Müller J.W.: Some second thoughts on error statements. Nucl. Instrum. Methods, 163, 1979, 241-251.
2. Giacomo P.: New from the BIPM. Metrologia, 17, 1981, 69.
3. Kaarls R.: Rapport du groupe de travail sur l'expression des incertitudes au CIPM. BIPM Proc. Verb. Com. Int. Poids et Mesures 49, Annexe A, 1981.
4. Allisy A.: W: Proc. Int. School of Phys. Enrico Fermi, Course LXVIII, Metrology and Fundamental Constants (FERRO MILONE A., GIACOMO P., LESCHIUTTRA S., Eds), North-Holland, Amsterdam, 1980, 20, również Biulletin BNM, No 53, 1983, 5.
5. Müller J.W.: W: IMEKO International Training Course. Seibersdorf, 1984.
6. DIN Norm 1319, Teil 4, Grundbegriffe der Messtechnik -
- Behandlung von Unsicherheiten bei der Auswertung von Messungen, Beuth Verlag, Berlin (West), 1985.

5. ZALECENIE 1 (CI-1986) CIPM - WYRAŻANIE NIEPEWNOŚCI
W PRACACH WYKONYWANYCH POD AUSPICJAMI CIPM¹⁾

Międzynarodowy Komitet Miar,

- biorąc pod uwagę zalecenie INC-1(1980) przyjęte w 1980 r. przez Grupę Roboczą w sprawie wyrażania niepewności i Zalecenie 1 (CI-1981) przyjęte przez CIPM w 1981 r. w tej samej sprawie,
- biorąc pod uwagę, że pewni członkowie komitetów doradczych mogą sobie życzyć wyjaśnień na temat tych Zaleceń w związku z przypadającymi na nich zadaniami, w szczególności z komparacjami międzynarodowymi,
- przyjmuje do wiadomości istnienie Grupy Roboczej Międzynarodowej Organizacji Normalizacyjnej (ISO), Grupy wspólnej ISO, OIML i IEC, z którą współpracuje CIPM i która zajmuje się specjalnymi zastosowaniami wspomnianymi w paragrafie 5 Zalecenia INC-1(1980), a między innymi - zastosowaniami, które mają znaczenie przemysłowe,
- zwraca się z prośbą do wszystkich uczestników komparacji międzynarodowych i innych prac prowadzonych pod auspicjami CIPM i jego Komitetów Doradczych o przestrzeganie dyrektyw podanych w paragrafie 4 Zalecenia INC-1(1980) i o podawanie - wraz ze swoimi wynikami - niepewności łącznej ("złożonej" - composée), wynikającej ze składników typu A i typu B, w postaci odchylenia standardowego.

¹⁾ Oryginalny tekst francuski tego zalecenia 1, jak też zalecenia 2 (por. pkt. 6 poniżej) został opublikowany w "Procès-Verbaux de séances du CIPM", vol. 54, 75, sesja w 1986, Sèvres, 1987.

6. ZALECENIE 2 (CI-1986) - ANALIZA WYNIKÓW KOMPARACJI
MIĘDZYNARODOWYCH¹⁾

Międzynarodowy Komitet Miar,

- biorąc pod uwagę rosnące znaczenie jednolitości etalonów i pomiarów fizycznych dla przemysłu krajowego,
- biorąc pod uwagę, że komparacje międzynarodowe, pociągające za sobą dużą różnorodność pomiarów wykonywanych z inicjatywy Komitetów Doradczych pod odpowiedzialnością CIPM, w sposób najbardziej bezpośredni kładą szczególny nacisk na zgodność między etalonami pomiarowymi różnych laboratoriów w całym świecie,
- biorąc pod uwagę, że komparacje międzynarodowe organizowane przez Komitety Doradcze są często przedstawiane bez komentarzy,
- zaleca, aby Komitety Doradcze i ich grupy robocze zwracały w przyszłości uwagę explicite na każdy obiektywny wniosek, jaki można wyciągnąć z tych komparacji.

¹⁾ Por. odsyłacz do pkt. 5 powyżej.

ISSN 0209-1046

